A new approach in the determination of Parton Distribution Functions. Application to processes with heavy quarks in the initial states.

Maria Ubiali

.

I would like to record my thanks to all those who have helped me in the course of this work. In particular to my supervisors; to Prof. L. Del Debbio for all I have learnt from him over these four years, for his constant support and guidance. I am also very grateful to Prof. R. Ball for his help and for the illuminating discussions.

I also want to mention those who have been maestri for me since the beginning of my adventure as a PPT student in Milano, in particular Prof. G. Ridolfi and Prof. S. Forte. I owe my deepest gratitude to them for having accompanied me during these years with their ideas and support. I am deeply grateful towards the other members of the NNPDF collaboration for all that they have patiently taught me. In particular I thank Andrea Piccione and Joan Rojo for having introduced me into the world of the computer coding when I was still unable to switch on a computer. I thank Alberto Guffanti for having answered to my thousands of questions during the year spent together in Edinburgh. I finally thank Jose' Ignacio for his illuminating ideas and for the great time spent in Barcelona. To conclude, I would like to thank Francesco and Valerio for having joined the collaboration and working with enthusiasm and reliability.

In my daily work I have been blessed with a friendly and cheerful group of fellow students and members of the staff. Thanks to all the people in the Edinburgh PPT group who were and are here, especially Thomas Binoth, whose memory is a constant inspiration for me, Einan Gardi, James Zanotti, Steffen Schumann and Alberto Guffanti. Thanks to all the present and former physics PhD students, especially Ben, Tom, Thomas, Dave, Simone, Chris, Eike, Enrico, Claudia, Gavin, Hajrah, Eric, Liam, Eoin and Matthew. I spent really good time with all of you and I am grateful for having been your colleague and friend.

It has been a pleasure to spend the last year of my PhD at the Universite' Catholique de Louvain in Belgium, working with Prof. Fabio Maltoni. I thank him for having taught me so much and for his incredible enthusiasm. I thank all people in the CP3 group, in particular Prof. J. M. Gerard, Michael, Marco, Priscilla, Same', Will, Suzanne, Larissa, Olivier, Celine, Carine and Ginette. They made me feel immediately at home and have been very supportive with me.

To conclude I want to thank my husband Giacomo, for being always with me, in the hard and joyful times. Nobody describes better what I want to express than R. G. Rilke: In the experience of a great love, everything comes together; everything that happens becomes an event within that relationship. I thank my parents, my sisters and all members of my family, who are always with me, despite the distance. I also want to mention all the new friends that I have met here, for their faithful presence and for all the meals and the time spent together, in particular Joseph, Megan, Robert, Chiara, Fr. John, Lucia, Dan and many others. My heart is always with the friends whom I left in Milan four years ago, whose friendship is now even more profound than it used to be. I finally thank God who has given me everything.

These doctoral studies would not have been possible without the financial support of the Scottish Universities Physics Alliance (SUPA) and the Institut Interuniversitaire des Sciences Nucleaires (IISN).

Contents

Introduction

Hadron colliders like the Large Hadron Collider (LHC) at CERN or the Tevatron at Fermilab probe our understanding of the theory which describes the subnuclear interaction. For the past few decades, physicists have been able to describe with increasing details the fundamental particles that constitute the Universe and the interactions between them. This understanding is encapsulated in the Standard Model of particle physics, but there are still important gaps in our knownledge. The upcoming experimental data from the LHC might produce unexpected results and unveil new scenarios in our understanding of the model of elementary particles. However, the correct identification of any signal of new physics requires a careful assessment of the Standard Model backgrounds. Given that the vast majority of events are due to strong interactions, a deep understanding of the phenomenology of strong interactions is fundamental in order to fully exploit the physics potential of modern colliders.

The theory which currently describes the strong interaction is the result of the assembling of many theoretical ideas and experimental results. The search for such a theory started half-century ago. In 1963, Gell-Mann, Ne'eman and Zweig proposed a model based on symmetry principles which was able to make sense of the chaos caused by the proliferation of new hadrons produced in nuclear experiments [1, 2, 3, 4]. In a few words, they recognised that the known hadrons could be associated to some representations of the special unitary $SU(3)$ group. This lead to the concept of *quarks* as the building block of hadrons. Mesons were expected to be quark-antiquark bound states, while baryons were interpreted as bound states of three quarks. This "constituent" quark model still successfully describes most of the qualitative features of the baryon spectroscopy.

At that time it was not obvious that such a picture of strong interactions could succeed. The deeply inelastic scattering experiments carried out in the 60's, probing the inner structure of the nucleon with beams of highly energetic electrons, represented a rev-

olution in the conception of the strong force. The astonishing result was that a much larger than expected number of electrons was observed at large deflection angle. Feynman gave a simple phenomenological explanation for this result: at the small distance probed by the electrons, the nucleon has to be considered as a gas of non–interacting point–like particles, called partons [5]. The electron simply scatters elastically the collinear components of the proton, each of them carrying a fraction x of the nucleon momentum, such that

$$\sum_{\text{partons}} x = 1. \tag{1}$$

According to this model, a hadron could be represented by mean of functions, the so–called parton densities, representing the probability that a parton of given kind (or flavor) carries a fraction x of the longitudinal momentum of the hadron. In this simple picture, the observed hadronic cross–section is simply given by the convolution between the parton densities and point–like partonic cross sections, which assume free partons. Therefore all the dynamical effect of strong interactions is contained in the specific form of the parton densities.

Pretty soon after the formulation of the parton model, these partons were then identified with the quarks introduced by Gell–Mann and Zweig to interpret the spectroscopy experiments. Partons were assigned spin 1/2 and electric charges of 2/3 and -1/3. This gave rise of a series of sum rules relating parton densities of different hadrons, which were experimentally satisfied. However, postulating the existence of "sea" quarks, *i.e.* of short living pairs of quark–antiquark in the nucleon during the collision, was not sufficient to fulfil the sum rule Eq. (1). It was also necessary to admit the existence in the hadron of neutral particles, called gluons, which do not interact with the electron probes. At the same time, it was recognised that the quark model seemed to require a new quantum number for the quarks, the colour. Originally introduced to solve a problem of Fermi statistic for the spin one-half quarks in the Δ^{++} baryon, the colour provided a natural set of currents to which gluons might couple.

The assembling of the ideas and experimental evidences mentioned above gave rise to the formulation of Quantum Chromo–Dynamics (QCD). Introduced in 1973, QCD is the renormalisable non–abelian theory based on the group SU(3) containing gluons and quarks as elementary fields [6, 7, 8][1]. The corresponding classical Lagrangian which exhibits explicitly the SU(3) symmetry is given by the Yang–Mill Lagrangian

$$\mathcal{L}_{\text{class}} = \sum_{\text{flavors}} \bar{\Psi}_a \left(i\gamma_\mu D^\mu - m\right)_{ab} \Psi_b - \frac{1}{4} \text{Tr}\, G^A_{\mu\nu} G^{\mu\nu}_A \,, \tag{2}$$

[1] For more references see Refs. [9, 10, 11] and references therein

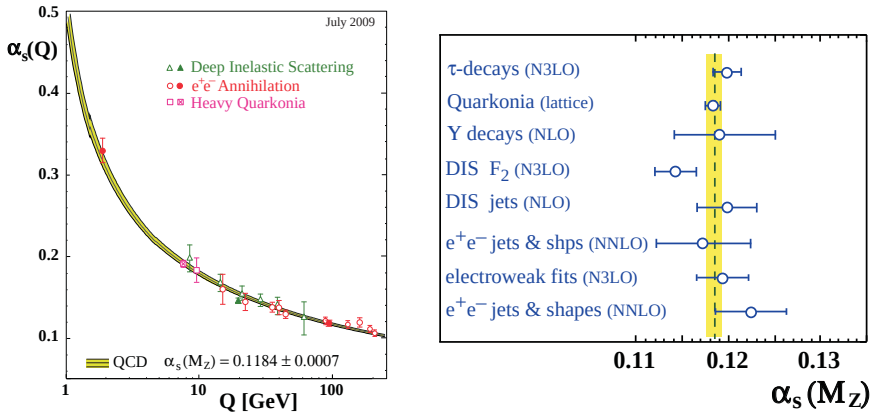


Figure 1: Experimental measurements of the universal QCD running coupling constant $\alpha_s(Q)$: (a) as a function of the scale Q; (b) summary of the measurements converted to an equivalent $\alpha_s(M_Z^2)$. The vertical line and shaded band mark indicate the final world average value from these measurements [12].

where Ψ_a are the quark fields, $G^A_{\mu\nu}$ is the field strength tensor derived from the gluon field A^A and D^μ is the covariant derivative, defined in Appendix A.

This theory possesses a number of important properties. In the low–energy regime it is strongly–interacting and produces an attractive force in a quark–antiquark and a three–quarks system. On the other hand, it predicts the *asymptotic freedom* [13, 14], according to which the coupling decreases as the energy increases, as it is shown in Fig. 1.

The scale dependence of the coupling constant encoded in the β function which has the following perturbative expansion

$$\beta(\alpha_s) = -\frac{\beta_0}{4\pi}\left(1 + \frac{\beta_1}{4\pi\beta_0}\alpha_s + \mathcal{O}(\alpha_s^2)\right) \tag{3}$$

where n_f is the number of light active flavors and

$$\begin{aligned} \beta_0 &= \frac{33 - 2n_f}{3} \\ \beta_1 &= \frac{2(153 - 19n_f)}{3}. \end{aligned} \tag{4}$$

Neglecting β_1 and higher order coefficients, the leading–order solution for the running of $\alpha_s(Q^2)$ is given by

$$\alpha_s(Q^2) = \frac{\alpha_s(\mu^2)}{1 + \alpha_s(\mu^2)\beta_0 \log(Q^2/\mu^2)}. \tag{5}$$

Contrarily to what we have in QED, the coefficient of the expansion of the β function in a non–abelian theory like QCD are negative and this is what makes α_s decrease as the energy increases, Eq. (5).

An important consequence of asymptotic freedom is that strong cross sections are computable in perturbation theory if a sufficiently high energy scale is involved in the computation. This would still not be enough to apply the results of perturbative calculations on partons to the world of observed hadrons. We need two other fundamental ingredients: the concept of infrared safety and the factorisation theorem. The former guarantees the cancellation between the singularities arising at the boundary of the phase space in each perturbative calculation and allows us to ignore the long–distance effects into the perturbative calculations. On the other hand, the factorisation theorem enables us to write down the hadronic cross–sections as convolutions between a hard part computable in perturbative QCD and a low–energy non–perturbative part up to $\mathcal{O}\left(\Lambda^2_{\rm QCD}/Q^2\right)$ corrections corresponding to higher–twist operators. The latter is given by the parton distribution functions (PDFs). They cannot be derived from first principles, but are process–independent and their scale dependence is predicted by perturbative QCD. Therefore we may extract PDFs from the available experimental data and evolve them to the scale of a new experiment for which PDFs represent a theoretical input. The kinematic plane (x, Q^2) with the region which is going to be explored by the LHC as compared to the region covered by some typical experiments from which parton distribution functions are extracted is shown in Fig. 2. Parton distributions and their uncertainties are fundamental inputs of any phenomenological prediction at the LHC and at the Tevatron. More details on the collinear factorisation theorem and on the evolution of parton distribution functions are given in Chapters 2 and 5. In the latter a method for a fast computation of the parton evolution and of the hadronic observables is described in details. An overview on how PDFs are extracted from data and on the statistical topics involved in global analyses are given in Chapter 3. In Chapter 4, the approach developed within the NNPDF collaboration is explained in all details and the physical results are exposed. The method lead to several phenomenological analyses which are discussed in Chapter 6.

In the real world, three of the six quarks present in Nature are light, while three of them have a mass which is not negligible with respect to the $\Lambda_{\rm QCD}$ scale at which the perturbative approach to QCD breaks down. The precise definition of the PDFs and the formulation of the factorisation does depend on considerations about the rela-

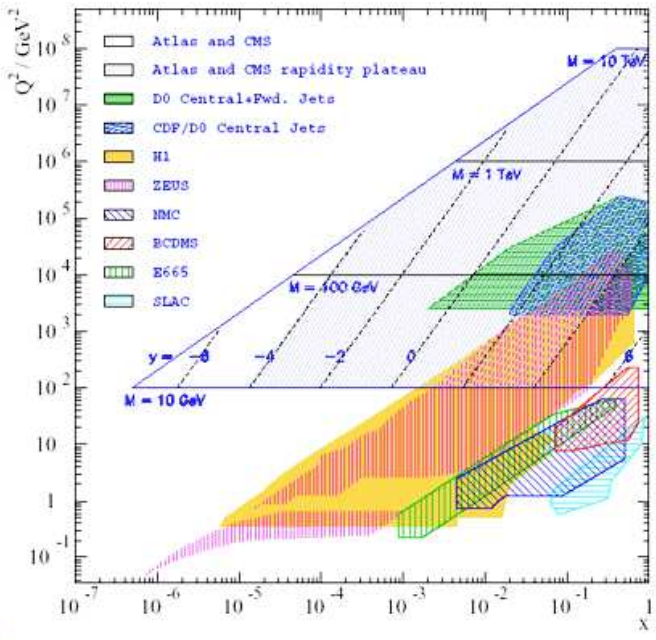


Figure 2: This plot shows the kinematical coverage of the LHC when the centre–of–mass energy will reach 14 TeV in the plane (x, Q^2) compared to the kinematical coverage of previous colliders.

tive magnitude of the scales involved in the process with respect to the quark masses (m_c, m_b, m_t). This is not a trivial issue because different regimes, depending on whether a quark is considered light or heavy, correspond to different theories and require a suitable definition of the renormalised quantities. The general framework and the definition of renormalisation schemes able to deal with this multi–scale problem are given in the second part of Chapter 2.

As in perturbative QCD we model the hadron as an object constituted by massless partons, it is not trivial to perform calculations of processes initiated by a heavy quark. The schemes elaborated in literature in order to perform such calculations exploit what is already well–known about perturbative QCD with light quarks in the initial states and heavy particles in the final states. In particular, there are two extreme and complementary schemes which are generally employed depending on the relative size of the heavy quark mass with respect to the hard scale of the process. In the so–called massless scheme the heavy quark is treated as a parton in the initial state, while in the opposite massive scheme the heavy quark is treated as a massive particle in the final state. Both schemes present several advantages and disadvantages. In the massless scheme

theoretical calculations at next-to-leading orders are highly simplified and logarithmic corrections associated to the collinear splitting of the heavy quarks are resummed to all orders. However the calculation might be very inaccurate in kinematic regions where the effect of the non-logarithmic terms are important. On the other hand, the massive scheme does not resum the logarithmic corrections but it treats correctly the leading–order kinematics and employs coherently the mass threshold effects. A detailed analysis of their difference is presented in Chapter 7. This study leads to a better understanding of the size of the potentially large logarithms resummed by the DGLAP equations and allows a systematic estimate of their impact.

Chapter 1

Perturbative Quantum Chromo–Dynamics

In this chapter some essential features of perturbative QCD are discussed. The collinear factorisation theorem is introduced by using the deep inelastic scattering process as a paradigm. The flavor decomposition and the solution of the DGLAP equations are described in detail. In the second part of the chapter, the discussion focuses on the description of a renormalisation scheme able to deal with the finite heavy quarks masses. The deriving schemes used to perform calculations involving heavy quarks are then reviewed.

1.1 The QCD improved parton model

Quantum Chromo–Dynamics (QCD) is the universally accepted theory of the strong interactions. The success of the theory has been confirmed by the comparison between the experimental data collected in the last forty years and its predictions.
Although we cannot see directly the quarks and the gluons as we do with the other elementary particles, their presence is unmistakably revealed in high energy interactions as distinctive "jets", *i.e.* as a bunch of collimated hadrons. In Fig. 1.1 a typical signature from the annihilation of electrons and positrons at the centre–of–mass energy of 91 GeV at LEP is shown. The signature can be interpreted as illustrated on the right–hand side of the same figure: the e^+e^- pair annihilates into a quark–antiquark pair which then "hadronise" into the observed pencil–like jets. The same experimental

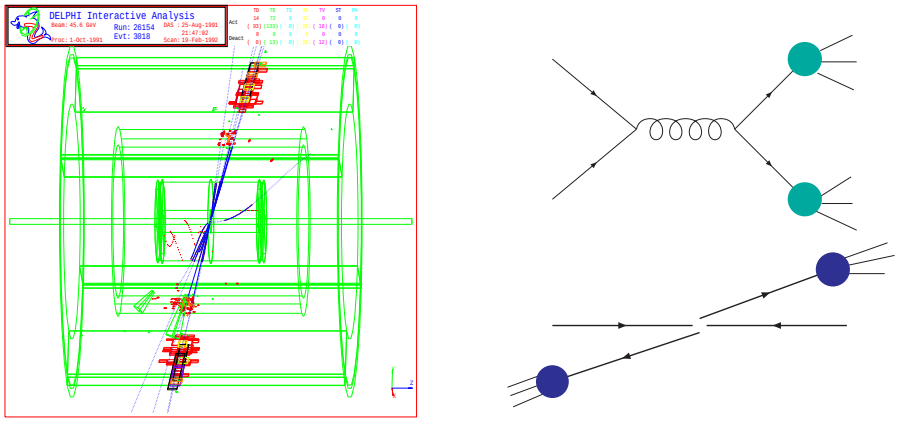


Figure 1.1: Typical event in $e^+e^- \to$ hadrons with 2–jets in final state at LEP. Clear evidence for $q\bar{q}$ creation.

evidence was reached for the existence of gluon. In Fig. 1.2 a 3–jets event in e^+e^- collision is shown. The natural interpretation is drawn on the right of the experimental trace: the event correspond to a $q\bar{q}g$ final state with subsequent hadronisation of the quarks and gluon into hadronic jets. Detailed studies of the angular distribution of the three jets confirm this picture, the spin of the underlying third jet being consistent with one. This proves that the gluon is a vector–boson.

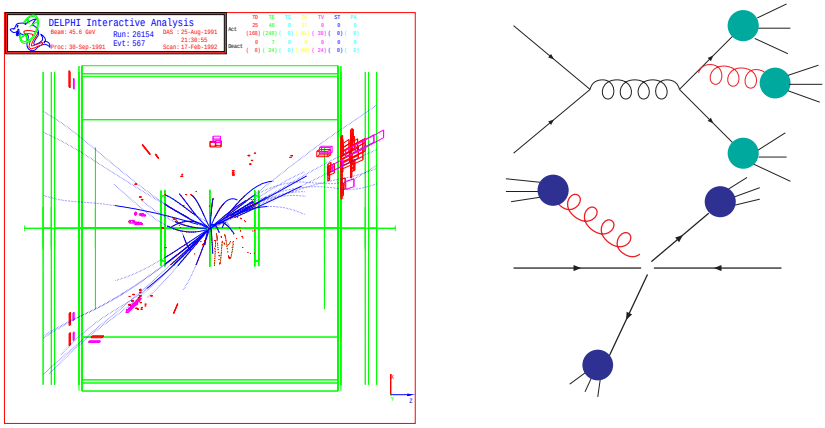


Figure 1.2: Typical event in $e^+e^- \to$ hadrons with 3–jets in final state at LEP. Clear evidence for $q\bar{q}g$ creation.

The unique feature of the underlying Quantum Field Theory which makes the perturbative approach applicable to QCD is the concept of asymptotic freedom. The strong processes computable in perturbation theory are those which involve a high energy scale so that the coupling is sufficiently small. Some examples which will be discussed in this thesis are: the deep inelastic scattering (DIS) of an electron off a proton, where the hard scale Q^2 is given by the transferred momentum, and the production of heavy particles, such as weak bosons, where the hard scale is given by the mass of the produced particles.

Even though QCD is asymptotically free, the computation of any strong cross section does involve non-perturbative contributions, since the initial and final states are not the fundamental degrees of freedom of the theory but compound states of quarks and gluons. An important property of QCD is the factorisation theorem, which basically enables us to separate in every process a hard part, computable in perturbation theory, from a low energy one, which is process-independent and can be taken as a phenomenological input. The latter, given by the Parton Distribution Functions (PDFs), parametrise our ignorance on the inner structure of the nucleons. More details on the way in which PDFs are extracted from the existing data and on the related phenomenological issues are given in the next Chapter.

The possibility of separating long and short distance effects largely explains the success of the parton model, a predecessor of QCD, introduced by Feynman and Bjorken in the late Sixties [15]-[16]. It began as a quasi–classical model for DIS, based upon the idea that the hadron can be described as a bunch of collinear independent partons, carrying a fraction of the total longitudinal momentum of the hadron, off which a lepton can scatter via the exchange of a vector boson. Nowadays the parton model is understood as the lowest order approximation of a perturbative QCD calculation. To maintain the separation between long and short distance effects in the presence of QCD corrections, one is obliged to make PDFs scale dependent, that is, functions of both x and Q^2. The dependence of PDFs on Q^2 has been confirmed and accurately measured by experiments, especially by the large amount of data collected at the HERA collider [17]. The dependence of PDFs on the scale can be understood pictorially with the idea that the parton carries within it further daughter–partons and that these are revealed when the energy of the probing vector boson is increased. Fortunately the scale dependence is predicted by perturbative QCD and it is governed by the DGLAP equations. The latter allow one to deduce PDFs at a given scale from a set of PDFs extracted at any other given scale.

In the following sub–sections we first provide the Leading–Order picture of QCD through the parton model description of Deep–Inelastic Scattering. From there we give a heuristic development of the NLO perturbative QCD corrections to DIS and we use it to discuss collinear factorisation. After that we focus on the DGLAP evolution

equations by providing details on their solution. This is followed by a discussion on how these equations resum large logarithmic enhancements to a cross section at all orders in perturbative QCD. Finally, we show how the factorisation formalism is applied to the vector production process in hadron–hadron collisions.

1.1.1 Deep Inelastic Scattering

Lepton–hadron scattering is the traditional method for probing the structure of hadrons and the first testing ground of perturbative QCD [11] as well as having established the first evidence of partons. It consists in the scattering of a high-energy charged lepton off a hadron target as shown in Fig. 1.3:

$$l(k) + h(P) \longrightarrow l'(k') + X(P_X).$$

At the leading order, it is the manifestation of the partonic sub–process:

$$l(k) + q(zP) \longrightarrow l'(k') + q'(p'),$$

where the quark q carries a fraction z of the momentum of the incoming hadron. A basic classification of the events is based on the nature of the boson exchanged by the initial lepton and quark. In neutral current events $l = l'$ and the boson exchanged is either a photon or a Z. In charged current events, characterised by $l = e, \mu, \tau$ and $l' = \nu_e, \nu_\mu, \nu_\tau$ or vice versa, a $W^\pm$ boson is exchanged. If the initial lepton is a neutrino, only weak interactions occur, making this class of measurements a useful probe to disentangle the contributions from quarks and anti–quarks.

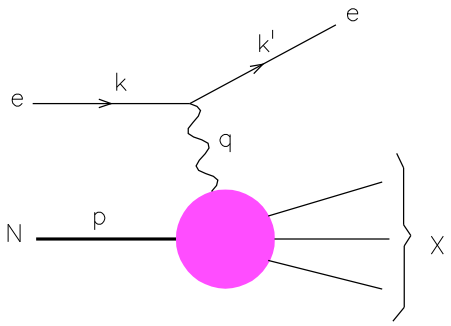


Figure 1.3: Schematic representation of deep–inelastic charged lepton–nucleon scattering.

If we label the momentum of the exchanged boson by $q^\mu = k^\mu - k'^\mu$, neglecting the relatively small lepton masses, the standard kinematic variables are defined by

$$\begin{aligned} Q^2 &\equiv -q^2 = -(k-k')^2 = +2E_l E_{l'}(1-\cos\theta) \\ \nu &\equiv (P\cdot q)/M_N = (E_l - E_{l'}), \end{aligned} \tag{1.1}$$

where the energies refer to the target rest frame and M_N is the target mass. The scale Q^2 characterises the resolving power of the probe, while ν represents the energy transferred to the target hadron. Two other Lorentz–invariants, often used to define the DIS kinematics, are the Bjorken–variable x_B and inelasticity y, which take values between 0 and 1

$$\begin{aligned} x_B &= \frac{Q^2}{2M_N\nu} \\ y &= \frac{q\cdot P}{k\cdot P} = \frac{E_l - E_{l'}}{E_l}. \end{aligned} \tag{1.2}$$

Once the energy of the lepton–nucleon centre–of–mass S is set, only two variables out of x_B, y and Q^2 are independent of each others, being

$$S = (P+k)^2 = \frac{Q^2}{x_B y} + M_N^2. \tag{1.3}$$

The matrix element of the hadronic process may be written in terms of the leptonic and hadronic currents as

$$\mathcal{M}(lh \to l'X) \;=\; \left\langle l'|J^l_\mu|l\right\rangle g_{lV}\frac{-g^{\mu\nu}}{q^2 - M_V^2} g_{hV}\left\langle X|\mathcal{J}^h_\nu|h\right\rangle, \tag{1.4}$$

where M_V is the mass of the exchanged vector boson, J^l and $\mathcal{J}^h$ are respectively the conserved leptonic and hadronic currents, g_{hV} is the coupling between the hadron and the vector boson and $g_{lV} = k_V^2(V_{lV}^2 + A_{lV}^2)$ is the electroweak coupling, whose expression is shown in Table 1.1 for each type of vector boson. Also notice that the propagator in Eq. (1.4) is implicitly defined in the Feynman gauge. Eq. (1.4) suggests that the inclusive lepton–hadron scattering cross section may be written in terms of two tensors $L_{\mu\nu}$ and $H^{\mu\nu}$ as

$$d\sigma^{(lh)} \;=\; \frac{1}{4(k\cdot P)}\frac{(g_{lV}g_{hV})^2}{(Q^2+M_V^2)^2}\, L_{\mu\nu}H^{\mu\nu}(4\pi)\frac{d^3k'}{2E_{k'}(2\pi)^3} \tag{1.5}$$

Bosons	k_V	V_{lV}	A_{lV}
γ	e_l	1	0
Z	$1/(2\sin\theta_W\cos\theta_W)$	$I_3^l - 2e_l\sin^2\theta_W$	-I_3^l
$W^\pm$	$V_{ll'}/(2\sqrt{2}\sin\theta_W)$	1	-1

Table 1.1: Coupling of fermions to the weak bosons. Here e_l is the electric charge measured in unit of the positron charge, I_3^l is the third component of the weak isospin, $+1/2$ for up–type quarks or neutrinos and $-1/2$ for down–type quarks or charged leptons. For charged current interactions involving quarks, the coefficients $V_{ll'}$ of the Cabibbo–Kobayashi–Maskawa matrix [18]-[19] are involved. The parameter $\sin\theta_W$ is the Weinberg mixing angle.

which, evaluating the Jacobian between $(E_{l'}, \cos\theta)$ and (x_B, Q^2) by mean of Eqs. (1.1) and (1.2), becomes

$$\frac{d\sigma^{(lh)}}{dx_B dQ^2} = \frac{Q^2}{2x_B^2(k+P)^2}\frac{(g_{lV}g_{hV})^2}{(Q^2+M_V^2)^2}\frac{1}{4\pi}L_{\mu\nu}H^{\mu\nu} \tag{1.6}$$

with

$$L_{\mu\nu} = \frac{1}{2}\left\langle l|J_\mu^{l\dagger}|l'\right\rangle\left\langle l'|J_\nu^l|l\right\rangle \tag{1.7}$$

$$H^{\mu\nu} = \frac{1}{2}\frac{1}{4\pi}\sum_X \left\langle h|\mathcal{J}^{h\dagger\mu}|X\right\rangle\left\langle X|\mathcal{J}^{h\nu}|h\right\rangle(2\pi)^4\delta(P_X - k - P). \tag{1.8}$$

The leptonic tensor $L_{\mu\nu}$ is easily calculable,

$$L_{\mu\nu} = \left[k_\mu k'_\nu + k'_\mu k_\nu - \frac{Q^2}{2}g_{\mu\nu} + iC_{lV}\epsilon_{\mu\nu}^{\alpha\beta}\,k_\alpha k'_\beta\right]. \tag{1.9}$$

The last term is associated to the parity violation of the weak interaction, while the coefficient

$$C_{lV} = \frac{2A_{lV}V_{lV}}{(V_{lV}^2 + A_{lV}^2)}$$

depends only on the type of vector boson exchanged and can be read out of Table 1.1. As QED does not violate parity, $C_{l\gamma} = 0$. In Eq. (1.9) we have neglected the term associated to the parity violation proportional to the masses of the leptons.

The hadronic tensor is harder to calculate. It is summed over all allowed final states and by convention includes a factor $(4\pi)^{-1}$ and an overall four-momentum conserving δ–function. Since it is constructed from the only two available four–vectors P^μ and

q^μ and since it must be Lorentz–invariant, we may write down its general form as:

$$\begin{aligned} H^{\mu\nu} &= -g^{\mu\nu}F_1 + (P\cdot q)^{-1}\Big[P^\mu P^\nu F_2 + i\epsilon^{\mu\nu}_{\alpha\beta}P^\alpha q^\beta F_3 + P^\mu q^\nu(F_4 + iF_5) \\ &\quad + q^\mu P^\nu(F_4 - iF_5) + q^\mu q^\nu F_6\Big], \end{aligned} \tag{1.10}$$

where $F_i \equiv F_i(x, Q^2)$ are the so–called structure functions. If the spin of the colliding particles is specified, there are extra terms involved. The general form of the hadronic tensor for unpolarised scattering simplifies further if we impose time–reversal invariance, which sets $F_5 = 0$, and electro–magnetic gauge invariance, which implies electro–magnetic current conservation, $q_\mu H^{\mu\nu} = 0$ and $H^{\mu\nu}q_\nu = 0$. As a consequence, the form of the hadronic tensor is further simplified to

$$\begin{aligned} H^{\mu\nu} &= \left(-g^{\mu\nu} + \frac{q^\mu q^\nu}{q^2}\right)F_1 + (P\cdot q)^{-1}\left(P^\mu - \frac{P\cdot q}{q^2}q^\mu\right)\left(P^\nu - \frac{P\cdot q}{q^2}q^\nu\right)F_2 \\ &\quad + i(P\cdot q)^{-1}\epsilon^{\mu\nu\alpha\beta}\,P_\alpha q_\beta F_3. \end{aligned} \tag{1.11}$$

In order to simplify the following calculations it is convenient to project out of the hadronic tensor two combinations of structure functions, the transverse and longitudinal components in $d = (4 - 2\epsilon)$ dimensions:

$$\begin{aligned} H_T &\equiv -g_{\mu\nu}H^{\mu\nu} && (1.12)\\ &= (d-2)\frac{F_2}{2x_B}\left(1 + \frac{(2x_BM_N)^2}{Q^2}\right) - (d-1)\left[\frac{F_2}{2x_B}\left(1 + \frac{(2x_BM_N)^2}{Q^2}\right) - F_1\right] \\ H_L &\equiv P_\mu P_\nu H^{\mu\nu} && (1.13)\\ &= \frac{Q^2}{(2x_B)^2}\left[\frac{F_2}{2x_B}\left(1 + \frac{(2x_BM_N)^2}{Q^2}\right) - F_1\right]\left(1 + \frac{(2x_BM_N)^2}{Q^2}\right). \end{aligned}$$

Combining Eqs. (1.5) and (1.11), we can write the general expression for the unpolarised and inclusive lepton–hadron scattering cross–section in terms of the structure functions F_i as

$$\begin{aligned} \frac{d^2\sigma^{(\mathrm{lh})}}{dx_B dQ^2} &= \frac{y}{Q^2}\frac{d^2\sigma^{(\mathrm{lh})}}{dx_B dy} \\ &= \frac{(4\pi)\alpha_{lV}\alpha_{hV}}{(Q^2 + M_V^2)^2}\left[y^2F_1 + \left(1 - y - \frac{(x_ByM_N)^2}{Q^2}\right)\frac{F_2}{x_B} - C_{lV}\left(y - \frac{y^2}{2}\right)F_3\right], \end{aligned} \tag{1.14}$$

where $\alpha_{lV} = g_{lV}^2/(4\pi)$ and $\alpha_{hV} = g_{hV}^2/(4\pi)$.

The information on the a priori unknown structure of the target as seen by the virtual boson is carried by the structure functions F_i. They can only be functions of Q^2, x_B or y as well as the mass of the nucleon. In the parton model the structure functions have a very simple expressions, since the hadron is described in terms of the probability density distributions for the momentum fractions of its parton constituents

$$f_{i,h}(z)dz \;=\; \mathcal{P}(z' \in [z, z+dz]) \qquad \text{with } i = q, \bar{q}, g, \tag{1.15}$$

which only depend on the hadron itself and are independent of the process. They also do not depend on the scale Q^2. In this simple picture, which was then understood as the $Q^2 \to \infty$ limit, the structure functions are observed to obey an approximate scaling law

$$F_i(x_B, Q^2, M_N^2) \to F_i(x_B). \tag{1.16}$$

The hadron cross section $d\sigma^{(lh)}$ is then given by the sum of point–like quark or anti–quark cross sections $d\hat{\sigma}^{(li)}$ weighted by the functions $f_{i,h}(z)$

$$d\sigma^{(lh)} = \sum_{i=q,\bar{q},g} \int_0^1 dz\, f_{i,h}(z)\, d\hat{\sigma}^{(li)} \left(\frac{x_B}{z}\right), \tag{1.17}$$

and the hadron tensor (1.8) can be written as a sum of partonic tensors $d\hat{H}^{(i)}_{\mu\nu}$

$$H_{\mu\nu}(P,q) = \sum_{i=q,\bar{q},g} \int_0^1 \frac{dz}{z}\, f_{i,h}(z)\, d\hat{H}^{(i)}_{\mu\nu}\,(zP,q)\,. \tag{1.18}$$

The same projections performed at the hadronic level in Eq. (1.13) can be done at partonic level, just keeping in mind that the longitudinal component must be projected by $p_\mu p_\nu = z^2 P_\mu P_\nu$ and therefore

$$\begin{aligned} H_T &= \sum_{i=q,\bar{q},g} \int_0^1 \frac{dz}{z}\, f_{i,h}(z)\, d\hat{H}^{(i)}_T\,(zP,q) \\ H_L &= \sum_{i=q,\bar{q},g} \int_0^1 \frac{dz}{z^3}\, f_{i,h}(z)\, d\hat{H}^{(i)}_L\,(zP,q)\,. \end{aligned} \tag{1.19}$$

The projections of the partonic tensor are related to the partonic matrix elements $\mathcal{M}$ by the following relations

$$\begin{aligned}\hat{H}_T &= \frac{1}{4\pi e^2}\,(-g_{\mu\nu})\int d\Phi \mathcal{M}^\mu \mathcal{M}^\nu, \\ \hat{H}_L &= \frac{1}{4\pi e^2}\,p_\mu p_\nu \int d\Phi \mathcal{M}^\mu \mathcal{M}^\nu,\end{aligned} \tag{1.20}$$

as it can be inferred from Eq. (1.6). From the above equations one can easily derive the expression of each structure function as a combination of PDFs, depending on the nature of the exchanged boson. Indeed, inverting Eq. (1.13), and applying the parton model factorisation, Eq. (1.18), one gets

$$\begin{aligned}\frac{F_2(x)}{x} &= \frac{1}{(1-\epsilon)}H_T + \frac{(3-2\epsilon)}{(1-\epsilon)}\frac{4z^2}{Q^2}H_L \\ &= \sum_{i=q,\bar{q},g}\int_x^1 \frac{dz}{z} f_{i,h}\left(\frac{x}{z}\right)\left[\frac{1}{(1-\epsilon)}\hat{H}_T^{(i)}(z) + \frac{(3-2\epsilon)}{(1-\epsilon)}\frac{4x^2}{Q^2}\hat{H}_L^{(i)}(z)\right], \\ F_1(x) &- \frac{F_2(x)}{2x} = -\frac{4x^2}{Q^2}H_L \\ &= \sum_{i=q,\bar{q},g}\int_x^1 \frac{dz}{z} f_{i,h}\left(\frac{x}{z}\right)\frac{4z^2}{Q^2}d\hat{H}_L^{(i)}(z).\end{aligned} \tag{1.21}$$

To be more explicit, we may consider the partonic process $\gamma^* q \to q$. First we evaluate the spin and colors averaged matrix element squared in d–dimensions, with $d = (4 - 2\epsilon)$

$$-g_{\mu\nu}\overline{\sum}\mathcal{M}^\mu(\gamma^* q \to q)\mathcal{M}^\nu(\gamma^* q \to q) \;=\; 2e^2 e_q^2(1-\epsilon)Q^2. \tag{1.22}$$

Integrating over the one–body phase space one obtains

$$-g_{\mu\nu}\int d\Phi_1 \overline{\sum}\mathcal{M}^\mu(\gamma^* q \to q)\mathcal{M}^\nu(\gamma^* q \to q) \;=\; 2e^2 e_q^2(1-\epsilon)Q^2(2\pi)\delta(p'^2). \tag{1.23}$$

Since the struck quark carries a fraction z of the momentum of the parent's hadron momentum

$$p'^2 = (zP+q)^2 = \frac{Q^2}{x_B}(z - x_B) \quad \Rightarrow \quad \delta(p'^2) = \frac{Q^2}{x_B}\delta(z - x_B), \tag{1.24}$$

which implies that the fraction of the momentum carried by the struck parton is equal to the Bjorken-x. This is quite remarkable, namely a macroscopic parameter controls the momentum of the parton involved in the process. As far as the longitudinal

projection is concerned, for massless leptons it vanishes

$$q_\mu q_\nu \overline{\sum} \mathcal{M}^\mu(\gamma^* q \to q)\mathcal{M}^\nu(\gamma^* q \to q) \; = 0. \tag{1.25}$$

Finally, from Eqs. (1.19) and (1.21), one gets

$$\begin{aligned} F_2^\gamma(x) &= x \sum_{i=q,\bar{q}} e_q^2 f_{i,h}(x) \\ xF_3^\gamma(x) &= 0 \\ 2xF_1^\gamma(x) &= F_2(x). \end{aligned} \tag{1.26}$$

The last equality in Eq. (1.26) is the so–called Callan-Gross relation and it is a direct consequence of the spin-$\frac{1}{2}$ property of the quarks. F_1 and $(F_2 - 2xF_1)$ correspond respectively to the absorption of transversely and longitudinally polarised virtual photons. The combination

$$F_L(x,Q^2) = \left(1 + \frac{4M_N^2 x^2}{Q^2}\right) F_2(x,Q^2) - 2xF_1(x,Q^2) \to_{Q^2\to\infty} \; F_2(x) - 2xF_1(x) \tag{1.27}$$

is called longitudinal structure function. Structure function measurements show that $F_L \ll F_2$ confirming the spin-1/2 property of quarks.
The treatment of Z exchange involves only minor modifications due to the different coupling and the presence of a non–zero axial contribution. Keeping them into account one ends up with the following expressions

$$\begin{aligned} F_2^Z(x) &= \sum_{i=q,\bar{q}} xB_i(Q^2) f_{i,h}(x), \\ xF_3^Z(x) &= \sum_{i=q} xD_i(Q^2)\left[f_{i,h}(x) - f_{\bar{i},h}(x)\right], \end{aligned} \tag{1.28}$$

with

$$\begin{aligned} B_i(Q^2) &= -2e_i V_{eZ} V_{iZ} P_Z + (V_{eZ}^2 + A_{eZ}^2)(V_{iZ}^2 + A_{iZ}^2) P_Z^2, \\ D_i(Q^2) &= -2e_i A_{eZ} A_{iZ} P_Z + 4V_{eZ} A_{eZ} V_{iZ} A_{iZ} P_Z^2, \\ P_Z &= \frac{Q^2}{(Q^2 + M_Z^2)(4\sin^2\theta_W \cos^2\theta_W)}, \end{aligned} \tag{1.29}$$

where V_{iZ} and A_{iZ} are respectively the vector and the axial couplings of Table 1.1. Finally the charged current contribution associated to the W^+ exchange $l^+ N \to \nu X$

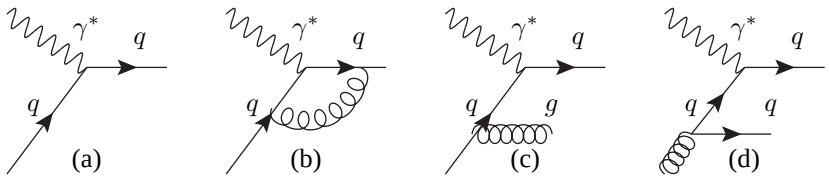
reads as

$$\begin{aligned} F_2^{W^+}(x) &= 2x \sum_i \left[D_{i,h}(x) + \bar{U}_{i,h}(x) \right] \\ xF_3^{W^+}(x) &= \sum_i \left[D_{i,h}(x) - \bar{U}_{i,h}(x) \right], \end{aligned} \tag{1.30}$$

where U includes all up–type quarks and D all the down–type quarks. For the exchange of W^-, the expression is the same as Eq. (1.31) where $U \to D$ and vice versa.

1.1.2 Next–to–leading order QCD corrections

The parton model existed before the formulation of QCD. As soon as QCD corrections are added on top of tree-level processes, singularities appear in the calculation. In the following we go through the calculation of the next–to–leading order QCD corrections to the DIS process, performing all calculations in $d = (4 - 2\epsilon)$ dimensions. This provides a simple ground where the collinear factorisation theorem can be discussed in more details. In the following we consider a process where a virtual photon is exchanged, given that only minor modifications are involved if we were considering a weak boson exchange. We further assume that ultra–violet singularities have been already taken care of by the renormalisation of the bare constants of the theory.
At $\mathcal{O}(\alpha_s)$ two classes of contributions appear; the interference between the tree–level amplitude (a) and the one–loop correction (b) have to be considered, together with the emission of a real gluon at tree–level (c). Moreover, one has to consider a process initiated by a gluon which then splits into a pair of quark and anti–quark, the so–called boson–gluon fusion process (d).



We first concentrate on the $\mathcal{O}(\alpha_s)$ contribution coming from the process $q\gamma^* \to q$ at one–loop and the same contribution at tree–level. The structure of the one–loop vertex and the tree level diagram are the same, so, as we do in the renormalisation of the UV divergencies to define the renormalised coupling, we define an effective vertex [11]

$$i\Gamma^\mu = -iee_q \left[1 - \frac{\alpha_s}{4\pi} C_F \left(\frac{4\pi\mu}{Q^2} \right)^\epsilon \frac{\Gamma(1-\epsilon)}{\Gamma(1-2\epsilon)} \left(\frac{2}{\epsilon^2} + \frac{3}{\epsilon} + 8 + \frac{\pi^2}{3} + \mathcal{O}(\epsilon) \right) \right]. \tag{1.31}$$

The double pole in ϵ originates from the region of the loop integration where the exchanged virtual gluon is simultaneously soft and collinear to the massless quark lines. As we are going to see explicitly, the singularity is cancelled by the analogous contribution from the emission of one real gluon. This is an example of infrared safety of QCD: virtual and real corrections cancel the infrared singularities. The KNL theorem [20, 21] and its generalisation to QCD [22, 23] guarantees that what we have seen in the case of DIS at next–to–leading order happens to all orders for any QCD inclusive observable.

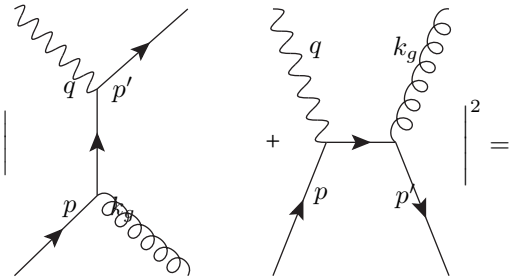
The calculation of the transverse component of the hadronic tensor coming from this virtual contribution is straightforward, it only requires the integration over the one–body phase space as in Eq. (1.23) and the multiplication by the conventional factor $1/(4\pi e^2)$, yielding

$$\begin{aligned} \hat{H}^{(\gamma q)}_{T,\mathrm{virt}} &= e_q^2(1-\epsilon)\delta(1-\xi) \\ &\quad \left\{1 - \frac{\alpha_s}{2\pi} C_F \left(\frac{4\pi\mu^2}{Q^2}\right)^{\epsilon} \frac{\Gamma(1-\epsilon)}{\Gamma(1-2\epsilon)} \left(\frac{2}{\epsilon^2} + \frac{3}{\epsilon} + 8 + \frac{\pi^2}{3}\right) \delta(1-\xi)\right\}. \end{aligned} \tag{1.32}$$

We now concentrate on the real contribution

$$\gamma^*(q) + q(p) \to q(p') + g(k_g),$$

providing some more details on the calculation, which contains examples of all singularities that we have to deal with. The square of the matrix element summed over the polarisations of the outgoing lepton and averaged over the incoming spins and colors is given by

$\overline{\sum}|\mathcal{M}(\gamma^* q \to q' g)|^2 =$ 

$$= 8N_c C_F e^2 e_q^2 (g_s \mu^{\epsilon})^2 \left[(1-\epsilon)\left(\frac{k_g\cdot p'}{k_g\cdot p} + \frac{k_g\cdot p}{k_g\cdot p'}\right) + \frac{Q^2(p\cdot p')}{(k_g\cdot p)(k_g\cdot p)} + 2\epsilon\right].$$

The three terms above correspond to the emission of a gluon off the outgoing quark (s–channel), the emission of the gluon off the incoming quark (t-channel) and the

interference between the two contributions. To obtain $\hat{H}_T^{(\gamma q)}$, we need to integrate over the two–body phase space of the final state particles,

$$\begin{aligned}\hat{H}_{T,\mathrm{real}}^{(\gamma q)} &= \frac{1}{4\pi e^2}\int d\Phi_2^{\epsilon}\overline{\sum}|\mathcal{M}(\gamma^* q\to q' g)|^2 \\ &= 4e_q^2\alpha_s C_F\int d\Phi_2^{\epsilon}\left[\frac{k_g\cdot p'}{k_g\cdot p}+\frac{k_g\cdot p}{k_g\cdot p'}+\frac{Q^2(p\cdot p')}{(k_g\cdot p)(k_g\cdot p)}\right], \quad (1.33)\end{aligned}$$

where the two–body phase space $d\Phi_2^{\epsilon}$ is evaluated in $(4-2\epsilon)$ dimensions. Looking at the denominators of the terms appearing in the integrand, we see that there are two kind of divergences: collinear, when the gluon is emitted parallel to the incoming or outgoing lepton, and soft, when the energy of the gluon vanishes. In the partonic centre–of–mass frame one may write explicitly the four-momenta of the particles in terms of s, the centre–of–mass energy in the γ^*–parton system,

$$\hat{s}=(zP+q)^2=Q^2(1-\xi)/\xi, \qquad \text{with} \qquad \xi=x_B/z$$

and the variable v related to the angle θ^* between incoming and outgoing particles in the centre–of–mass frame by $v\equiv(1+\cos\theta^*)/2$. In terms of these variables, Eq.(1.33) becomes

$$\begin{aligned}\hat{H}_{T,\mathrm{real}}^{(\gamma q)} &= \frac{e_q^2\alpha_s C_F}{8\pi}\left(\frac{2\pi\mu^2}{\sqrt{\hat{s}}}\right)^2\frac{(1-\epsilon)}{\Gamma(1-\epsilon)} \quad (1.34)\\ &\int_0^{+1} dv\, v^{-\epsilon}(1-v)^{-\epsilon}\left\{(1-\epsilon)\left[\frac{v}{(1-\xi)}+\frac{(1-\xi)}{v}\right]+\frac{2\xi}{(1-\xi)}\frac{(1-v)}{v}\right\}.\end{aligned}$$

where the soft gluon and the initial state collinear singularities manifest themselves as $\xi\to 1$ and $v\to 0$ respectively. The final state collinear singularity belongs to the same class as the soft singularity because both of them imply a zero–mass particle in the final state. Indeed

$$(zP+q)^2=Q^2\frac{(1-\xi)}{\xi}=(k_g+p')^2=2(k_g\cdot p').$$

The limit $(k_g\cdot p')\to 0$ is the same as $\xi\to 1$, both of them involving the kinematics of the lower order scattering to which they can be associated to.
Using the definition of the standard Euler β–function and the following identity

$$\frac{\xi^{\epsilon}}{(1-\xi)^{\epsilon}}=-\frac{1}{\epsilon}\delta(1-\xi)+\left(\frac{1}{1-\xi}\right)_+-\epsilon\left(\frac{\log(1-\xi)}{1-\xi}\right)_+ +\epsilon\frac{\log\xi}{1-\xi}, \quad (1.35)$$

we may solve the integration in Eq. (1.34) by expanding in ϵ. We end up with

$$\begin{aligned}\hat{H}^{(\gamma q)}_{T,\mathrm{real}} &= e_q^2 \frac{\alpha_s}{2\pi} C_F \left(\frac{4\pi\mu^2}{Q^2}\right)^{\epsilon} (1-\epsilon)\frac{\Gamma(1-\epsilon)}{\Gamma(1-2\epsilon)} \\ &\quad \left\{ \left(\frac{2}{\epsilon^2}+\frac{3}{2\epsilon}+\frac{7}{2}\right)\delta(1-\xi) - \frac{1}{\epsilon}\frac{(1+\xi^2)}{(1-\xi)_+} + (1+\xi^2)\left(\frac{\log(1-\xi)}{1-\xi}\right)_+ \right. \\ &\quad \left. -\frac{1+\xi^2}{(1-\xi)}\log\xi - \frac{3}{2}\frac{1}{(1-\xi)_+} + 3 - \xi + \mathcal{O}(\epsilon)\right\}, \end{aligned} \tag{1.36}$$

where the plus–prescription is a distribution defined as

$$F(\xi)_+ = F(\xi) - \delta(1-\xi)\int_0^1 dy\, F(y). \tag{1.37}$$

The double pole associated to the soft gluon singularity, as anticipate before, is cancelled out by the analogous virtual soft singularity. We are left with the single pole associated to the emission of gluons collinear to the incoming parton. To derive F_2, one needs also the longitudinal component of the partonic tensor, which is non–singular and easy to deduce. As in the leading–order calculation, it vanishes in the virtual contribution, while for the real contribution, it is finite and is given by

$$\begin{aligned}\hat{H}^{(\gamma q)}_{L,\mathrm{real}} &= \frac{1}{4\pi e^2}\int d\Phi_2 \overline{\sum} |p_\mu \mathcal{M}^\mu(\gamma^* q \to q' g)|^2 \\ &= \frac{e_q^2}{4}\frac{\alpha_s}{2\pi} C_F \frac{Q^2}{\xi}\left(\frac{4\pi\mu^2}{Q^2}\frac{\xi}{1-\xi}\right)^{\epsilon}\frac{\Gamma(2-\epsilon)}{\Gamma(2-2\epsilon)} \\ &= \frac{e_q^2}{4}\frac{\alpha_s}{2\pi} C_F \frac{Q^2}{\xi} + \mathcal{O}(\epsilon). \end{aligned} \tag{1.38}$$

Summing up Eqs. (1.32) and (1.36) and combining them with Eq. (1.38), we can write the structure functions F_2 and F_1 as

$$\begin{aligned} F_2^{\gamma q}(x) &= \frac{\alpha_s}{2\pi}\int_x^1 \frac{d\xi}{\xi} q\left(\frac{x}{\xi}\right)\left\{ -P_{qq}^{(0)}(\xi)\left[\frac{1}{\epsilon} - \gamma_E + \log(4\pi) - \log\left(\frac{Q^2}{\mu^2}\right)\right]\right. \\ &\qquad \left. C_F\left[\frac{1+\xi^2}{1-\xi}\left(\log\left(\frac{1-\xi}{\xi}\right) - \frac{3}{4}\right) + \frac{5\xi+9}{4}\right]_+ \right\} \\ F_1^{\gamma q}(x) &= \frac{F_2^{\gamma q}(x)}{2x} - e_q^2\frac{\alpha_s}{2\pi}\int_x^1 \frac{d\xi}{\xi} q\left(\frac{x}{\xi}\right) C_F \xi. \end{aligned} \tag{1.39}$$

where $P_{qq}^{(0)}$ is the Altarelli–Parisi splitting function which universally describes the splitting of a quark into a quark. Eq. (1.39) exhibits a singularity in ϵ which need to be renormalised. In order to do that, we can regard the quark distribution q as a d–dimensional bare distribution and therefore absorb the collinear pole into this unmeasurable quantity defining a renormalised physical object as

$$q^F(x,\mu_F) = q(x,\epsilon) - \frac{\alpha_s}{2\pi}\int_x^1 \frac{d\xi}{\xi}\, q\left(\frac{x}{\xi},\epsilon\right)\left[P_{qq}^{(0)}(\xi)\left(\frac{1}{\epsilon} - \log\left(\frac{\mu^2}{\mu_F^2}\right) + R_{qq}(\xi)\right)\right] \tag{1.40}$$

The second term in Eq. (1.40) is divergent but, as we do when we renormalise the UV divergencies, we can compensate this divergence with the divergence of the bare partonic distribution $q(x,\epsilon)$. In the same equation we introduced the scale μ_F, the factorisation scale by simply splitting the logarithm appearing in (1.39) as

$$\log\frac{Q^2}{\mu^2} = \log\frac{Q^2}{\mu_F^2} + \log\frac{\mu_F^2}{\mu^2}, \tag{1.41}$$

and pushing the unphysical scale μ into the bare parton distributions. We can picture the redefinition of PDFs as follows. As the scale increases the photon starts to "see" evidence for the point-like valence quarks within the proton. If the quarks were non-interacting, no further structure would be resolved increasing the resolving scale: the Bjorken scaling would set in, and the naive parton model would be satisfactory. However, QCD predicts that on increasing the resolution, one should see that each quark is itself surrounded by a cloud of partons. The number of resolved partons which share the proton's momentum increases with the scale.

Of course in Eq. (1.40) there is an arbitrariness in the choice of the finite contribution $R_{qq}(\xi)$: different choices correspond to different factorisation schemes. For instance two popular schemes take two opposite directions: in the $\overline{MS}$ scheme only the singular term plus some coefficients are absorbed into the PDF, while in the DIS schemes all terms are absorbed into the PDF so that $F_2(x,Q^2) = xe_q^2 q^{F,DIS}(x,Q^2)$, namely

$$R_{qq} = -\gamma_E + \log(4\pi) \qquad \overline{MS} \text{ scheme}$$

$$R_{qq} = C_F\left[\frac{1+\xi^2}{1-\xi}\left(\log\left(\frac{1-\xi}{\xi}\right) - \frac{3}{4}\right) + \frac{5\xi+9}{4}\right]_+ \qquad \text{DIS scheme}$$

Therefore PDFs should not be regarded as physical quantities, since they depend on the scheme used to define them. However, their convolution with the appropriate coefficient functions, evaluated in the same factorisation scheme, gives rise to physical, measurable structure functions.

The calculation of the transverse and longitudinal components for the process initiated by a gluon $\gamma^* g \to q\bar{q}$ follows the same lines as that for $\gamma^* q \to qg$. We end up with

the following expressions for the gluon contributions to F_2 and F_1

$$\begin{aligned} F_2^{\gamma g}(x) &= \frac{\alpha_s}{2\pi}\sum_{i=q,\bar{q}} e_i^2 \int_x^1 \frac{d\xi}{\xi} g\left(\frac{x}{\xi}\right)\Bigg\{ -P_{gi}^{(0)}(\xi)\left[\frac{1}{\epsilon} - \gamma_E + \log(4\pi) - \log\left(\frac{Q^2}{\mu^2}\right)\right] \\ &\qquad + T_R\left[[\xi^2+(1-\xi)^2]\log\left(\frac{1-\xi}{\xi}\right) - 1 + 8\xi(1-\xi)\right]\Bigg\} \\ F_1^{\gamma g}(x) &= \frac{F_2^{\gamma q}(x)}{2x} - \frac{\alpha_s}{2\pi} T_R \sum_{i=q,\bar{q}} e_i^2 \int_x^1 \frac{d\xi}{\xi} g\left(\frac{x}{\xi}\right) 4\xi(1-\xi). \end{aligned} \tag{1.42}$$

where $P_{gi}^{(0)}$ is the Altarelli–Parisi splitting function which universally describes the splitting of a gluon into a quark (or antiquark) of flavor i. Analogously to what was done in Eq. (1.40), we can define the finite gluon density as

$$g^F(x,\mu_F) = g(x,\epsilon) - \frac{\alpha_s}{2\pi}\int_x^1 \frac{d\xi}{\xi}\, g\left(\frac{x}{\xi},\epsilon\right)\left[P_{gq}^{(0)}(\xi)\left(\frac{1}{\epsilon} - \log\left(\frac{\mu^2}{\mu_F^2}\right) + R_{qg}(\xi)\right)\right], \tag{1.43}$$

where again the choice for R_{qg} depends on the factorisation scheme.
Generalising the results obtained from the exchange of a photon to the exchange of any vector boson, the NLO formula for $F_1^{(Vh)}$, $F_2^{(Vh)}$ and $F_3^{(Vh)}$ in the $\overline{MS}$ scheme reads

$$\begin{aligned} \frac{F_J(x,Q^2)}{x} &= \int_x^1 \frac{d\xi}{\xi}\sum_{i=q,\bar{q}} g_{Vi}^2 f_i^{\overline{\rm MS}}\left(\frac{x}{\xi},\mu_F^2\right)\Bigg[\delta(1-\xi) + \frac{\alpha_s}{2\pi}\Bigg(P_{qq}^{(0)}(\xi)\log\frac{Q^2}{\mu_F^2} \\ &\qquad + C_J^{Vi,\overline{\rm MS}}(\xi)\Bigg)\Bigg] + g^{\overline{\rm MS}}\left(\frac{x}{\xi},\mu_F^2\right)\frac{\alpha_s}{2\pi}\left[P_{qg}^{(0)}(\xi)\log\frac{Q^2}{\mu_F^2} + C_J^{Vg,\overline{\rm MS}}(\xi)\right], \end{aligned} \tag{1.44}$$

where,

$$
\begin{aligned}
C_1^{Vq,\overline{\rm MS}}(\xi) &= \frac{1}{2}C_2^{Vq,\overline{\rm MS}}(\xi) - C_F\xi, \\
C_2^{Vq,\overline{\rm MS}}(\xi) &= C_F\frac{1}{2}\left[\frac{1+\xi^2}{1-\xi}\left(\log\left(\frac{1-\xi}{\xi}\right) - \frac{3}{4}\right)\right], \\
C_3^{Vq,\overline{\rm MS}}(\xi) &= C_2^{Vq,\overline{\rm MS}}(\xi) - C_F(1+\xi), \\
C_1^{Vg,\overline{\rm MS}}(\xi) &= \frac{1}{2}C_2^{Vg,\overline{\rm MS}}(\xi) - T_F 4\xi(1-\xi), \\
C_2^{Vg,\overline{\rm MS}}(\xi) &= T_F\xi\left[[\xi^2+(1-\xi)^2]\log\left(\frac{1-\xi}{\xi}\right) - 1 + 8\xi(1-\xi)\right], \\
C_3^{Vg,\overline{\rm MS}}(\xi) &= 0, \\
C_i^{Vq,\overline{\rm MS}}(\xi) &= C_i^{V\bar{q},\overline{\rm MS}}(\xi).
\end{aligned}
\tag{1.45}
$$

The last identity is a direct consequence of charge conjugation. The corresponding PDFs in the $\overline{MS}$ scheme at NLO are given by

$$
\begin{aligned}
q^{\overline{MS}}(x,\mu_F^2) &= \sum_{i=q,g}\int_x^1 \frac{d\xi}{\xi} f_i\left(\frac{x}{\xi},\epsilon\right)\left[\delta(1-\xi)\delta_{qi} - \frac{\alpha_s}{2\pi}P_{qi}^{(0)}(\xi)\frac{1}{\epsilon}\left(\frac{4\pi\mu^2}{\mu_F^2 e^{\gamma_E}}\right)^{\epsilon}\right] \\
g^{\overline{MS}}(x,\mu_F^2) &= \sum_{i=q,g}\int_x^1 \frac{d\xi}{\xi} f_i\left(\frac{x}{\xi},\epsilon\right)\left[\frac{\alpha_s}{2\pi}P_{gi}^{(0)}(\xi)\frac{1}{\epsilon}\left(\frac{4\pi\mu^2}{\mu_F^2 e^{\gamma_E}}\right)^{\epsilon}\right].
\end{aligned}
\tag{1.46}
$$

1.1.3 DGLAP evolution equations

The PDFs which appear in Eq. (1.46) are not perturbatively calculable quantities, but one has to extract them from experimental data within a chosen factorisation scheme. Instead the structure functions are physical and measurable quantities which cannot depend on the arbitrary factorisation scale introduced in the redefinition of PDFs. The requirement that the physical cross section does not depend order by order on the factorisation scale, using renormalisation group techniques, leads to the well–known Dokshitzer- Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations [24, 25, 26]. Indeed, setting $\mu = \mu_F$ and differentiating Eq. (1.44) with respect to $\log\mu$, one ends up with an equation for the scale dependence of PDFs of the form

$$
\mu^2\frac{\partial q(x,\mu^2)}{\partial\mu^2} = \frac{\alpha_s}{2\pi}\int_x^1\frac{d\xi}{\xi}\left[P_{qq}^{(0)}(\xi)q\left(\frac{x}{\xi},\mu^2\right) + P_{qg}^{(0)}(\xi)g\left(\frac{x}{\xi},\mu^2\right)\right] + \mathcal{O}(\alpha_s^2), \tag{1.47}
$$

which is the basic form of the DGLAP equations. The above derivation is valid only at the lowest order in perturbation theory, but an all-order proof is possible [27].

The explicit calculation carried out in the previous section shows that, in the redefinition of the quark distribution, one has to include the contributions from the $q \to qg$ and the $g \to q\bar{q}$ splittings. Likewise the gluon involves contributions from the splitting of $q \to gq$, $\bar{q} \to g\bar{q}$ and $g \to gg$. More generally, at any given order in perturbative QCD, one has to consider all possible splittings $a \to b(cd)$ and higher order vertexes. If one takes for instance an initial quark, in the collinear limit, the radiation of $n-$partons out of the quark line needs to be considered,

$$\gamma^* + q \to q + \sum_{i=1}^{n} f_i.$$

A more accurate analysis reveals that collinear divergences arise only from the region in which the transverse momenta of the radiated partons are strongly ordered:

$$|k_{T,n}|^2 \gg |k_{T,n-1}|^2 \gg ... \gg |k_{T,1}|^2.$$

The computation of Feynman diagrams in the collinear limit leads to a squared amplitude proportional to:

$$|\mathcal{M}_n|^2 \sim \frac{(-1)^n}{n!} \left(\frac{\alpha_s}{2\pi}\right)^n \left(\frac{1}{\epsilon}\left(\frac{Q^2}{\mu^2}\right)^{-\epsilon}\right)^n P_{i1,j1} \otimes P_{i2,j2} \otimes ... \otimes P_{in,jn} \otimes C^{(0),Vq},$$

where $\otimes$ is the shorthand notation for the convolution product of Eq. (1.47). Analogously to the case of single emission, the collinear divergences can be absorbed into redefinition of the quark distribution, leaving a logarithmic dependence on the μ scale in the physical cross section:

$$\sigma(Q^2) \sim \left(\frac{\alpha_s}{2\pi}\right)^n \log^n\left(\frac{Q^2}{\mu^2}\right). \tag{1.48}$$

These logarithms are potentially dangerous because for a very large Q^2/μ^2, the product $\alpha_s \log(Q^2/\mu^2)$ might be of $\mathcal{O}(1)$ thus spoiling the perturbative approach.

In what follows, I show that DGLAP equations take care of the problem due to the large collinear logarithms because they resum to all orders them into the evolution of parton distribution functions. In order to show it explicitly, in the following I sketch the solution of the DGLAP equations at leading and next–to–leading orders. In the presence of n_f active quark flavors, where the precise definition of active flavor is going to be given in the following section, Eq. (1.47) is generalised to a system of $(2n_f + 1)$ couple integro–differential equations of the form

$$\frac{d}{dt}\begin{pmatrix} q_i(x,t) \\ g(x,t) \end{pmatrix} = \frac{\alpha_s(t)}{2\pi} \int_x^1 \sum_{j=q,\bar{q}} \frac{d\xi}{\xi} \begin{pmatrix} P_{ij}\left(\frac{x}{\xi}, \alpha_s(t)\right) & P_{ig}\left(\frac{x}{\xi}, \alpha_s(t)\right) \\ P_{gj}\left(\frac{x}{\xi}, \alpha_s(t)\right) & P_{gg}\left(\frac{x}{\xi}, \alpha_s(t)\right) \end{pmatrix} \otimes \begin{pmatrix} q_j(\xi,t) \\ g(\xi,t) \end{pmatrix},$$

(1.49)

where $t = \log \frac{Q^2}{\mu^2}$ and $\alpha_s(t)$ is the running coupling constant. Although PDFs are non-perturbative objects, the evolution kernels P_{ij} can be computed in perturbation theory:

$$P_{ij}(x, \alpha_s(t)) = \sum_{n=0}^{\infty} \left(\frac{\alpha_s(t)}{2\pi}\right)^n P_{ij}^{(n)}(x). \tag{1.50}$$

The coefficients of the expansion have been calculated perturbatively at next-to-leading order [28] and more recently they have been computed up to the NNLO in α_s [29, 30, 31]. At the leading-order they read

$$\begin{aligned}
P_{qq}^{(0)}(x) &= C_F \left[\frac{(1+x^2)}{(1-x)_+} + \frac{3}{2}\delta(1-x)\right], \\
P_{qg}^{(0)}(x) &= T_R \left[x^2 + (1-x)^2\right], \\
P_{gq}^{(0)}(x) &= C_F \left[\frac{1+(1-x)^2}{x}\right], \\
P_{gg}^{(0)}(x) &= 2N \left[\frac{x}{(1-x)_+} + \frac{1-x}{x} + x(1-x)\right] + \delta(1-x)\frac{(11N - 4n_f T_R)}{6}.
\end{aligned} \tag{1.51}$$

The leading order splitting functions can be interpreted as the probability of finding a parton of type i in a parton of type j with a fraction ξ of the longitudinal momentum of the parent parton and a transverse momentum squared much less than μ^2. The interpretation as probabilities implies that the splitting functions are positive definite for $x < 1$, and satisfy the sum rules

$$\begin{aligned}
&\int_0^1 dx\, P_{qq}(x) = 0, \\
&\int_0^1 dx\, x\, [P_{qq}(x) + P_{gq}(x)] = 0, \\
&\int_0^1 dx\, x\, [2n_f P_{qg}(x) + P_{gg}(x)] = 0,
\end{aligned} \tag{1.52}$$

which correspond to quark number conservation and momentum conservation in the splittings of the quark ad the gluon respectively.
The structure of the solution of these coupled integro-differential equations may be

written as

$$f_i(x,Q^2) = \sum_j \Gamma_{ij}(x,\alpha_s,\alpha_s^0) \otimes f_j(x,Q_0^2), \tag{1.53}$$

where $f_j(x,Q_0^2)$ are the input PDFs, to be determined empirically, $\Gamma_{ij}(x,\alpha_s,\alpha_s^0)$ are the evolution factors. From now on, I use the shorthand notation

$$\alpha_s \equiv \alpha_s(Q^2), \qquad \alpha_s^0 \equiv \alpha_s(Q_0^2). \tag{1.54}$$

The evolution factors also satisfy evolution equations:

$$Q^2 \frac{\partial}{\partial Q^2} \Gamma_{ij}(x,\alpha_s,\alpha_s^0) = \sum_k P_{ik}(x,\alpha_s) \otimes \Gamma_{kj}(x,\alpha_s,\alpha_s^0), \tag{1.55}$$

with boundary conditions $\Gamma_{ij}(x,\alpha_s^0,\alpha_s^0) = \delta_{ij}\delta(1-x)$.
An efficient method to solve the DGLAP equations consists in defining particular linear combinations of the individual quark distributions in such a way that the $(2n_f+1)$ equations (1.49) maximally decouple from each others. Indeed, from considerations based on charge conjugation and flavour symmetry, it is possible to write down n_f non–singlet (NS) flavour combinations which evolve independently [32]

$$\begin{aligned} q^{\pm}_{NS,ij} &= q_i^{\pm} - q_j^{\pm}, \\ q^{v}_{NS} &= \sum_{i=1}^{n_f} q_i^{-}, \end{aligned} \tag{1.56}$$

where $q_i^{\pm} = q_i \pm \overline{q}_i$. Only one combination of flavors is left which couples to the gluon, the singlet combination, Σ, defined as

$$\Sigma = \sum_{i=1}^{n_f} (q_i + \overline{q}_i)\,. \tag{1.57}$$

A convenient basis, in presence of $n_f = 6$ active flavors, can be explicitly written as

$$\begin{aligned}
V &= \sum_{i=1}^{6} q_i^-, \qquad (1.58)\\
V_3 &= u^- + d^-,\\
V_8 &= u^- + d^- - 2s^-,\\
V_{15} &= u^- + d^- + s^- - 3c^-,\\
V_{24} &= u^- + d^- + s^- + c^- - 4b^-,\\
V_{35} &= u^- + d^- + s^- + c^- + b^- - 5t^-,\\
T_3 &= u^+ - d^+,\\
T_8 &= u^+ + d^+ - 2s^+,\\
T_{15} &= u^+ + d^+ + s^+ - 3c^+,\\
T_{24} &= u^+ + d^+ + s^+ + c^+ - 4b^+,\\
T_{35} &= u^+ + d^+ + s^+ + c^+ + b^+ - 5t^+,
\end{aligned}$$

where V is the total valence, V_i are q_{NS}^-–like combinations and T_i are q_{NS}^+–like combinations of the flavour distributions. In this basis, the DGLAP equations are simplified and, using the short–hand notation for the convolution product, they read

$$\frac{d}{dt} q_{NS}^{\pm,v}(x,t) = P_{NS}^{\pm,v} \otimes q_{NS}^{\pm,v}(x,t) \qquad (1.59)$$

$$\frac{d}{dt} \begin{pmatrix} \Sigma \\ g \end{pmatrix}(x,t) = \begin{pmatrix} P_{qq} & 2n_f P_{qg} \\ P_{gq} & P_{gg} \end{pmatrix} \otimes \begin{pmatrix} \Sigma \\ g \end{pmatrix}(x,t). \qquad (1.60)$$

The combinations V_i and T_j evolve according to Eq. (1.59) with P_{NS}^- and P_{NS}^+ respectively, while the total valence V evolves with P_{NS}^v. At LO $P_{NS}^{(0),+} = P_{NS}^{(0),-} = P_{NS}^{(0),v} = P_{qq}^{(0)}$. At NLO $P_{NS}^{(0),-} = P_{NS}^{(0),v}$ while the other splitting functions are different. Starting from $\mathcal{O}(\alpha_s^2)$ all splitting functions are different from each others. Having solved the equations for the $q^{\pm,v}$ combinations, it is straightforward to invert the linear system Eq. (1.58) and obtain the individual PDFs evolved at any scale Q^2.

Before solving the DGLAP equations it is useful to introduce a technical tool. The expression for a physical observable consists of a convolution between a hard coefficient function and parton distribution functions as in Eq. (1.44). A theorem states that the convolution product of two functions g and h can be turned into an ordinary product by taking Mellin moments of the functions. Suppose $f = g \otimes h$, then the Mellin transforms of f is simply given by the ordinary product of the Mellin transform of g

and h

$$M[f] = M[g]M[h],$$

where the Mellin transform of a function f is defined as

$$M[f][N] = \int_0^1 dy\, y^{N-1} f(y). \tag{1.61}$$

In Mellin space, the evolution equation for the evolution factor, Eq. (1.55), can be written as

$$Q^2 \frac{\partial}{\partial Q^2} \Gamma_{ij}(N, \alpha_s, \alpha_s^0) = \sum_k \gamma_{ik}(N, \alpha_s)\, \Gamma_{kj}(N, \alpha_s, \alpha_s^0)\ , \tag{1.62}$$

where $\Gamma_{ij}(N, \alpha_s, \alpha_s^0)$ are the Mellin moments of the x–space evolution factor

$$\Gamma_{ij}(N, \alpha_s, \alpha_s^0) \equiv \int_0^1 dx\, x^{N-1} \Gamma_{ij}(x, \alpha_s, \alpha_s^0)\ , \tag{1.63}$$

and γ_{ij} are the anomalous dimensions, defined as the Mellin moments of the splitting functions $P_{ij}(x)$:

$$\gamma_{ij}(N, \alpha_s(t)) = \int_0^1 dx\, x^{N-1} \frac{\alpha_s(t)}{2\pi}\, x\, P_{ij}(x, \alpha_s(t)). \tag{1.64}$$

Like the splitting functions, also the anomalous dimensions can be expanded in series of α_s as

$$\gamma_{ij}(N, \alpha_s(t)) = \sum_{n=0}^{\infty} \left(\frac{\alpha_s}{2\pi}\right)^n \gamma_{ij}^{(n)}(N). \tag{1.65}$$

Since splitting functions (and therefore anomalous dimensions) depend on the scale only through the coupling constant, an easy way to solve Eq. (1.49) consists in differentiating the above equation with respect to α_s and rewrite it as

$$\alpha_s \frac{\partial}{\partial \alpha_s} \Gamma_{ij}(N, \alpha_s, \alpha_s^0) = -\sum_k R_{ik}(N, \alpha_s) \Gamma_{kj}(N, \alpha_s, \alpha_s^0). \tag{1.66}$$

where the matrix $R_{ij} \equiv (\mathbf{R})_{ij}$ has the perturbative expansion

$$\mathbf{R}(N, \alpha_s) = \mathbf{R}_0(N) + \alpha_s \mathbf{R}_1(N) + \mathcal{O}(\alpha_s^2). \tag{1.67}$$

The matrices in the expansion of R_{ij} are recursively defined in terms of the coefficients in the perturbative expansion of the anomalous dimensions as

$$\mathbf{R}_0 \equiv \frac{\boldsymbol{\gamma}^{(0)}}{\beta_0} \qquad \mathbf{R}_k \equiv \frac{\boldsymbol{\gamma}^{(k)}}{\beta_0} - \sum_{i=1}^{k} b_1 R_{k-i}\,, \tag{1.68}$$

where $b_1 \equiv \beta_1/\beta_0$ and β_i are the coefficients of the QCD $\beta-$function defined in Eq. (5).

The complete matrix of anomalous dimensions $\boldsymbol{\gamma}$, and thus the matrices $\mathbf{R}$ are in fact almost completely diagonal: all the flavor nonsinglet and valence quark distributions evolve multiplicatively, and only the singlet quark and gluon actually mix. Thus we only need to solve Eq. (1.66) for the non–singlet scalar evolution factors and the two by two singlet evolution matrix.

We consider first the simplest case of the evolution of flavor nonsinglet distributions: the evolution factor then satisfies the simple first order equation

$$\frac{\partial}{\partial \ln \alpha_s}\Gamma_{\rm NS}(N,\alpha_s,\alpha_s^0) = -R_{\rm NS}(N,\alpha_s)\Gamma_{\rm NS}(N,\alpha_s,\alpha_s^0). \tag{1.69}$$

At LO the solution is trivial:

$$\Gamma_{\rm NS,LO}(N,\alpha_s,\alpha_s^0) = \left(\frac{\alpha_s}{\alpha_s^0}\right)^{-R_{\rm NS}^{(0)}}, \tag{1.70}$$

while at NLO we need to work a little harder: using Eq. (1.68) one finds

$$\Gamma_{\rm NS,NLO}(N,\alpha_s,\alpha_s^0) = \exp\left\{-\frac{R_{\rm NS}^{(1)}}{b_1}\ln\left(\frac{1+b_1\alpha_s}{1+b_1\alpha_s^0}\right)\right\}\left(\frac{\alpha_s}{\alpha_s^0}\right)^{-R_{\rm NS}^{(0)}}. \tag{1.71}$$

This exact solution is equivalent up to subleading terms to the linearised solution

$$\Gamma^{\rm lin}_{\rm NS,NLO}(N,\alpha_s,\alpha_s^0) = \left(1 - R^{(1)_{\rm NS}}(\alpha_s-\alpha_s^0)\right)\left(\frac{\alpha_s}{\alpha_s^0}\right)^{-R_{\rm NS}^{(0)}}, \tag{1.72}$$

which is in turn the exact solution to Eq. (1.69) with $R_{\rm NS} = R_{\rm NS}^{(0)} + \alpha_s R_{\rm NS}^{(1)}$.

Turning to the singlet sector, we need to solve Eq. (1.66) when $\mathbf{R}$ are two by two matrices, corresponding to coupled singlet quarks and gluons:

$$\frac{\partial}{\partial \ln \alpha_s}\boldsymbol{\Gamma}_{\rm S}(N,\alpha_s,\alpha_s^0) = -\mathbf{R}_{\rm S}(N,\alpha_s)\boldsymbol{\Gamma}_{\rm S}(N,\alpha_s,\alpha_s^0). \tag{1.73}$$

At LO we can proceed by diagonalisation:

$$\mathbf{\Gamma}_{\rm S,LO}(N,\alpha_s,\alpha_s^0) \equiv \mathbf{L}(N,\alpha_s,\alpha_s^0) = \mathbf{e}_+(N)\left(\frac{\alpha_s}{\alpha_s^0}\right)^{-\lambda_+(N)} + \mathbf{e}_-(N)\left(\frac{\alpha_s}{\alpha_s^0}\right)^{-\lambda_-(N)}, \tag{1.74}$$

where

$$\begin{aligned}\lambda_\pm(N) &= \frac{1}{2\beta_0}\Bigg[\gamma_{qq}^{(0)}(N)+\gamma_{gg}^{(0)}(N) \\ &\quad \pm\sqrt{\left(\gamma_{qq}^{(0)}(N)-\gamma_{gg}^{(0)}(N)\right)^2+4\gamma_{qg}^{(0)}(N)\gamma_{gq}^{(0)}(N)}\Bigg]\end{aligned} \tag{1.75}$$

are the eigenvalues of the two by two matrix $\mathbf{R}_{\rm S}^{(0)}(N)$ of singlet anomalous dimensions, and

$$\mathbf{e}_\pm(N) = \pm\frac{1}{\lambda_+(N)-\lambda_-(N)}(\mathbf{R}_{\rm S}^{(0)}(N)-\lambda_\mp(N)\mathbf{I}), \tag{1.76}$$

are the corresponding projectors. The full NLO solution is more complicated, and must be developed recursively as a perturbative expansion around the LO solution $\mathbf{L}(N,as,\alpha_s^0)$: writing

$$\mathbf{\Gamma}_{\rm S,NLO}(N,\alpha_s,\alpha_s^0) \equiv \mathbf{U}(N,\alpha_s)\mathbf{L}(N,\alpha_s,\alpha_s^0)\mathbf{U}(N,\alpha_s^0)^{-1}, \tag{1.77}$$

where $\mathbf{U}(N,\alpha_s)$

$$\mathbf{U}(N,\alpha_s) = \mathbf{1}+\alpha_s\mathbf{U}^{(1)}(N)+\alpha_s^2\mathbf{U}^{(2)}(N)+\cdots, \tag{1.78}$$

solves Eq. (1.66) provided

$$\mathbf{U}^{(k)} = \frac{\mathbf{e}_-\widetilde{\mathbf{R}}^{(k)}\mathbf{e}_+}{\lambda_+-\lambda_--k}+\frac{\mathbf{e}_+\widetilde{\mathbf{R}}^{(k)}\mathbf{e}_-}{\lambda_--\lambda_+-k}-\frac{1}{k}\left[\mathbf{e}_+\widetilde{\mathbf{R}}^{(k)}\mathbf{e}_+ + \mathbf{e}_-\widetilde{\mathbf{R}}^{(k)}\mathbf{e}_-\right], \tag{1.79}$$

where

$$\widetilde{\mathbf{R}}^{(0)} = \mathbf{R}_{\rm S}^{(0)}, \qquad \widetilde{\mathbf{R}}^{(k)} = \mathbf{R}_{\rm S}^{(k)}+\sum_{i=1}^{k-1}\mathbf{R}_{\rm S}^{(k-i)}\mathbf{U}^{(i)}. \tag{1.80}$$

By solving recursively Eqs. (1.79, 1.80) with the NLO approximation in Eq. (1.68), the NLO evolution factor can be computed. Just as in the nonsinglet case, the exact

NLO solution may be linearised to give

$$\begin{aligned}\mathbf{\Gamma}^{\rm lin}_{\rm S,NLO}(N,\alpha_s,\alpha_s^0) &= \mathbf{L}(N,\alpha_s,\alpha_s^0) \\ &\quad +\alpha_s\mathbf{U}^{(1)}(N)\mathbf{L}(N,\alpha_s,\alpha_s^0)-\alpha_s^0\mathbf{L}(N,\alpha_s,\alpha_s^0)\mathbf{U}^{(1)}(N),\end{aligned} \tag{1.81}$$

which again is an exact solution to the truncated evolution equation, and equivalent to the full solution Eq. (1.77) up to subleading terms.

Finally we want to see explicitly that the problem of the large collinear logarithms is solved by the DGLAP evolution. Looking at Eq. (1.70), we can easily verify this. Setting $Q_0^2=\mu^2$, it is straightforward to show that all the leading logarithms of Q^2/μ^2 are resummed in the evolution factor:

$$\begin{aligned}\Gamma_{LO}(N) &= \exp\left[\frac{-\gamma^{(0)}(N)}{\beta_0}\log\frac{\alpha_s(Q^2)}{\alpha_s(\mu^2)}\right] \\ &= \exp\left[\frac{-\gamma^{(0)}(N)}{\beta_0}\log\left(1-\alpha_s(Q^2)\beta_0\log\frac{Q^2}{\mu^2}\right)\right] \\ &= \exp\left[\alpha_s(Q^2)\gamma^{(0)}(N)\log\frac{Q^2}{\mu^2}\right] \\ &= \sum_{n=0}^{\infty}\left(\alpha_s\log\frac{Q^2}{\mu^2}\gamma^{(0)}(N)\right)^n,\end{aligned} \tag{1.82}$$

where we expanded in α_s and used the LO running of α_s, Eq. (5). If higher terms in the perturbative expansion of the anomalous dimension were included, also subleading logarithms would be resummed. The inclusion of the $p-$th order in the expansion of the anomalous dimension enables use to perform the resummation of N^{k}LO logarithms:

$$\sim\sum_{n=0}^{\infty}\sum_{k=0}^{p}\alpha_s^{n+k}(t)\log^n\frac{Q^2}{\mu^2}\gamma^{(k)}(N).$$

1.1.4 Collinear factorisation Theorem

In the previous sections deep inelastic scattering at one-loop was discussed, showing how it is possible to absorb the singularities arising from the emission of collinear partons into a redefinition of the parton densities. The factorisation theorem of collinear singularities states that it possible to write the hadronic cross section as a convolution between a partonic, process dependent, coefficient function and universal parton distributions. Corrections to the leading–twist factorisation are suppressed by powers

of Q^2. In QCD, structure functions have 'higher–twist' power corrections, which are difficult to estimate quantitatively

$$F_i(x,Q^2) = F_i^{(2)}(x,Q^2) + \frac{F^{(4)}(x,Q^2)}{Q^2} + \dots . \tag{1.83}$$

In the above equation, $^{(n)}$ refers to the twist, defined as "dimension minus spin", of the contributing operators. The factorisation of Eq. (1.91) includes only the twist–2 operator. The power corrections due to higher order twists are visible at low–Q^2 structure functions data.

A rigorous proof of factorisation to all orders exists for deep inelastic scattering in the context of the operator product expansion (for a review see Ref. [33]). In hadron–hadron collisions, the analysis is more complicated since the question arises whether the partons in hadron h_1, through the influence of their colour fields, change the distribution of partons in hadron h_2, thus spoiling the simple parton picture. Factorisation of the cross section into a pure short-distance contribution, computable in perturbation theory and non-perturbative, but universal, parton distribution functions is more complicated because of these colour correlations. Nevertheless it can be proved to all orders for sufficiently inclusive observables [27].

As an illustrative example of how factorisation works also in inclusive hadronic processes we consider the Drell–Yan production of a W boson. The Drell–Yan process is the production of a high–mass lepton pair from the decay of the electroweak boson produced in a hadron–hadron collision. Originally the most frequent Drell–Yan event corresponded to the creation of e^+e^- or $\mu^+\mu^-$ pairs from the decay of a virtual photon but, as energy of collisions increased, it includes the contribution from Z exchange and also $e\nu_e$ and $\mu\nu_\mu$ pairs coming from $W^\pm$ decays. For the full computation at NLO of the $W^\pm$ production, see Ref. [34]. Here we just sketch the calculation in order to show the power of the factorisation theorem.
In the parton model picture the cross–section is simply given by

$$\begin{aligned}\frac{d\sigma^{(0)}}{dM_W^2}(h_1,h_2 \to W^\pm) &= \int_0^1 dx_1 \int_0^1 dx_2 \sum_{i,j=1}^{n_f} g_{ij,W}\, f_i(x_1)\, \bar{f}_j(x_2) \\ &\quad \frac{d\hat{\sigma}^{(0)}}{dM_W^2}(f_1\hat{f}_2 \to W^\pm)\left(M_W^2,\hat{\tau}\right)\end{aligned} \tag{1.84}$$

where i, j runs over $U, \bar{D}$ in case of W^+ production and $D, \bar{U}$ in case of W^- production and $d\hat{\sigma}$ is the partonic cross–section given by

$$\frac{d\hat{\sigma}^{(0)}}{dM_W^2}(f_1\hat{f}_2 \to W^\pm) = \frac{\pi}{3}\sqrt{2}G_F|V_{ij}|^2\delta(\hat{s} - M_W^2) \equiv \frac{\hat{\sigma}^{(0)}_{W^\pm}}{M_W^2}\delta(1-\hat{\tau}). \quad (1.85)$$

In the above equations $\hat{\tau} = M_W^2/\hat{s}$ and $\hat{s} = x_1x_2S$. Plugging Eq. (1.85) into Eq. (1.84) gives

$$\frac{d\sigma^{(0)}}{dM_W^2}(h_1, h_2 \to W^\pm) = \frac{\hat{\sigma}^{(0)}_{W^\pm}}{M_W^2}\hat{\tau}\int_{\hat{\tau}}^1 \frac{dx}{x}\sum_{i,j=1}^{n_f} g_{ij,W}\, f_i(x)\, \bar{f}_j(\hat{\tau}/x) \quad (1.86)$$

At the next–to–leading order there are basically two contributions. There is the gluon bremsstrahlung correction to the lowest order process. The collinear singularity arises when the gluon becomes parallel to one of the incoming quark or anti–quark. Because of the KLM theorem, we know that soft singularities will be cancelled by the virtual correction to the lowest order process. There is also the gluon initiated process $gq \to W^\pm q'$. This will contain a collinear singularity when the scattered quark is parallel to the incoming quark in the centre–of–mass frame. The calculation of the Feynman amplitudes for these process is simplified by applying crossing symmetry to the DIS matrix elements calculated above. The phase space is different, but it can be easily evaluated, ending up with the following results:

$$\begin{aligned}
\frac{d\hat{\sigma}^{(1)}}{dM_W^2}(q\bar{q} \to W^\pm g) &= \frac{\hat{\sigma}^{(0)}_{W^\pm}}{M_W^2}\frac{\alpha_s}{2\pi}\left[-2P_{qq}^{(0)}(\hat{\tau})\left(\frac{1}{\epsilon} - 4\pi - \gamma_E + \log\left(\frac{Q^2}{\mu^2}\right)\right) + C_{qq'}(\hat{\tau})\right] \\
\frac{d\hat{\sigma}^{(1)}}{dM_W^2}(qg \to W^\pm q) &= \frac{\hat{\sigma}^{(0)}_{W^\pm}}{M_W^2}\frac{\alpha_s}{2\pi}\left[-P_{qg}^{(0)}(\hat{\tau})\left(\frac{1}{\epsilon} - 4\pi - \gamma_E + \log\left(\frac{Q^2}{\mu^2}\right)\right) + C_{qg}(\hat{\tau})\right],
\end{aligned} \quad (1.87)$$

where we encounter the collinear $1/\epsilon$ pole accompanied by the finite coefficients that we already found in Eqs. (1.39, 1.42). The coefficient functions $C_{qq'}$ and C_{qg} can be explicitly written as

$$\begin{aligned}
C_{qq'}(z) &= C_F\left[4(1+z^2)\left(\frac{\log(1-z)}{1-z}\right)_+ - 2\frac{1+z^2}{1-z}\log z + \left(\frac{2\pi^2}{3} - 8\right)\delta(1-z)\right], \\
C_{qg}(z) &= T_R\left[2[z^2 + (1-z)^2]\log\left(\frac{(1-z)^2}{z}\right) + 3 + 2z - 3z^2\right].
\end{aligned} \quad (1.88)$$

The NLO cross section for hadron–hadron collisions may be written as

$$\begin{aligned}\frac{d\sigma^{(1)}}{dM_W^2}(h_1,h_2\to W^+) &= \int_0^1 dx_1\int_0^1 dx_2\, g_{ij,W} \\ &\sum_{i,j=1}^{n_f}\left[f_i(x_1)\,\bar f_j(x_2)+\bar f_i(x_1)\,f_j(x_2)\right] \\ &\left[\frac{d\hat\sigma^{(0)}}{dM_W^2}(f_1\bar f_2\to W^+)+\frac{d\hat\sigma^{(1)}}{dM_W^2}(f_1\bar f_2\to W^+ g)\right]+ \\ &\sum_{i=1}^{n_f}\left[f_i(x_1)\,g(x_2)+g(x_1)\,f_i(x_2)\right]\,\frac{d\hat\sigma^{(1)}}{dM_W^2}(fg\to W^+ f).\end{aligned} \tag{1.89}$$

In case of W^- production, the structure is the same, where $f\to\bar f$. The structure of Eq. (1.90) is very similar to the corresponding expressions that arose in the NLO description of the DIS structure functions, Eqs. (1.39, 1.42). Introducing the $\overline{MS}$ PDFs defined in the previous section for the DIS process in Eq. (1.46), we obtain a finite expression for the NLO W^+ production

$$\begin{aligned}\frac{d\sigma^{(1)}}{dM_W^2} &= \frac{\hat\sigma^{(0)}_{W^+}}{M_W^2}\int_\rho^1\frac{dx_1}{x_1}\int_{x1\rho}^1\frac{dx_2}{x_2}\sum_{i,j=1}^{n_f} g_{ij,W} \\ &\left[f_i^{\overline{MS}}(x_1,\mu_F^2)\,\bar f_j^{\overline{MS}}(x_2,\mu_F^2)+\bar f_i^{\overline{MS}}(x_1,\mu_F^2)\,f_j^{\overline{MS}}(x_2,\mu_F^2)\right] \\ &\left[\delta(1-\hat\tau)+\frac{\alpha_s}{2\pi}\left(2P_{qq}^{(0)}(\hat\tau)\log\left(\frac{Q^2}{\mu_F^2}\right)+C_{qq'}(\tau)\right)\right]+ \\ &\sum_{i=1}^{n_f} g_{ij,W}\left[f_i^{\overline{MS}}(x_1,\mu_F^2)\,g_j^{\overline{MS}}(x_2,\mu_F^2)+g^{\overline{MS}}(x_1,\mu_F^2)+f_i^{\overline{MS}}(x_2,\mu_F^2)\right] \\ &\left[\frac{\alpha_s}{2\pi}\left(2P_{qg}^{(0)}(\hat\tau)\log\left(\frac{Q^2}{\mu_F^2}\right)+C_{qg}(\tau)\right)\right]\end{aligned} \tag{1.90}$$

where ρ is the ratio between the hard scale of the process M_W^2 and the centre-of-mass energy s. The power of the factorisation is that essentially the separation of the cross section into long–distance scale–dependent PDFs and short–distance coefficient functions will treat in the same way the collinear singularities arising at NLO, independently of the considered process.

1.2 Heavy quarks

In the previous section we have seen that factorisation theorem enables one to write a wide class of relevant cross–sections characterised by a large scale Q^2 in terms of perturbatively calculable quantities and a limited set of non perturbative quantities that must be obtained from experiments. In this section I treat the case of cross–sections involving heavy quarks. Since in the framework of perturbative QCD we model the hadron as an object constituted by massless partons, calculations of processes involving heavy quarks require particular care. They are typically handled by combining perturbative QCD calculations with light quarks in the initial states and calculations of heavy particles in the final states. The combination of the two ingredients depends on the mass of the heavy quark with respect to the other scales involved in the considered process.

Processes involving heavy quarks are a good example of multi–scale processes. To be precise, we define a heavy quark to be one whose mass M_Q is large enough that the effective coupling at the scale of a heavy quark mass $\alpha_s(M_Q)$ is in the perturbative region. With this definition, the charm, bottom and top quarks are heavy quarks. If n_l is the number of light quarks, and n_f the total number of quarks, in the present state of knowledge of QCD, $n_l = 3$ and $n_f = 6$.

Whenever a cross–section is characterised by two scales, say the hard scale Q^2 and the mass of a heavy quark M_Q^2, perturbative calculations typically present both logarithms and powers of the ratio of the scales M_Q^2/Q^2. These contributions may spoil the accuracy of the calculation. As we have seen in the previous section, defining a heavy quark parton density would automatically resum the logarithms to all orders in perturbative QCD via the DGLAP evolution equations. However this approach is only adequate at high energies where the mass of the heavy quark is small with respect to the typical scale of the process. In a small or intermediate regimes, it ignores power suppressed contributions. On the other hand, a calculation where the heavy quark appears only as a final state ignores the logarithms and it is valid only at a relatively small scale. This kind of situation requires an adequate treatment for the mass of the heavy quark which ought to be valid in all cases, when $Q^2 \ll M^2$ as well as when $Q^2 \sim M^2$ and $Q^2 \gg M^2$.

In this section I first present the Collins–Wilczek and Zee (CWZ) renormalisation scheme [35], which shows explicit decoupling of the heavy quarks at low energies. Then I give some details about its application to QCD and discuss the matching conditions between different sub–schemes associated to different number of active partons. To conclude, I mention the main characteristics of the schemes proposed to deal with heavy quark masses in perturbative calculations.

1.2.1 CWZ renormalisation scheme

A detailed explanation of the renormalisation scheme proposed by Collins–Wilczek and Zee [35] is given in Chapter 8 of Ref. [36]. In this section I summarise the toy model described in Ref. [36] in order to show that the CWZ scheme exhibits explicit decoupling. Later I show that the same scheme may be applied to QCD in presence of one or more heavy quarks [37].

Let us consider a Φ^3 scalar theory with two fields in $d = 6$ dimensions. Let ϕ be a light– field of mass m and Φ a heavy–field of mass M. If we impose the symmetry under $\Phi \to -\Phi$, the Lagrangian density of the theory can be written as

$$\mathcal{L} = \frac{(\partial\phi)^2}{2} + \frac{(\partial\Phi)^2}{2} - m^2\frac{\phi^2}{2} - M^2\frac{\Phi^2}{2} - \mu^{3-d/2}\left[g_1\frac{\phi^3}{6} + g_2\frac{\phi\Phi^2}{2}\right] + \mathcal{L}_{c.t.}, \quad (1.91)$$

where $\mathcal{L}_{c.t.}$ includes the counterterms which cancel the divergences of the bare quantities of the theory. It can be explicitly written as

$$\begin{aligned}\mathcal{L}_{c.t.} &= (Z_l - 1)\frac{(\partial\phi)^2}{2} + (Z_h - 1)\frac{(\partial\Phi)^2}{2} - [m^2(Z_m - 1) + M^2 Z_{mM}]\frac{\phi^2}{2} \quad (1.92)\\ &\quad -[M^2(Z_M - 1) + m^2 Z_{Mm}]\frac{\Phi^2}{2} - \mu^{3-d/2}\left[(g_{1B} - g_1)\frac{\phi^3}{6} + (g_{2B} - g_2)\frac{\phi\Phi^2}{2}\right].\end{aligned}$$

In the above Lagrangian density I have ignored the linear term that is typically introduced in order to cancel the tadpoles.

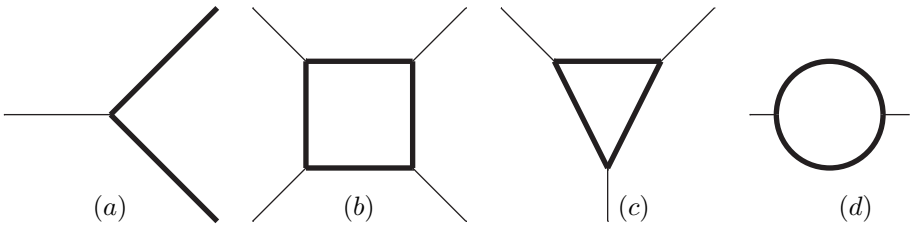
According to the decoupling theorem, phenomena on energy scales much less than M are described by an effective low–energy theory whose Lagrangian has the form

$$\mathcal{L}_{\rm eff} = z\frac{(\partial\phi)^2}{2} - m^{*2}z\frac{\phi^2}{2} - \mu^{3-d/2}g^* z^{3/2}\frac{\phi^3}{6} + \mathcal{L}_{\rm eff,c.t.}, \quad (1.93)$$

where the counterterms Lagrangian has a similar structure as the one of Eq. (1.92) with only light fields included.

To prove the decoupling theorem, we have to show that m^*, g^* and z might be chosen so that the Green functions of ϕ obtained from the effective Lagrangian differ from those obtained in the full Lagrangian only by terms $\mathcal{O}(p/M)$, where p are the external momenta to the Green functions. The $\overline{MS}$ renormalisation scheme does not correspond to the right choice. On the contrary a zero–momentum subtraction scheme (BPH) exhibits explicit decoupling. In the proof we consider only the LO (a) and NLO diagrams (b,c,d) shown below. The argument can be generalised to higher orders [36]. We consider only those Feynman amplitudes with external light fields and internal heavy fields, because those are the only ones which might contribute to the

renormalisation of the couplings of the effective Lagrangian with terms depending on the mass of the heavy fields which are present only in the full Lagrangian.



It is straightforward to show that, for $M^2 \to \infty$ as the external momenta are fixed, the graphs (a) and (b) are suppressed by powers of M^2. Indeed, in graph (a) each of the n internal heavy quark line contributes with a propagator $\sim 1/(p^2 - M^2)$ which, in the $M^2 \to \infty$ limit gives a power–suppressed contribution $\sim 1/(M^2)^n$. Analogously, in the diagram (b), the four point loop amplitude is not divergent and, in the limit $M^2 \to \infty$, we can ignore the momenta of the external particles ending up with a contribution

$$\Gamma_2 = g_2^4 \int \frac{d^6k}{(2\pi)^6} \frac{1}{(k^2 - M^2)^4} \propto \frac{1}{M^2}$$

which vanishes as $M^2 \to \infty$. If we were considering its leading contribution, it would be an effective four–point vertex. The non–renormalisability of the coupling is tied to the negative power of M^2.

Diagrams (c) and (d) are divergent. The divergent three–point function (c), being p_1, p_2 and p_3 the incoming momenta and k the loop momentum, can be written in $d = 6 - 2\epsilon$ dimensions as

$$\begin{aligned}
\Gamma_3 &= g_2^3 \int \frac{d^dk}{(2\pi)^d} \frac{\mu^{9-3d/2}}{[(k+p_1)^2 - M^2 + i\varepsilon][(k-p_2)^2 - M^2 + i\varepsilon][k^2 - M^2 + i\varepsilon]} \qquad (1.94)\\
&= 2ig_2^3 \int_0^1 dx \int_0^{1-x} dy \int \frac{d^dk}{(2\pi)^d} \frac{\mu^{9-3d/2}}{[k^2 - M^2 + x(p_1^2 + 2k\dot{p}_1) + y(p_2^2 - 2k\cdot p_2) + i\varepsilon]^3}\\
&= g_2^3 \frac{\Gamma(\epsilon)}{(4\pi)^2} \mu^\epsilon \int_0^1 dx \int_0^{1-x} dy \left(\frac{M^2 - p_1^2 x(1-x) - p_2^2 y(1-y) - 2xyp_1p_2}{4\pi\mu^2} \right)^\epsilon .
\end{aligned}$$

The divergence in Eq. (1.94) may be normalised in several ways. In the $\overline{MS}$ scheme one subtracts only the pole in $1/\epsilon$ and the associated $\gamma_E + \log(4\pi)$ finite terms. The key point is that the subtraction in the $\overline{MS}$ scheme is independent of the mass of the particles involved in the process. In this way the renormalised three–point function is

given by

$$R^{\overline{MS}}(\Gamma_3) = i\frac{g_2^3}{(4\pi)^2}\left[\log\left(\frac{M^2}{\mu^2}\right) + \mathcal{O}\left(\frac{p_i^2}{M^2}\right)\right]. \tag{1.95}$$

For large M this gives a logarithmic correction to the effective coupling constant according to

$$z^{3/2}g^* = g_1 - \frac{g_2^3}{2(4\pi)^2}\log\left(\frac{M^2}{\mu^2}\right) + \mathcal{O}(g^5). \tag{1.96}$$

Instead, in the zero–momentum subtraction scheme, we get

$$\begin{aligned} R^{ZM}(\Gamma_3) &= g_2^3\frac{\Gamma(\epsilon)}{(4\pi)^2}\mu^\epsilon\int_0^1 dx\int_0^{1-x} dy \\ &\quad\left[\left(\frac{M^2 - p_1^2x(1-x) - p_2^2y(1-y) - 2xyp_1p_2}{4\pi\mu^2}\right)^\epsilon - \left(\frac{M^2}{4\pi\mu^2}\right)^\epsilon\right] \\ &= g_2^3\frac{\mu^\epsilon}{(4\pi)^2}\int_0^1 dx\int_0^{1-x} dy \\ &\quad\left(1 - \frac{p_1^2}{M^2}x(1-x) - \frac{p_2^2}{M^2}y(1-y) - 2xy\frac{p_1p_2}{M^2}\right) = \mathcal{O}\left(\frac{p^2}{M^2}\right) \end{aligned} \tag{1.97}$$

and therefore the renormalised effective coupling would not depend on the mass of the heavy fields if not by power–suppressed terms

$$z^{3/2}g^* = g_1 + \mathcal{O}(p^2/M^2). \tag{1.98}$$

The same can be shown for the two–point function in the diagram (d). In the $\overline{MS}$ subtraction scheme, we obtain the following renormalisation conditions:

$$\begin{aligned} z &= 1 - \frac{g_2^2}{768\pi^3}\log\frac{M^2}{\mu^2} + \mathcal{O}(g^4) \\ zm^{*2} &= m^2 - \frac{g_2^2M^2}{128\pi^3}\left(\log\frac{M^2}{\mu^2} - 1\right) + \mathcal{O}(g^4), \end{aligned} \tag{1.99}$$

and consequently we have to fine–tune m in order to get a finite value for m^* in the limit $M^2 \to \infty$. Instead, in the zero–momentum subtraction scheme, one gets $R^{ZM}(\Gamma_4) = \mathcal{O}(p^2/M^2)$ and therefore

$$\begin{aligned} z &= 1 + \mathcal{O}(p^2/M^2) \\ zm^{*2} &= m^2 + \mathcal{O}(p^2/M^2), \end{aligned} \tag{1.100}$$

i.e. we see a manifest decoupling of the low–energy effective theory with respect to Φ.

The idea of Ref. [35] is to define a renormalisation scheme which combines the desirable features of the zero–momentum subtraction renormalisation and those of the $\overline{MS}$ schemes. The reason why BHP alone is less convenient is related to the treatment of the infrared divergencies and to the large $\log(m^2/p^2)$ which appear in the calculations when $p^2 \gg m^2$. The $\overline{MS}$ scheme is far more convenient at high–energies. The idea then consists in using the $\overline{MS}$ scheme for high–momentum calculations, where all masses are neglected, while employing a mixed scheme at low–momentum, where $\overline{MS}$ is applied to all graphs containing only light fields, while the BPH is applied to all graphs containing at least one heavy line.

In the intermediate region between masses, the operators evolve according to the renormalisation group equations as the scale μ varies. In the CWZ scheme the renormalisation group coefficients are mass–independent and thus renormalisation–group equations are highly simplified. Moreover it can be demonstrated that this scheme does not introduce extra infrared divergencies and that it preserves gauge invariance.

These features make the CWZ scheme particularly suitable to deal with calculations involving heavy quarks. Indeed, if we consider processes in which all external momentum scales are much smaller than the masses of the heavy quarks, we can omit all graphs containing heavy quark lines and only make a power–suppressed error. Furthermore, given that the CWZ scheme reduces to the $\overline{MS}$ renormalisation scheme at high energy, the same holds in the opposite situation when the masses of heavy quarks are negligible with respect to the high scale of the process. In that case the quark is considered a massless parton and the renormalisation group equations keep the same form as in the $\overline{MS}$ scheme, with the only difference that the total number of flavors is substituted by the number of light flavors.

The application of the CWZ scheme to QCD [37] requires to generalise what we have discussed until now to a theory where more than one heavy field is considered. First of all it is useful to specify the notation: by n_A we refer to the number of active flavors, i. e. of all flavors which are light with respect to the hard scale of the process Q^2. By definition $n_A \geq n_l$. The CWZ renormalisation and factorisation scheme applied to QCD consists then in series of sub–schemes labelled by the number of active flavors. Therefore, in order to define the parton densities and the coupling, the number of active flavors used in the definition must be specified. In each sub–scheme characterised by n_A active flavors

- Graphs that contain only gluons or active quarks are renormalised by $\overline{MS}$ counterterms.

- Graphs with external non–partonic lines (like leptons or vector bosons) are renormalised by $\overline{MS}$ counterterms.
- Graphs whose external lines are gluons or active quarks but which have internal heavy quark lines are renormalised by zero-momentum subtraction.
- Heavy quark masses are defined as pole masses, defined as the position of the pole in the quark propagator in perturbation theory.

These rules apply in the renormalisation of the strong coupling constant, of the parton densities and fragmentation functions. As a consequence, when $n_A = n_f$ the scheme coincides with the $\overline{MS}$ scheme. Moreover it has the advantage that evolution equations for the coupling and parton densities are the same as for QCD with n_A flavors and pure $\overline{MS}$ subtraction.

The term "variable flavor number scheme" (VFNS) refers to the sequence of sub–schemes that we just described. It is typically implemented by using $\overline{MS}$ evolution with a number of active flavors that varies as one crosses the boundaries $\mu \sim M_Q$, where M_Q is the mass of one of the heavy quarks. Thus, for a given scale μ, all quarks with mass less than μ are treated as partons and have associated QCD–evolved parton distributions. The heavy quark parton distribution f_Q vanishes when $\mu^2 < M_Q^2$ [1].

If for simplicity we consider only the first threshold, corresponding to the charm mass, at first sight this approach might be confused with the so–called intrinsic charm approach, where the charm quark appears in the initial state and it is considered as a part of the hadronic wave function. However the two approaches are completely different. In the VFNS charm (or any other heavy quark) is treated in a similar way to a massless parton but the fixed–order perturbation theory (FOPT) calculation performed in a 3 flavors scheme, where the charm appears only as a final state, imposes some conditions on the charm–quark density in the VFN scheme. These conditions, also called matching–conditions, do not exist in the intrinsic charm approach, where the charm density is just an arbitrary function fitted to experimental data [2]. What we have said for the charm density holds for all heavy flavor densities. The $(n_A + 1)$–th distribution is determined by matching the FOPT calculation in n_A flavors with the $(n_A + 1)$ calculation in the massless limit.

These conditions between schemes with different n_A derived from the FOPT calculations are just a case of a transformation between different renormalisation and factorisation schemes. The matching conditions for parton distribution functions can be

[1] Even if commonly the VFN scheme is implemented by choosing the matching point and the switching point between sub–schemes to be equal to the relevant heavy quark mass, this choice is not essential. The difference in theoretical predictions due to the scale is of higher order in α_s with respect to the order of the calculation.

[2] The possibility of having an intrinsic c distribution has been studied for instance in Ref. [38]

found in Ref. [39]. They have the form

$$f_i^{(n_A+1)}(x,\mu^2) = \int_x^1 \frac{dy}{y} \sum_{j=q,\bar{q},g} K_{ij}\left(\frac{x}{y}, \frac{M^2_{n_A+1}}{\mu^2}, \alpha_s^{(n_A)}\right) f_j^{(n_A)}(x,\mu^2). \quad (1.101)$$

The functions K_{ij} can be expressed as an expansion in terms of α_s with coefficients which are polynomial in $\log(M^2_{n_A+1}/\mu^2)$. They are known up to two loops for the parton densities [39]. In Eq. (1.101) i runs over all quarks, light and heavy, while the sum over j runs only on the gluon and the light quarks. This means that the heavy quark parton density in the (n_A+1)–scheme is expressed in terms of the light densities, which physically means that the heavy quark PDFs is generated perturbatively at threshold. If there were also an intrinsic heavy–flavor contribution, the sum on the right–hand side of the equation would extend also to the intrinsic heavy–flavors.

The same relation between sub–schemes are imposed to the strong coupling constant. The β function coefficients in Eq. (5) depend on the number of flavors and, if n_f is interpreted as the number of active flavors at a given scale, the β coefficients are given for the coupling of an effective theory in which n_A flavors are considered light and the heavy flavors decouple from the theory. The coupling for the theory with n_A o flavors and the one for the theory with (n_A+1) flavors are related by an equation of the form

$$\alpha_s^{(n_A+1)}(\mu^2) = \alpha_s^{(n_A)}(\mu^2)\left(1 + \sum_{n=1}^{\infty}\sum_{l=0}^{n} c_{nl}[\alpha_s^{(n_A)}(\mu^2)]^n \log^l\left(\frac{\mu^2}{M^2_{n_A+1}}\right)\right), \quad (1.102)$$

where mass M_{n_A+1} is the mass of the n_A+1–th quark. The matching coefficients are known up to $n=4$, see for instance Ref. [40]. At leading order in α_s it is easy to derive the coefficient c_{10}. Indeed

$$\begin{aligned}
\alpha_s^{(n_A+1)}(\mu^2) &= \alpha_s(M^2_{n_A+1})(1+\alpha_s(M^2_{n_A+1})b_0(n_A+1)\log(M^2_{n_A+1}/\mu^2)) + \mathcal{O}(\alpha_s^2) \\
\alpha_s^{(n_A)}(\mu^2) &= \alpha_s(M^2_{n_A+1})(1+\alpha_s(M^2_{n_A+1})b_0(n_A)\log(M^2_{n_A+1}/\mu^2)) + \mathcal{O}(\alpha_s^2) \\
\Rightarrow \quad & \alpha_s^{(n_A+1)}(\mu^2) = \alpha_s^{(n_A)}(\mu^2) + \frac{\alpha_s^{2,(n_A)}(\mu^2)}{6\pi}\log\left(\frac{M^2_{n_A+1}}{\mu^2}\right) + \mathcal{O}(\alpha_s^3),
\end{aligned} \quad (1.103)$$

where we used the definition of $b_0(n) = (33-2n)/12\pi$. From Eq. (1.103) we observe that $\alpha_s^{(n_A+1)}(M^2_{n_A+1}) = \alpha_s^{(n_A)}(M^2_{n_A+1}) + \mathcal{O}(\alpha_s^3)$ and that the matching between quantities in the subschemes with n_A and (n_A+1) active flavors does not involves large logarithms of masses, provided that the renormalisation/factorisation scale μ is of order the mass of the (n_A+1)–th quark.

1.2.2 Heavy quark mass schemes

Theoretical predictions in QCD have been performed according to a variety of schemes for dealing with the heavy quark masses. A proper treatment of heavy flavors is essential for precision measurements at hadron colliders, especially in global QCD analyses. Recent studies show that the W and Z productions at the LHC are sensitive to detailed features of PDFs which depend on heavy quark mass effects [41, 42, 43]. The theoretical framework described in the previous section is conceptually simple. However its implementation in the calculation of processes involving heavy quarks requires attention to a number of details, both kinematical and dynamical, that may be implemented in different ways leading to different schemes. In this section I identify the different but self–consistent choices that can be made by briefly outlining differences and similarities.

For sake of illustration let us write down the general pQCD factorisation for high–energy hard processes, exemplified in the inclusive DIS structure functions $F_i(x, Q^2)$ as

$$F_i(x,Q^2) = \sum_{a=1}^{n_{f,i}} \sum_{b=1}^{n_{f,f}} \int_\chi^1 \frac{d\xi}{\xi} f_a(\xi\mu) C^a_{b,\lambda} \left(\frac{\chi}{\xi}, \frac{Q}{\mu}, \frac{m_Q}{\mu}, \alpha_s(\mu) \right), \tag{1.104}$$

where a runs over the partons in the initial state, b over the partons in the final states and m_Q the masses of the heavy quarks.

FFNS The conceptually simplest way for taking into account heavy quark mass effects is the fixed–flavor number scheme (FFNS) in which all flavors below m_Q are treated as massless and they are the only ones which enter in the DGLAP evolution equations and in the running of the coupling constant. The heavy quarks enter only in the massive computation of the coefficient functions. For $Q = c, b, t$ the fixed number of light flavors is $n_l = 3, 4, 5$ respectively; the index a in Eq. (1.104) runs from 1 to n_l, while the index b run from 1 to the total number of known flavors, $n_f = 6$.

Historically, higher-order $\mathcal{O}(\alpha_s^2)$ computations of the heavy quark production were all performed in the FFNS [44]. These calculations provide reliable results when the scale of the process is of the order of m_Q compared to the conventional massless calculations. However, at any finite order in a perturbative calculation, the FFNS results become increasingly unreliable as the scale becomes large compared to m_Q: the Wilson coefficients at order n contain logarithmic terms of the form $\alpha_s^n \log^m(Q/m_Q)$, where $m = 1, ..., n$ which might spoil the perturbative expansion. Thus, even if all n_l-flavor FFNS are mathematically equivalent, in practice, the 3-flavor scheme yields the most reliable results in

the region $Q \sim m_c$, the 4-flavor scheme in the region $m_c \leq Q \leq m_b$, the 5-flavor scheme in the region $m_b \leq Q \leq m_t$, and, if needed, the 6-flavor scheme in $Q \geq m_t$.

This naturally leads to the formulation of the variable flavor number schemes (VFNS) which consists of a sequence of n_f-flavor FFNS, each in its region of validity, consistently matched at the transition points. There are several ways of implementing a VFNS.

ZMVFNS The simplest way for implementing the VFNS in parton analyses is the so–called zero–mass variable flavor number scheme (ZMVFNS). In this scheme all quarks are treated as massless. Heavy quarks are absent below threshold and they are radiatively generated at threshold by the subprocess $g \to Q\bar{Q}$ with $m_Q = 0$[3]. The ZMVFNS is therefore a combination of massless $\overline{MS}$ schemes characterised by different numbers of light flavors n_l. Basically the only mass effects are due to the change of number of flavors in the QCD β function and in the anomalous dimensions as one crosses the heavy quark thresholds. The coefficient functions are calculated under the assumption that all active partons are massless, with the associated singularities subtracted in the $\overline{MS}$ scheme

$$C^a_{b,\lambda}\left(\frac{\chi}{\xi}, \frac{Q}{\mu}, \frac{m_Q}{\mu}, \alpha_s(\mu)\right) = C^a_{b,\lambda}\left(\frac{\chi}{\xi}, \frac{Q}{\mu}, 0, \alpha_s(\mu)\right)_{\overline{MS}}.$$

Their expressions are the usual massless coefficient functions which are well–known to NNLO in the QCD coupling α_s for DIS and Drell-Yan processes, and to NLO for many other processes. Hence, due to the simplicity of its implementation, till a few years ago this scheme was the most broadly used in parton analyses.

However this scheme neglects terms above threshold which are proportional to powers of $\frac{m_Q^2}{Q^2}$, thereby losing accuracy for scales close to the thresholds. This does not depend uniquely on the fact that $\mathcal{O}(m_Q^2/Q^2)$ terms are neglected in the coefficient function, but also on the approximate treatment of the phase space. Indeed in Eq. (1.104) both a and b run over all active parton flavors at scale μ. This convention is problematic since the initial state sum runs over the active parton flavors, while the sum over b involves summation over physical final states. This leads to inconsistencies for the conventional ZMVFNS calculation. For example, consider a typical small-x kinematic configuration, say $x = 10^{-4}$ and $Q = 3$ GeV, corresponding to virtual Compton scattering centre-of-mass

[3]This is implemented through a step evolution from the initial scale to the final one passing through the thresholds. Regarding *c*, *b* and *t* as "heavy" implies that they may be computed perturbatively, if no intrinsic heavy flavor combinations are assumed.

(CM) energy $W = \sqrt{Q^2(1-x)/x} = 300$ GeV. Since the factorisation scale is smaller than the bottom quark mass m_b, the bottom quark b in this calculation scheme is not counted as an active parton; thus the final-state parton flavor summation does not include b, whence instead bottom quarks are easily produced at this centre–of–mass energy, as is indeed experimentally observed. Finally another problem in the ZMVFN scheme is related to the scaling variable. In the conventional formulation of the ZMVFN scheme, the extreme of integration in Eq. (1.104) χ is identified with the Bjorken x_B. Since this integral originates from summing over final-state phase space of real particles, this practice leads to violation of Lorentz kinematics in the case of heavy-flavor production. For instance, in neutral-current DIS at $Q \sim m_b$ and $W \sim 2m_b$, this calculation scheme will predict similar contributions from b quark production and d sea quark production (since ZM hard matrix elements are flavor-independent), whereas, in fact, this kinematical regime is below the b-production threshold, and the d scattering cross section is much less suppressed than the b one. It follows that, by extending the lower bound of the convolution integral to x_B in Eq. (1.104), the ZM formalism grossly overestimates the contributions from the region of phase space near the physical thresholds.

I-ZMVFNS In order to overcome the inconsistencies of the ZM formalism mentioned above, while preserving the simplicity of its coefficient functions, in Ref. [45] an intermediate-mass scheme has been proposed. It may be considered either as an improved ZM formulation (I-ZMVFNS) which corrects the kinematic treatment of the final phase space or as a simplified general–mass VFNS. The improvement over the ZM scheme is achieved by replacing the the lower limits of the convolution integral in Eq. (1.104) with the equivalent rescaling variable

$$\chi(x, Q^2) = x_B(1 + M_f^2/Q^2), \tag{1.105}$$

where M_f^2 indicate the total mass of the final states. The choice of this variable restricts the phase space to the physically allowed region. It is shown in Ref. [45] that there are some subtleties related to the definition of χ and a flexible rescaling variable ζ that generalises the mass-dependent rescaling variable of Eq. (1.105) is proposed and discussed. Moreover in the same reference it is shown that global analysis carried out in the I-ZMVFNS can approximate the general–mass scheme results quite well, both in terms of the resulting PDFs, and in terms of predictions of standard observable at the LHC.

If the advantage of the ZM and the I-ZM schemes is the simplicity of their implementation and their consistent resummation of large logarithms of $\mathcal{O}(Q^2/m_Q^2)$, nevertheless they do not fully implement the effect of the heavy quark masses. For this

reason, schemes which interpolate smoothly between the FFNS, which gives a correct description of the threshold region, and ZM-VFNS which accounts for large energy logarithms, have been formulated: they are the so–called general–mass VFN schemes (GMVFNS). Here we briefly describe the available formulations.

ACOT: A technique for the inclusion of mass-suppressed contributions, built upon the CWZ renormalisation scheme [35], was developed long ago and it is the so–called ACOT scheme [46]. It provides a mechanism to incorporate the heavy quark masses in theoretical calculations both kinematically and dynamically. It yields the complete quark mass dependence from the low to high energy regimes; for $m_Q \gg Q$ it ensures manifest decoupling, and in the limit $m_Q \ll Q$ it reduces precisely to the $\overline{\rm MS}$ scheme without any finite renormalisation term. The key ingredient provided by the ACOT procedure is the subtraction term, the third diagram in Fig. 1.4, which removes the "double counting" arising from the regions of phase space where the LO and NLO contributions overlap. Specifically, in the case shown in Fig. 1.4, the subtraction term is given by

$$\sigma_{SUB} = f_g \otimes \tilde{P}_{g\to Q} \otimes \sigma_{QV\to Q}, \tag{1.106}$$

where σ_{SUB} represents a gluon emitted from a proton (f_g) which undergoes a collinear splitting to a heavy quark ($\tilde{P}_{g\to Q}$) convoluted with the LO quark-boson scattering $\sigma_{QV\to Q}$. Here, $\tilde{P}_{g\to Q}(x,\mu) = \frac{\alpha_s}{2\pi}\,\log(\mu^2/m_c^2)\,P_{g\to c}(x)$ where $P_{g\to c}(x)$ is the usual $\overline{MS}$ splitting kernel.

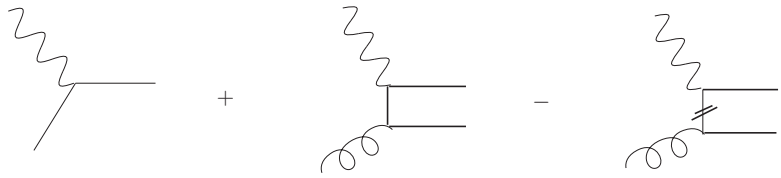


Figure 1.4: Schematic representation of the ACOT scheme implementation at NLO in the heavy quark production process.

Several variants of this method were subsequently proposed, such as **S-ACOT** [47] and **ACOT-χ** [48, 49]. The former was formulated after observing that the heavy quark mass could be set to zero in certain pieces of the hard scattering terms without any loss of accuracy. For instance, in the NLO calculation of the process illustrated in Fig.1.4, one can set $m_Q = 0$ for both the LO terms ($QV \to Q$) and the NLO quark-initiated terms (both the real $QV \to Qg$ and

the virtual $QV \to Q$) as this involves an incoming heavy quark. One can also set $m_Q = 0$ for the subtraction terms as this has an on-shell cut on an internal heavy quark line. Hence, the only contribution which requires calculation with m_Q retained is the NLO $gV \to Q\bar{Q}$ process, drawn in the second diagram in Fig. 1.4. Instead the ACOT-χ scheme is a modified version of the ACOT scheme obtained by rescaling $x_B \to \chi$, where χ^2 is defined in Eq. (1.105) [48]. The factor $(1 + (M_f)^2/Q^2)$ represents a kinematic suppression factor which suppresses the heavy quark production relative to the lighter quarks.

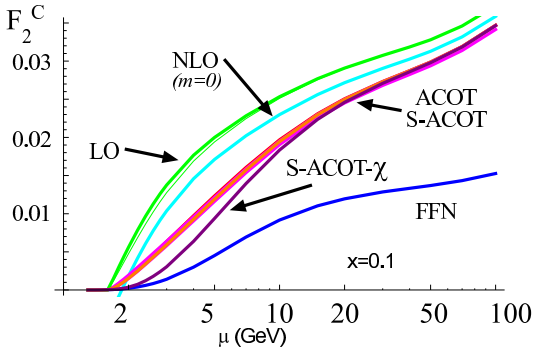


Figure 1.5: Calculation of DIS heavy quark production for a variety of schemes: LO represents $QV \to Q$ process. NLO includes the $\mathcal{O}(\alpha_s^1)$ processes in the massless approximation. In the FFNS the heavy quark PDF is set to zero and it only receives contributions from $gV \to Q\bar{Q}$. ACOT and S-ACOT are described in the text. Taken from Ref. [50].

In Fig. 1.5 the shapes of charm structure function $F_2^c(x, \mu)$, computed in a variety of scheme and perturbative orders, are compared. This comparison has been performed in Ref. [50]: LO represents the $\mathcal{O}(\alpha_s^0)$ $QV \to Q$ process. NLO includes the $\mathcal{O}(\alpha_s^1)$ processes (primarily $gV \to Q\bar{Q}$) in the massless approximation. In the FFNS the heavy quark PDF is set to zero; hence, at $\mathcal{O}(\alpha_s^1)$ this only receives contributions from $gV \to Q\bar{Q}$. The ACOT and S-ACOT schemes are virtually identical—the curves are indistinguishable in this plot. Finally, the implementation of the χ-prescription for the S-ACOT scheme provides some additional suppression in the region $\mu \sim m_Q$.

TR: An alternative method, sometimes called Thorne–Roberts (TR), was introduced in Ref. [51] as an alternative to ACOT [46] with more emphasis on correct threshold behaviour. Like the ACOT scheme it is based on there being two different regions separated by a transition point, by default set about the value of

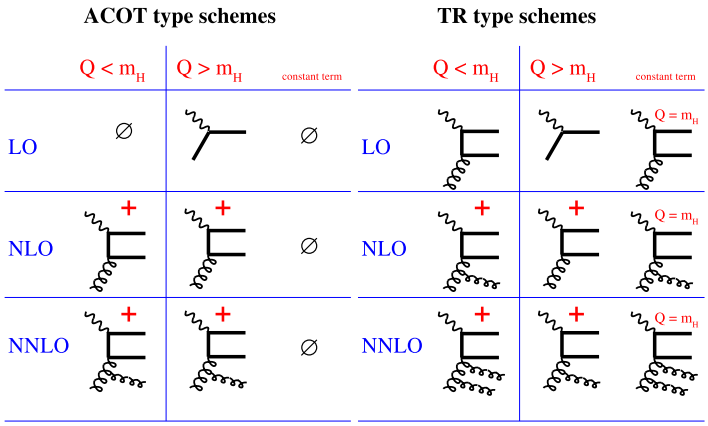


Figure 1.6: Diagrammatic comparison of TR and ACOT type schemes for the case of DIS. This diagram is schematic to emphasise the similarities and differences. Taken from Ref. [50].

the heavy quark mass. In Ref. [50] a diagrammatic comparison between the two approaches is displayed, Fig. 1.6, which helps in illuminating several aspects of the TR scheme. At a fixed order n in perturbation theory, the difference between ACOT and TR amounts to adding different $\mathcal{O}(\alpha_S^{n+1})$ higher order terms, therefore the difference is reduced as one increases the order of our perturbation theory. At LO, α_S^0, when the heavy quark PDF is an "active" parton, the LO contribution is $\gamma + Q \to Q$. However, when the heavy quark PDF is not an "active" parton the LO contribution vanishes. For the ACOT scheme, no higher order terms are added to this results. For the TR scheme, a portion of the $\gamma g \to Q\bar{Q}$ contribution is added; for $\mu < m_Q$ the full $\gamma g \to Q\bar{Q}$ term is included, and for $\mu > m_Q$ the $\gamma g \to Q\bar{Q}$ term frozen at $\mu = Q$ to avoid any difficulty with large logarithms of the form $\ln(m_Q/\mu)$. Consequently, in the $\mu < m_Q$ region, the TR scheme yields a finite LO result while the ACOT scheme yields zero. While both schemes formally agree at $\mathcal{O}(\alpha_S^0)$, the $\mathcal{O}(\alpha_S^1)$ terms can be important, particularly in the $\mu < m_Q$ region. At NLO, for the low μ region we now include $\gamma g \to Q\bar{Q}$ as well as the $\gamma + Q \to Q$ process. At NLO the ACOT scheme obtains a finite result in the region $\mu < m_Q$. For the TR scheme, in addition to the terms present in the ACOT scheme, a portion of the $\gamma g \to gQ\bar{Q}$ contribution is added; again, for $\mu > m_Q$ the $\gamma g \to gQ\bar{Q}$ term is frozen at $\mu = Q$. As before, both the TR scheme and ACOT scheme formally agree at $\mathcal{O}(\alpha_S^1)$, but they will differ by the NNLO $\mathcal{O}(\alpha_S^2)$ terms.

The TR scheme achieves in practice the same highest asymptotic order as ACOT by some modeling of terms below $Q^2 = m_Q^2$ which become (relatively) unim-

portant at high Q^2. These terms are allowed due to the fact that in the PDFs matching conditions there is an arbitrariness related to the definition of the heavy quark coefficient functions. As $m_Q^2/Q^2 \to 0$ all VFNS coefficient functions must tend to the massless $\overline{\rm MS}$-scheme limit, but at finite Q^2 there is a freedom in the heavy quark coefficient functions. In the TR scheme [51] the approach is to make a choice where all coefficient functions obey the correct threshold $W^2 \geq 4m_Q^2$ for heavy quark pair production. This was first imposed by defining the heavy quark coefficient functions such that the evolution of the observable σ, $\partial\sigma/\partial \log Q^2$, is continuous order-by-order at the transition point. However, it results in expressions which become increasingly complicated at higher order. In Ref. [52] the correct threshold behaviour was achieved by using the simple approach of replacing the limit of x for convolution integrals with χ, defined in Eq. (1.105). In the case that the heavy flavour coefficient functions are just the massless ones with this restriction one obtains the S-ACOT(χ) approach. A very similar definition for heavy flavour coefficients was adopted in Ref. [52], resulting in the TR$'$ scheme, and extended explicitly to NNLO.

FONLL: A somewhat different technique for the inclusion of heavy quark effects, the so-called FONLL method, was introduced in Ref. [53] in the context of hadroproduction of heavy quarks. The FONLL method only relies on standard QCD factorisation and it involves calculations with massive quarks in the decoupling scheme of Ref. [35] and with massless quarks in the $\overline{\rm MS}$ scheme. The name FONLL is motivated by the fact that the method was originally used to combine a fixed (second) order calculation with a next-to-leading log one; however, the method is entirely general, and it can be used to combine consistently a fixed order with a resummed calculations to any order of either. The application of the FONLL scheme to deep–inelastic structure functions was recently presented in Ref. [54]. The method is based upon the idea of looking at both the massless and massive scheme calculations as power expansions in the strong coupling constant, and replacing the coefficient of the expansion in the former with their exact massive counterpart in the latter, when available. In Ref. [54] three FONLL scheme implementations have been proposed: scheme A, where one uses the NLO massless scheme calculation, matched with the LO (i.e. $\mathcal{O}(\alpha_s)$) massive scheme calculation; scheme B, where one uses the NLO massless scheme calculation, matched with the NLO (i.e. $\mathcal{O}(\alpha_s^2)$) massive scheme calculation; and scheme C, where one uses the NNLO massless scheme calculation, matched with the NLO massive scheme calculation. Moreover, in order to suppress higher order contribution arising in the subtraction term near the threshold region, two prescriptions are proposed. One consists in damping the subtraction term by a threshold factor which differs from unity by power–suppressed terms; the other consists in using a rescaling variable such as the

one defined in Eq. (1.105). Both prescription introduce terms which are formally subleading with respect to the order of the calculation, therefore they do not change its nominal accuracy, but they may in practice improve the perturbative stability and smoothness of the results.

Thanks to its simplicity, the FONLL method provides a framework for understanding differences between other existing approaches, and for a study of the effect of different choices in the inclusion of subleading terms. It is easily seen that the scheme A in the FONLL calculation should be equivalent to the S-ACOT scheme. If a χ-scaling prescription is applied to all terms computed in the massless approximation, scheme A should become equivalent to the S-ACOT-χ prescription. A benchmark comparison presented at the Les Houches workshop [55] confirms this statement in a quantitative way. By adopting common settings the heavy quark structure functions F_{2c} and F_{Lc}, as implemented in the ACOT, TR and FONLL schemes, have been computed and compared. In Fig. 1.7 the S-ACOT, ACOT and FONLL-A computations for the F_2^c structure function are compared. It is clear that, without applying any threshold or rescaling prescription, the S-ACOT and the FONLL-A schemes are exactly equivalent. On the other hand the difference between the full ACOT and the S-ACOT scheme vanishes for $Q^2 \sim 10$ GeV^2, as it should be, given that they differ only by mass–suppressed terms. Scheme C should again be equivalent to a NNLO

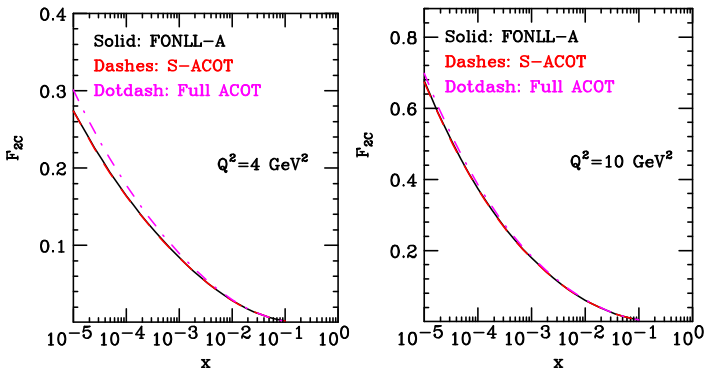


Figure 1.7: The F_{2c} structure function for $Q^2 = 4$ (left) and 10 (right) GeV^2 in the FONLL scheme A (plain) compared to the Simplified ACOT (S-ACOT) and full ACOT schemes. Taken from Ref. [54].

generalisation of the S-ACOT scheme. Scheme B instead does not correspond to any S-ACOT calculation. It is more reminiscent of the TR method, where at the

NLO level the full NLO massive result is also used[4]. However from Figs. 1.8 no obvious similarities can be identified between the two families of schemes, neither at NLO nor at NNLO at small Q^2. Clearly the difference between schemes decrease when Q^2 is increased. Furthermore it was observed that, switching off the threshold prescriptions in both cases, FONLL-A gets rather close to TR' at NLO and FONLL-C gets rather close to TR' at NNLO.

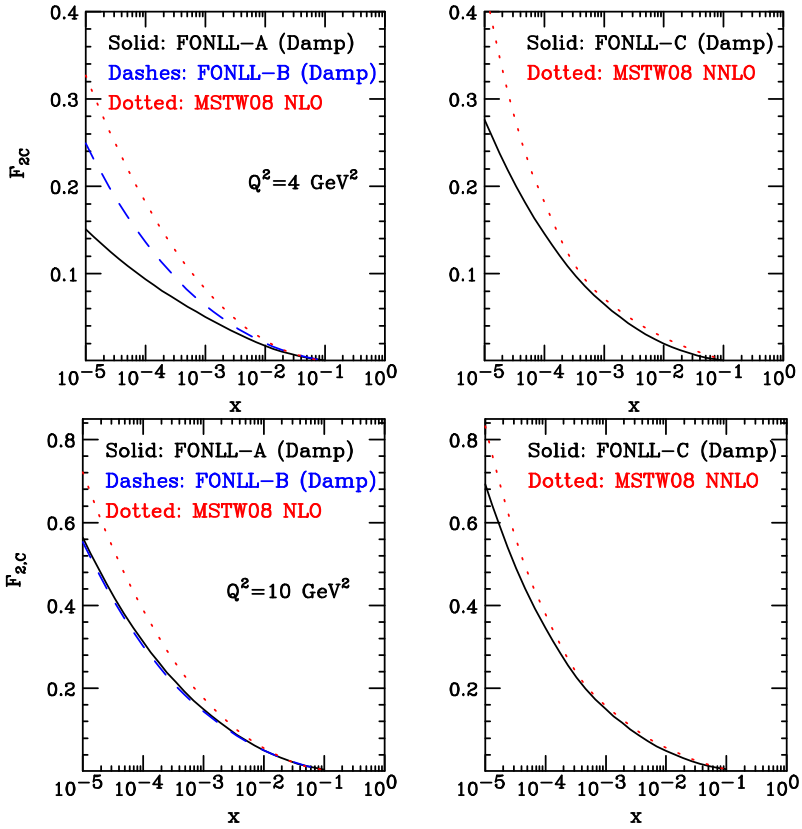


Figure 1.8: The F_{2c} structure function for $Q^2 = 4$ (top) and 10 (bottom) GeV^2 in FONLL and in TR$'$, both for the NLO schemes (left plots) and for the NNLO schemes (right plots). In both cases the default threshold prescriptions are used: χ-scaling for TR$'$ and a damping factor for FONLL. Taken from Ref. [54].

[4]This is only true at $Q^2 = m_c^2$, since in the TR method the higher order term in the massive calculation is frozen at threshold.

Chapter 2

Parton Distribution Functions

In this chapter several aspects of the determination of parton distribution functions are explored. Starting from an historical overview of the method for extracting PDFs from experimental data, I emphasise the progress that has been made, not only due to the increased theoretical and experimental accuracy but also due to the refinement of the statistical tools used. I describe both experimental and theoretical inputs of the so–called global analyses and the statistical issues associated to the determination of a set of functions from a finite number of data points. Then some of the delicate aspects related to the determination of the best–fit in a space of functions are discussed, especially the inclusion of the normalisation uncertainties. I show that an improper treatment of normalisation uncertainties might lead to a bias and propose a prescription able to eliminate the latter. The final section is devoted to the description of the several benchmark studies which have been recently performed in order to clarify the differences between various approaches used to fit PDFs in a simplified context where common settings are used.

2.1 Global QCD analyses

Parton Distribution Functions are universal, process–independent, non–perturbative quantities which must be extracted at a given scale from comparison to the available experimental data. They are then evolved by mean of the DGLAP equations to another scale where they are used as an input for theoretical predictions. Any calculation of cross sections with hadrons in the initial states involves the choice of a set of PDFs,

as well as the experimental simulations performed by mean of Monte Carlo event generators.

The needs of precision physics at hadron colliders have determined a revolution in the approach to the determination of PDFs over the last few years. While the Tevatron has been providing data for a variety of hard hadronic processes which establish the validity of collinear factorisation and parton universality to the level of precision physics, the LHC requires a precision approach to the structure of the nucleon in the context of searches for new physics. Given that the uncertainty of PDFs represents the dominant systematic uncertainty in the theoretical prediction of some key processes at the LHC, such as the production of electroweak bosons, a reliable knowledge of their errors is essential for the full exploitation of the LHC physics potential.

At a very early stage [56, 57, 58], parton distributions were determined through a combination of general physical principles (as embodied in sum rules), model assumptions and the first crude experimental information coming from Bjorken scaling and its violation. These determinations aimed to show the compatibility of the data with the partonic interpretation of hard processes. Thanks to a second generation of high–precision DIS and hadron collider experiments, QCD gradually evolved towards being viewed as precision physics. This required an approach to parton densities determination based on the next-to-leading order theory (in order to have perturbative uncertainties under control), and also based on fairly wide "global" sets of data of a varied nature, in order to minimise as much as possible the role of theoretical prejudice in the determination of the shape of the parton distributions. Next-to-leading order parton sets became standard analysis tools and were constantly updated. In particular, the wealth of data from the HERA collider [17] led to a considerable extension of the kinematic region over which parton distributions could be determined, along with a substantial improvement in accuracy.

With parton distributions becoming a tool for precision physics, it became important to be able to assess accurately the uncertainty on any given parton set. The procedure for estimating the PDFs uncertainties was progressively refined and more data were included into the analyses. There are currently at least four global sets of parton distributions with uncertainties available and constantly updated by the CTEQ–TEA [59, 60, 41, 61, 62], MRST–MSTW [63, 64, 65, 66, 43], NNPDF [67, 68, 69, 70, 71] and Alekhin–ABKM [72, 73, 74, 75] groups. There are many similarities in the approach of the first two collaborations, both in the choice of the parametrisation and in the determination of PDFs uncertainties. The ABKM collaboration uses the same kind of fixed functional parametrisation for parton densities but a somewhat different procedure in the determination of the error. The NNPDF collaboration adopts a radically different approach, which is going to be described in Chap. 4. On the other hand,

among the non–global analyses, there are the PDF sets determined by the H1-ZEUS collaboration from the self–consistent and accurate HERA data [17].

2.1.1 Experimental input

The essential input for a global QCD analysis comes from experimental data. In principle, one would like to include as many experimental datasets as possible in order to maximally constrain PDFs. However the complexity of each dataset makes the application of standard statistical tools difficult and the differences between datasets pose a challenge for global analyses. First of all, when many datasets of the same or similar processes are included, they may not be consistent according to standard statistical tests, even if the individual ones appear to be self-consistent. Besides some experiments are performed using different nuclear targets. Although the assumption of incoherent scattering off individual partons inside the nucleus is valid as a first approximation, nuclear effects might significantly affect the analysis and theoretical calculations of nuclear correction factors are model-dependent, hence often controversial. On the other hand, there are issues related to the kinematic cuts in variables Q^2 and $W^2 = Q^2(1-x)/x$. Although in principle one would like to have a range as wide as possible, there are some restrictions that must be taken into account: theory and experiments should be compared only where the theory is expected to be reliable and this implies that some kinematic region must be excluded, typically the data whose $Q^2 < 2$ GeV^2 and $W^2 \lesssim 10$ GeV^2.

To give an idea of the wide range of potential experiments available for global QCD analysis, in Tab. 2.1 we list the typical input data, the measured physical observables and the PDFs constrained by each of them. The datasets listed in Tab. 2.1 cover a wide kinematical range and each of them has different features. The HERA experiments are high statistics, high precision experiments. They consist of many hundreds of data points with statistical and systematic errors of only a few percents. The combined H1–ZEUS data of neutral and charged current reduced cross sections improve on the accuracy of the separate H1 and ZEUS separate datasets, due to the cross–calibration between detectors. HERA data are mostly at low–x, while fixed target data, like NMC, BCDMS and SLAC, cover a higher–x region. Even if the latter have considerably lower statistics and often large systematic errors, their inclusion enables one to disentangle isospin triplet and isospin singlet contributions due to the combination of deuteron and proton data. Moreover charged current scattering data from charged lepton beams and neutrino scattering data disentangle the quark and antiquark distributions. On top of these data, H1 and ZEUS have also provided some measurements of $F_{2,c}$ and $F_{2,b}$ which help in constraining the charm and bottom distributions.

Experiment	Observable	Constrain	Ref.
DIS experiments			
SLAC	$F_2^{e^-p}, F_2^{e^-d}$	medium–,large–$x\ q, \bar{q}, g$	[76]
BCDMS	$F_2^{\mu p}, F_2^{\mu d}$	medium–,large–$x\ q, \bar{q}, g$	[77, 78]
NMC	$F_2^{\mu p}, F_2^{\mu d}$	medium–,large–$x\ q, \bar{q}, g$	[79]
	$F_2^{\mu p}/F_2^{\mu d}$	large–$x\ d_v/u_v$ ratio	[80]
E665	$F_2^{\mu p}, F_2^{\mu d}$	medium–,large–$x\ q, \bar{q}, g$	[81]
H1,ZEUS	$F_2^{e^\pm p}, \tilde{\sigma}^{NC,e^\pm}, \tilde{\sigma}^{CC,e^\pm}$	small–$x\ q\bar{q}, d_v$ and small–,medium– $x\ g$	[17]
	$F_{2,c}^{e^\pm,p}, F_{2,b}^{e^\pm,p}$	c, b	[82]
	F_L	small–$x\ g$	[83]
HERAI-AV	$\tilde{\sigma}^{NC,e^\pm}, \tilde{\sigma}^{CC,e^\pm}$	all $x\ q\bar{q}$ small–$x\ g$	[84]
ZEUS-H2	$\tilde{\sigma}^{NC,e^-}, \tilde{\sigma}^{CC,e^-}$	medium–,large– $x\ q\bar{q}$	[85]
CCFR	$F_{2,3}^{\nu(\bar{\nu})}$	medium–$x\ u_v + d_v$	[86]
	$\tilde{\sigma}^{\nu(\bar{\nu}),c}$	strange sea, $s, \bar{s}$	[87]
CHORUS	$\tilde{\sigma}^{\nu(\bar{\nu})}$	medium–,large–$x\ q, \bar{q}, g$	[88]
NuTeV	$F_{2,3}^{\nu(\bar{\nu})}$	medium–,large–$x\ q, \bar{q}, g$	[89]
	$\tilde{\sigma}^{\nu(\bar{\nu}),c}$	strange sea, $s, \bar{s}$	[89]
Fixed Target Drell-Yan production			
E605	$\frac{d^2\sigma^{\rm Cu}}{dM^2dy}$	$\bar{q}, g$	[90]
E772	$\frac{d^2\sigma^{\rm Cu}}{dM^2dy}$	$\bar{q}, g$	[91]
E866	$\frac{d^2\sigma^{p,d}}{dM^2dx_F}$	$\bar{q}, g$	[92, 93]
	$\frac{d^2\sigma^{d}}{dM^2dx_F}/\frac{d^2\sigma^{p}}{dM^2dx_F}$	$\bar{u}, \bar{d}$	[94]
Collider vector boson production			
CDF	$\frac{\sigma^{W^+}-\sigma^{W^-}}{\sigma^{W^+}+\sigma^{W^-}}$	large–$x\ u/d$ ratio	[95]
	$\frac{d\sigma^Z}{dy}$	u, d	[96]
D0	$\frac{d\sigma^Z}{dy}$	u, d	[97]
Collider inclusive jet production			
H1, ZEUS	DIS+jet data	medium–$x\ g$	[98]
CDF	$\sigma^{\rm jet,inc}$	large–$x\ g$	[99]
D0	$\sigma^{\rm jet,inc}$	large–$x\ g$	[100]

Table 2.1: Table of the experimental datasets and of the measured observables included in a typical global QCD analyses of parton distribution functions.

On the other hand, the dimuon data from the CCFR and NuTeV collaborations help in constraining the strange sea, as it will be discussed in Chap. 6.

If DIS data provide important constraints to the quarks and anti–quark distributions, as well as to the gluon at medium and small–x, there are other regions of the PDF space which need to be constrained by the hadronic data. Drell–Yan data, like the fixed–target E605, E772 and E866 experiments help in disentangling the anti–quark contributions. On the other hand, the collider vector boson production data helps in constraining the u/d ratio at high–x and the u and d distributions. Finally collider jet data cover a broad range in x and Q^2 by themselves and are particularly important in the determination of the high–x gluon distribution.

2.1.2 Theoretical input

There are several sources of uncertainty in a parton analysis, which may be divided in two classes: those associated to the experimental errors and those due to the so–called theory errors, which depend on the theoretical inputs of the analysis. It is important to remark that the error bands associated to the PDFs do not include the theoretical error[1]. The latter may induce systematic shifts in the central values and the error bands.

The main theoretical inputs to global analyses are the perturbatively calculated hard cross sections, along with the QCD evolution equations which control the scale dependence of the PDFs. The $\overline{MS}$ scheme is used almost universally but some PDFs are also available in the DIS scheme. One of the main theoretical errors comes from the higher–order contributions. Most global PDF analyses are carried out at NLO. Recently, the DGLAP evolution kernels have been calculated at NNLO [30, 31], allowing a full NNLO evolution to be carried out, and parton sets at NNLO are available [43, 74]. However, not all processes in the global fits, and specifically the inclusive jet production, are available at NNLO. Thus, current NNLO global PDF analyses are still approximate.

Other sources of possible large corrections to the standard QCD parton model formula may come from the lack of resummation of logarithms arising at the boundaries of the phase space, from effects beyond the standard DGLAP expansion, from power–law corrections like higher–twists, target–mass corrections and, if applicable, nuclear corrections.

Another significant source of theoretical error is the treatment of the heavy quark masses. The various schemes for including their effect in a global analysis have been

[1]The inclusion of the theoretical uncertainty in the standard PDF uncertainty bands has been broadly discussed but there is no agreement about its precise definition. For a broader discussion see Refs. [64, 101]

described in the previous chapter. Here I summarise the heavy quark schemes used by the parton fitting collaborations in their most recent analyses.

- The ABKM collaboration has recently released three sets of NLO and NNLO PDFs evaluated in a FFNS with $n_f = 3, 4, 5$ [102]. The set with $n_f = 4$ is obtained by using the set with $n_f = 3$ as input and using the Buza–Matiounine–Smith–van Neerven matching conditions of Ref. [39] at $Q^2 = m_c^2$. Analogously the $n_f = 5$ set is obtained from the 4–flavors PDFs.

- While the ACOT scheme [46] was formulated long time ago, it was first used for an actual general-purpose global parton fit only recently, in Refs. [49, 41]. All previous CTEQ analyses were performed in the ZMVFNS, even if the full ACOT scheme had been used in specific studies in the CTEQ HQ series of fits, HQ4 [103], HQ5 [104] and HQ6 [49].

- The Thorne Roberts (TR) scheme [51] was used in MRST global analyses up to MRST 2004 [65]. The TR′ scheme [52] was first used in the MRST 2006 analysis [66], and has been used in all subsequent MSTW analyses [43].

- The FONLL GM-VFN scheme [53, 54] is currently being implemented in the NNPDF family of fits [105], which up to now have been obtained in the ZMVFNS [68, 70, 71].

On top of the mentioned theoretical errors there are theoretical assumptions, implicit in the choice of the parametrisation and in the flavour decomposition. In principle one should parametrise and extract all the $(2n_f + 1)$ independent parton distributions by fitting observables from experimental data involving different linear combinations of PDFs. In practice, however, one has to rely on some assumptions since the available data cannot constrain all of them. For example, an approximation used in early fit consisted in considering the sea $\bar{u} = \bar{d}$, due to the lack of experimental information. For the same reason, until a few years ago, the strange valence distribution $(s-\overline{s})$ was typically set to zero. Nowadays the availability of Drell–Yan and as well as charged–current and the dimuon production data, allows to disentangle the sea quarks contribution and to determine independent $\bar{d}, \bar{u}, s, \bar{s}$ distributions.

The choice of the parametrisation is another crucial point [106]. In principle, since PDFs represent our ignorance of the non-perturbative nucleon structure, there should be complete freedom in choosing their parametric form. However, in order to carry out a parton fit, one needs to choose a particular functional form for the PDFs at the initial scale. The typical approach adopted by most of the PDFs fitting collaborations consists in setting a functional parametrisation which takes into account some handle

we have on PDFs behaviour; the standard parametrisation has a form such as

$$f_i(x, Q_0^2) = a_0\, x^{a_1}\, (1-x)^{a_2} P(x, a_3, a_4, ...),$$

where for some PDF combinations the normalisation parameters a_0 are typically determined by imposing the momentum and valence sum rules. The factor a_1 is motivated by physics considerations in the small x limit where the power a_1 has a Regge interpretation. The factor a_2 is motivated by physical considerations in the limit $x \to 1$ which is the resonance region where PDFs are supposed to be zero; its value can be related to the quark counting rules [107]. However it is unclear to what value of Q^2 these considerations should apply and therefore such predictions can only be taken as a rough guide for the values to be expected. Finally the function $P(x)$ is a smooth polynomial function in x which interpolates between the small-x and the large-x regions.

The functional form has to be flexible enough to accommodate observed experimental data without introducing a theoretical bias. Given that the flexibility is intrinsically related to the number of parameters, the latter should be large enough. However, adding more parameters to the parametrisation might complicate the fitting procedure due to the lack of experimental constraints.

Another issue is related to the question: how do we estimate the error associated to the choice of a particular functional form? Are the results independent on its choice? This kind of question cannot be easily addressed within the framework of the traditional parton analysis. In Chap. 4 we will see that in the NNPDF approach, where a more general parametrisation than the simple polynomial one is used, this question may be easily addressed. Another attempt in this direction has been performed in the recent HERAPDF analysis [108], where in addition to the model uncertainty a new error contribution is introduced resulting from the parametrisation choice. Alternative parametrisations leading to good fit quality but peculiar behaviour at large–x are used to estimate the parametrisation uncertainties. An envelope of these fit solutions is built, which is added in quadrature to the contributions of the experimental and model uncertainties, as it is shown in Fig. 2.1.

2.1.3 Fitting procedure

When performing a global fit one needs a criterion for evaluating the "goodness of fit" for a particular set of PDFs. All experimental uncertainties are generally assumed to be Gaussian. Then the least squares method estimates the quality of the fit by mean of

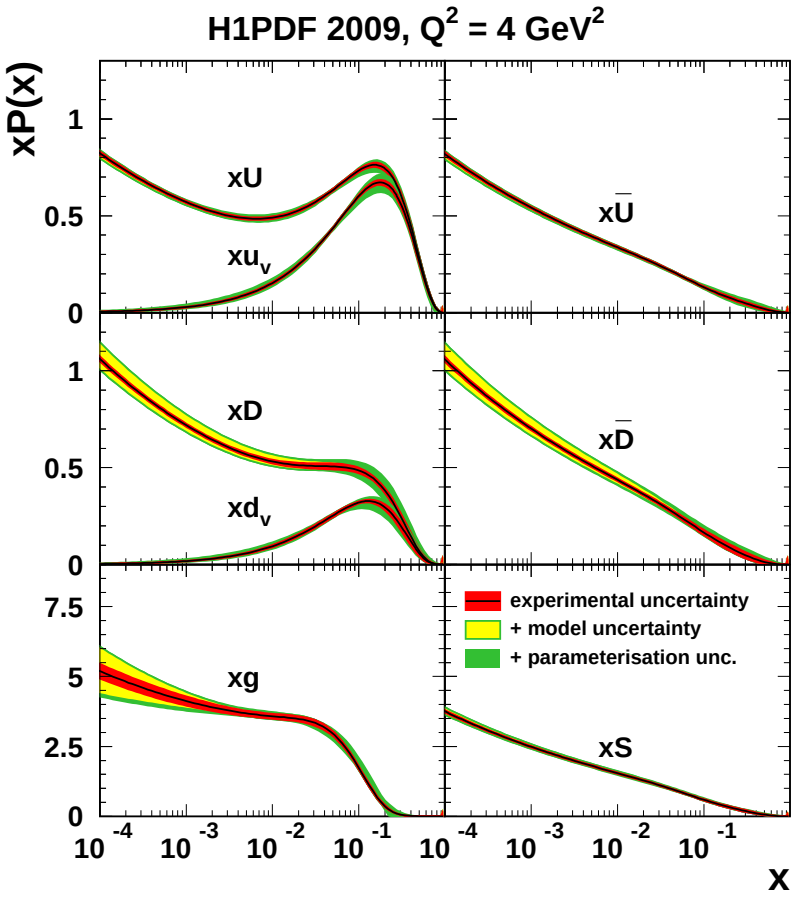


Figure 2.1: The parton distribution functions from HERAPDF1.0, xU, xu_v, $x\bar{U}$, xD, xd_v, $xS = 2x(U^- + D^-)$ and xg, at Q^2 = 10 GeV^2. The experimental, model and parametrisation uncertainties are shown separately. Taken from Ref [108].

the χ^2 function [109]. If only statistical errors are included, the latter is defined as

$$\chi^2\left(\{f\}\right) = \sum_{i=1}^{N_{dat}} \frac{\left(F_i^{(\exp)} - F_i^{(\text{theo})}\left(\{f\}\right)\right)^2}{\sigma_{i,\text{stat}}^2} \,. \tag{2.1}$$

where $F_i^{(\exp)}$ is the experimental measurement of some observable F, $F_i^{(\text{theo})}\left(\{f\}\right)$ is the theoretical prediction as a function of the initial scale PDFs $\{f\}$, hence function of the parameters $\vec{a}$ which describe the set of parton distributions, and $\sigma_{i,\text{stat}}$ is the total uncorrelated statistical uncertainty. The parameters $\vec{a}$ are determined by minimising the χ^2, *i.e.* one wants to determine the set of parameters $\vec{a}^0$ which satisfy the condition

$$\chi^2\left(\vec{a}^0\right) \equiv \min_{\vec{a}}\left[\chi^2\left(\vec{a}\right)\right] \,. \tag{2.2}$$

The sum in quadrature of statistical and systematic uncertainties is an acceptable approximation only if statistical errors are dominant and correlations negligible. However the statistical accuracy of current data is much higher than it used to be, and the experimental groups provide the contributions from the different sources of systematic errors along with the experimental data; this leads to consider the effects of the correlated systematic uncertainties. This can be done by using the explicit form of the experimental covariance matrix, which collects all information about statistical, systematic and normalisation uncertainties. The latter is defined as

$$[\text{cov}]_{ij} = \left(\sum_{l=1}^{N_c} \sigma_{i,l}\sigma_{j,l} + \sum_{n=1}^{N_a} \sigma_{i,n}\sigma_{j,n} + \sum_{n=1}^{N_r} \sigma_{i,n}\sigma_{j,n} + \delta_{ij}\sigma_{i,s}^2\right) F_i^{(\exp)} F_j^{(\exp)} \,, \tag{2.3}$$

where i and j run over the experimental points, $F_i^{(\exp)}$ and $F_j^{(\exp)}$ are the measured central values, and the various uncertainties, given as relative values, are: $\sigma_{i,l}$, the N_c correlated systematic uncertainties; $\sigma_{i,n}$, the N_a (N_r) absolute (relative) normalisation uncertainties; $\sigma_{i,s}$ the statistical uncertainty.

With this definition, the χ^2 which is minimised during the fitting procedure is given by

$$\chi^2\left(\vec{a}\right) = \sum_{i,j=1}^{N_{dat}} \left(F_i^{(\exp)} - F_i^{(\text{theo})}\left(\vec{a}\right)\right) \left[\text{cov}^{-1}\right]_{ij} \left(F_j^{(\exp)} - F_j^{(\text{theo})}\left(\vec{a}\right)\right) . \tag{2.4}$$

A special treatment is required for the normalisation and multiplicative uncertainties which, as we are going to see in Sect. 3.2, should be treated separately from other systematic errors in order to avoid systematic biases.

2.1.4 Determination of errors

Until a few years ago, only the central values were provided and errors on PDFs were estimated by varying some sets of parameters, or comparing different determinations. They were generally considered to be negligible in comparison to other sources of theoretical or experimental error. Now, the simultaneous progress in higher-order theoretical calculations and experimental results requires a realistic estimate of the error of the released PDF sets. Indeed, all modern parton fitting collaborations provide a set of PDFs at the initial scale and their errors.

The determination of the PDFs error is a difficult problem, not only because of the difficulties in collecting and propagating uncertainties contained in large experimental covariance matrices, but also because a PDF set is a set of functions that should be inferred from a finite amount of data points. Therefore one is faced with the problem of constructing a probability measure in a space of functions. The mean of an observable $\mathcal{O}$ depending on a set of PDFs $\{f\}$ would then be given by

$$\langle \mathcal{O}[\{f\}] \rangle = \int [\mathcal{D}f]\, \mathcal{O}[\{f\}] \mathcal{P}[\{f\}], \tag{2.5}$$

where $\mathcal{O}[\{f\}]$ is the value of $\mathcal{O}$ based on the PDF set $\{f\}$ and $\mathcal{P}[\{f\}][\mathcal{D}f]$ is the probability density measure in the space of PDFs.

A technique for estimating the probability density $\mathcal{P}[\{f\}][\mathcal{D}f]$ and to deduce the errors and the correlation of quantities depending on PDFs is provided by the use of Monte Carlo (MC) methods. Using random sampling, *i.e.* a MC evaluation of the integral which appears in Eq. (2.5), these methods succeed in calculating such integral without relying on Gaussian approximations or linearisations of uncertainties. The underlying idea is that, if a sufficiently high number of PDFs replicas is generated according to $\mathcal{P}[\{f\}][\mathcal{D}f]$, the integral of Eq. (2.5) can be approximated by the sum

$$\langle \mathcal{O}[\{f\}] \rangle \simeq \frac{1}{N_{\rm rep}} \sum_{i=1}^{N_{\rm rep}} \mathcal{O}[f_i], \tag{2.6}$$

where i runs over the members of the Monte Carlo ensemble. In the original work where this technique was presented, Ref. [110], the probability density $\mathcal{P}[\{f\}][\mathcal{D}f]$, was projected from the functional space of all possible functional forms assumed by the parton densities into the $N_{\rm par}$–dimensional space of parameters characterising the chosen functional form; the initial probability density is then a function $\mathcal{P}(a)d\vec{a}$, with $\vec{a} \equiv a_1, a_2, ..., a_{N_{\rm par}}$. The space of parameters is sampled by generating a Monte Carlo ensemble consisting of $N_{\rm rep}$ random sets of parameters $\{\vec{a}\}^{(k)}$, $k = 1, 2, ..., N_{\rm rep}$ distributed according to $\mathcal{P}[\vec{a}]$. However, while this can be numeri-

cally implemented, the limitations of the method proposed in Ref. [110] and developed in Ref. [111] are related to the generation of the Monte Carlo ensemble. Indeed, the sampling of the parameter space $\{\vec{a}\}^{(k)}$ is not particularly efficient due to the presence of many flat directions, which would require the generation of very large ensembles which cannot be used for practical purposes. In the NNPDF approach [67, 68, 70, 71] this problem is solved by generating the Monte Carlo ensemble in the space of the experimental data included into the fit rather than in the space of parameters, as discussed in detail in Chap. 4.

The most widely used approach in the PDF fitting community is the Hessian method. A detailed description can be found in Refs. [112, 63]. Given a set of $N_{\rm par}$ parameters $\vec{a}$ determining the PDFs at the initial scale, one first determines the best fit parameters $\vec{a}^0$ from the minimisation of the fully correlated χ^2. Then, in order to estimate the associated uncertainty, one assumes that the deviation in χ^2 from the minimum value χ^2_0 is quadratic in the deviation of the parameters specifying the input parton distributions $\vec{a}$ from their values at the minimum $\vec{a}^0$ and assumes that

$$\Delta\chi^2 \equiv \chi^2 - \chi^2_0 \sim \sum_{i=1}^{N_{\rm par}} \sum_{j=1}^{N_{\rm par}} H_{ij} \left(a_i - a_i^0\right) \left(a_i - a_j^0\right) \ , \tag{2.7}$$

where H is the Hessian matrix, defined as

$$H_{ij} = \frac{\partial^2 \chi^2 \left(\vec{a}\right)}{\partial a_i \partial a_j} \ . \tag{2.8}$$

H_{ij} has a complete set of $N_{\rm par}$ orthonormal eigenvectors v_{jk} with eigenvalues ε_k

$$\sum_{j=1}^{N_{\rm par}} H_{ij} \left(\vec{a}\right) v_{jk} = \varepsilon_k v_{ik} \qquad i,k = 1, \ldots, N_{\rm par} \ . \tag{2.9}$$

The eigenvectors provide a natural basis to express arbitrary variations about the minimum. One has to take into account that since variations in some directions in the parameter space lead to deterioration of the quality of the fit far more quickly than others, the eigenvalues ε_k span several orders of magnitude and therefore the diagonalisation might be hard, especially when the fully correlated χ^2 involves large covariance matrices. In terms of the diagonalised set of parameters defined with respect to the eigenvectors v_{ij}

$$z_i = \sqrt{\frac{\varepsilon_i}{2}} \sum_j (a_j - a_j^0) v_{ij} \tag{2.10}$$

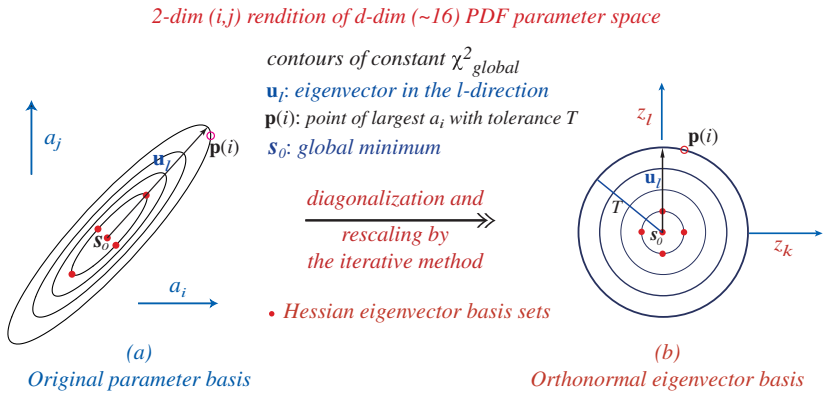


Figure 2.2: A schematic representation of the transformation from the PDFs parameter basis to the orthonormal eigenvector basis defined in Eq. (2.9). Taken from Ref. [113]

one obtains $\Delta\chi^2 = \sum z_i^2$, i.e. the surfaces of constant χ^2 are spheres in the $\vec{z}$ space of radius $\sqrt{\Delta\chi^2}$. The diagonalisation procedure is illustrated in Fig. 2.2. If one then considers an observable $\mathcal{O}$ depending on PDFs $\{f\}$, hence on the parameters $\vec{a}$, in the neighbourhood of the global minimum, assuming the first term of the Taylor–series expansion of $\mathcal{O}$ gives an adequate approximation, the deviation of $\mathcal{O}$ from its best estimate is given by $\Delta\mathcal{O} = \mathcal{O} - \mathcal{O}_0 \approx \sum_i O_i z_i$ with $O_i \equiv \partial\mathcal{O}/\partial z_i|_{\vec{0}}$. For a given tolerance $\Delta\chi^2$ the uncertainty on the physical observable can be evaluated by the simple formula

$$(\Delta\mathcal{O})^2 = \Delta\chi^2 \sum_i O_i^2 = \sum_i (\mathcal{O}(S_i^+) - \mathcal{O}(S_i^-))^2 \tag{2.11}$$

where in the above equality $\mathcal{O}_i$ is evaluated by finite difference and $S_i^\pm$ are PDF sets which correspond to two points in the $\vec{z}$ parameter space specified by

$$z_j^\pm = \pm\delta_{ij}\sqrt{\Delta\chi^2}/2,$$

as one can see in Fig. 2.2. Thus, with the calculation of the Hessian matrix and its eigenvectors, one obtains $(2N_{\rm par}+1)$ sets of parton distributions, $S_0, S_1^\pm, \cdots, S_{N_{\rm par}}^\pm$ from which one can evaluate the uncertainties of PDFs themselves and of the observables depending on them by using the master formula, Eq. (2.11).

Possible drawbacks of this method are the assumption that the linearised approximation in error propagation is valid, and the introduction of non-standard tolerance criteria $\Delta\chi^2 > 1$ which does not allow to give a statistically rigorous meaning to the

resulting uncertainties. Indeed results depend on the choice of the tolerance $\sqrt{\Delta\chi^2}$, which is the allowed variation in χ^2. Textbook statistics implies that, if the measurements belong to experimental sets which are compatible with each others, one should have $\Delta\chi^2 = 1$. However it has been argued that a higher value is required in order to keep into account the inconsistencies between datasets included into a global analysis. The CTEQ collaboration has investigated this problem in detail [114]. They consider that the choice $\Delta\chi^2 = 1$ is not a reasonable tolerance on a global fit with about 2000 data points from diverse sources, with theoretical and model uncertainties which are hard to quantify and experimental uncertainties which may not be Gaussianly distributed. They have tried to formulate a more reasonable setting for the tolerance T by estimating the range of overall χ^2 along each of the eigenvector directions within which a good fit to all datasets (within 90% C.L.) can be obtained, and then averaging the ranges over the $N_{\rm par}$ eigenvector directions. According to this study $T \sim \sqrt{100}$. The MRST collaboration suggested a slightly smaller value for $T \sim \sqrt{50}$ confirmed by a similar analysis performed by the ZEUS collaboration [115]. More recently, the MSTW collaboration has substituted the choice $T = \sqrt{50}$ by a new procedure which enables a *dynamic* determination of the tolerance for each eigenvector direction, by demanding that each dataset must be described within its one-sigma (or 90% C.L.) limits according to a hypothesis-testing criterion, after rescaling the χ^2 for each dataset so that the value at the global minimum corresponds to the most probable value. Application of this procedure to the MSTW benchmark fit gives $T \sim 3$ for 1σ uncertainties and $T \sim 5$ for 90% C.L. uncertainties. Nevertheless the use of $T > 1$ is still controversial given that there is no rigorous statistical proof for the criterions adopted to estimate it. The introduction of a tolerance larger than one is equivalent to the PDG error rescaling procedure for dealing with incompatible datasets and practically corresponds to inflate the experimental and PDFs uncertainties by a factor $\sqrt{\Delta\chi^2/2.7}$ [109].

Another technique for estimating PDFs errors which does not rely on the quadratic approximation of the χ^2 is the Lagrange multiplier method. The technique has been applied in literature to determine the PDFs induced uncertainties on several physical observables [63, 116]. However it is not practical for global fits of parton distributions since it depends on the considered physical observable and is not handy for an external user.

2.2 Normalisation uncertainty

Within the context of global analyses the treatment of normalisation uncertainties becomes relevant. Indeed, when combining datasets from independent experiments it is necessary to take into account the overall normalisation uncertainty associated with each experiment: an experiment with large normalisation uncertainty should con-

tribute less to the fit than one with a small uncertainty. Normalisation uncertainties are usually multiplicative, in the sense that each data point within the set has a normalisation uncertainty proportional to the measurement at that point. All these normalisation uncertainties are however correlated across the whole set of data points.

Fitting the data by using the complete covariance matrix, Eq. (2.3), leads to a substantial bias in the fitted value due to the fact that smaller data points are assigned a smaller uncertainty than larger ones [117]. This problem is usually avoided by using the "penalty trick": the normalisation of each dataset is treated as a free parameter to be determined during the fit, within a range restricted by the quoted experimental uncertainty. While this method gives correct results when fitting data from a single experiment, in Ref. [118] we showed that it remains biased when used in fits which combine several different datasets. In the same Ref. the so–called "t_0-method" was introduced, which provides a completely unbiased and rapidly convergent method for including the normalisation uncertainties in a fit.

Before proceeding, I define a simplified notation. In this section I consider a simple situation, where we have n measurements m_i of a single observable t with experimental uncertainties given by the covariance matrix $[\mathrm{cov}]_{ij}$ which takes the form of Eq. (2.3). With this notation, the fully correlated χ^2 reads

$$\chi^2(t) = \sum_{i=1}^{n}(t-m_i)(\mathrm{cov}^{-1})_{ij}(t-m_j) \tag{2.12}$$

and thus the least squares estimate for t is given by

$$t = \frac{\sum_{i,j=1}^{n}(\mathrm{cov}^{-1})_{ij}m_j}{\sum_{i,j=1}^{n}(\mathrm{cov}^{-1})_{ij}} \tag{2.13}$$

and its variance is found through

$$V_{tt} = \left(\frac{1}{2}\frac{\partial^2\chi^2}{\partial t^2}\right)^{-1} = \frac{1}{\sum_{i,j=1}^{n}(\mathrm{cov}^{-1})_{ij}}. \tag{2.14}$$

The features of the results obtained with this method depend on the choice of function which is minimised in order to determine the best–estimate of t.

2.2.1 The D'Agostini bias and the penalty trick

When datasets have overall multiplicative uncertainties, such as normalisation uncertainties there are biases arising from the rescaling of errors [117]. The analysis in Ref. [118] shows that with the Monte Carlo method the effect of this bias starts when

more than one experiment is considered, while in the Hessian method the bias is there also in the case of a single experiment. Here we only consider the Hessian approach in a very simple model of a single experiment with only two data points. In this case the covariance matrix is simply given by

$$(\mathrm{cov}_m)_{ij} = \begin{pmatrix} \sigma_1^2 + s^2 m_1^2 & s^2 m_1 m_2 \\ s^2 m_1 m_2 & \sigma_2^2 + s^2 m_2^2 \end{pmatrix}, \tag{2.15}$$

where s^2 is the normalisation uncertainty of the considered experiment. Therefore the χ^2 Eq. (2.12) is

$$\chi^2_m(t) = \frac{(t-m_1)^2(\sigma_2^2 + m_2^2 s^2) + (t-m_2)^2(\sigma_1^2 + m_1^2 s^2) - 2(t-m_1)(t-m_2)m_1 m_2 s^2}{\sigma_1^2\sigma_2^2 + (m_1^2\sigma_2^2 + m_2^2\sigma_1^2)s^2}. \tag{2.16}$$

Minimising this χ^2-function with respect to t yields

$$t = \frac{m_1/\sigma_1^2 + m_2/\sigma_2^2}{1/\sigma_1^2 + 1/\sigma_2^2 + (m_1-m_2)^2 s^2/\sigma_1^2\sigma_2^2} = \frac{w}{1 + (m_1-m_2)^2 s^2/\Sigma^2}, \tag{2.17}$$

where w is the weighted mean

$$w \equiv \Sigma^2 \sum_{i=1}^{n} \frac{m_i}{\sigma_i^2} \qquad \Sigma = \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \tag{2.18}$$

with $n = 2$.

It follows that, when $m_1 \neq m_2$ and $s \neq 0$, the result for t has a downward shift with respect to the unbiased result w. That this shift is a bias can be seen for instance by considering the simple case $\sigma_1 = \sigma_2 = \sigma$. Then Eq. (2.17) gives

$$t = \frac{\bar{m}}{1 + 2r^2 s^2 \bar{m}^2/\sigma^2} = \bar{m}(1 - 2r^2 s^2 \bar{m}^2/\sigma^2 + O(r^4)). \tag{2.19}$$

where we have defined

$$\bar{m} \equiv \frac{1}{2}(m_1 + m_2), \qquad r \equiv \frac{m_1 - m_2}{m_1 + m_2}. \tag{2.20}$$

Thus simply minimising the χ^2 of Eq. (2.12) with the fully correlated covariance matrix leads to a central value which is shifted downwards: for a sufficiently large s^2/σ^2 one can get an average which is lower than either of the two values which are being averaged, so one concludes that the result is biased. The variance of t is afflicted

by the same downward bias with respect to the expected $\Sigma^2 + s^2 w^2 (1 + r^2)$ result.

$$V_{tt} = \frac{\Sigma^2 + s^2 w^2 (1 + r^2)}{1 + r^2 s^2 w^2 / \Sigma^2}. \tag{2.21}$$

This is usually referred as the "d'Agostini bias", after Ref. [117] where it was studied and explained. It can be shown that this bias gets worse as the number of data points increases. The origin of this bias is clear: smaller values of m_i have a smaller normalisation uncertainty $m_i s$, and are thus preferred in the fit.

The standard way to include normalisation uncertainties in the Hessian approach by avoiding the d'Agostini bias consists of including the normalisations of the data n_i as parameters in the fit, with penalty terms to fix their estimated value close to one with variance s_i^2. This is usually referred as the "penalty trick". In Ref. [118] it is shown that, while it gives correct results for a single experiment, when used to combine results from several experiments it is biased.

In order to show this explicitly, here we first consider a single experiment, with diagonal covariance matrix, but now with an overall normalisation uncertainty with variance s^2. The value of t according to the penalty trick is obtained by minimising the error function

$$E_{\rm Hess}(t, n) = \sum_{i=1}^{n} \frac{(t/n - m_i)^2}{\sigma_i^2} + \frac{(n-1)^2}{s^2}\,. \tag{2.22}$$

where the last term is called the penalty term. The parameters t and n are determined by minimising this error function: minimising with respect to t gives $t = nw$, with w as defined in Eq. (2.18), while minimisation with respect to n fixes $n = 1$. To compute the error on the fitted quantity in this approach, one needs to evaluate the Hessian matrix:

$$V^{-1} = \frac{1}{2}\begin{pmatrix} \frac{\partial^2 \chi^2}{\partial t^2} & \frac{\partial^2 \chi^2}{\partial t \partial n} \\ \frac{\partial^2 \chi^2}{\partial n \partial t} & \frac{\partial^2 \chi^2}{\partial n^2} \end{pmatrix} = \frac{1}{\Sigma^2}\begin{pmatrix} 1 & -t \\ -t & \Sigma^2/s^2 + t^2 \end{pmatrix}. \tag{2.23}$$

The covariance matrix is obtained by inverting V^{-1}. In this way one recovers

$$V_{tt} = \Sigma^2 + s^2 w^2\,, \tag{2.24}$$

i.e. the result is unbiased. However if we turn to the more complex situation where we have several data points from different experiments, and thus with independent normalisations, $n_i = 1 \pm s_i$, we see that the penalty trick leads to a result which is

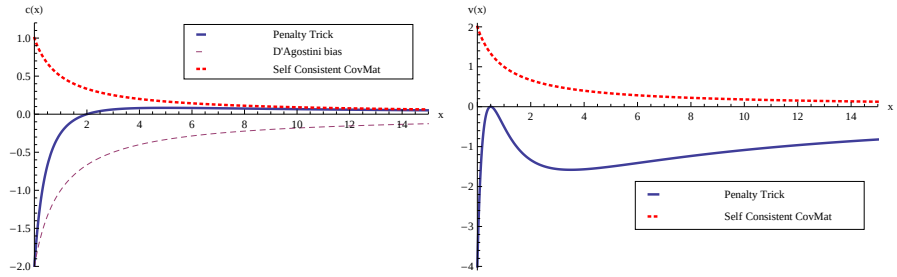


Figure 2.3: The "bias" functions $c(x)$ for the central value (left plot) and $v(x)$ for the variance (right plot), corresponding to the results obtained when the d'Agostini bias is present, using the penalty trick, and using the self-consistent covariance matrix method. The unbiased result corresponds to $c = v = 0$.

biased when all uncertainties are equal. To see it explicitly, we set up the error function

$$E_{\rm Hess}(t, n_i) = \sum_{i=1}^{n} \frac{(t/n_i - m_i)^2}{\sigma_i^2} + \sum_{i=1}^{n} \frac{(n_i - 1)^2}{s_i^2}\,, \tag{2.25}$$

where now there is a separate penalty term for each of the normalisations to be fitted. The minimum is obtained for

$$t = \frac{\sum_{i=1}^{n} \frac{m_i}{n_i \sigma_i^2}}{\sum_{i=1}^{n} \frac{1}{n_i^2 \sigma_i^2}}, \tag{2.26}$$

$$n_i = 1 + \frac{s_i^2 t}{n_i^2 \sigma_i^2} \left(\frac{t}{n_i} - m_i \right)\,. \tag{2.27}$$

These $(n + 1)$ equations are complicated nonlinear relations which must be solved for t and n_i. However it is possible to find solutions for certain special cases, which are sufficient to show that the approach is biased. In Ref. [118] two special cases are considered: the case of only two experiments with $\sigma_1 = \sigma_2 \equiv \sigma$ and $s_1 = s_2 = s$ and the case of only two experiments where the normalisation error dominates, i.e. $s_i \gg \sigma_i$. In both cases, it is shown that one gets a biased result for both central value and variance of the quantity t. This bias is more subtle than the d'Agostini bias, since it is caused by nonlinearities in the error function rather than by a consistent bias in the variances. It can thus have either sign, depending on the relative weight of statistical

and normalisation uncertainties. The bias may be written in the form

$$t = \bar{m}\big(1 + c(\frac{\sigma^2}{s^2\bar{m}^2})r^2 + O(r^4)\big), \tag{2.28}$$

$$V_{tt} = \frac{1}{n}\sigma^2 + \frac{1}{n}s^2\bar{m}^2\big(1 + r^2 + v(\frac{\sigma^2}{s^2\bar{m}^2})r^2 + O(r^4)\big), \tag{2.29}$$

with the functions c and v given by

$$c(x) = \frac{x-2}{(x+1)^2}, \qquad v(x) = -4\frac{(2x-1)^2}{(x+1)^3}. \tag{2.30}$$

The unbiased results would correspond to $c = v = 0$. The d'Agostini–biased results can also be cast in the form of Eqs. (2.28, 2.29), but now with $c(x) = -2/(x+1)$, $v(x) = 0$ (since in this case the variance is unbiased at $O(r^2)$). The "bias" functions $c(x)$ and $v(x)$ for these two cases are compared in Fig. 2.3. Thanks to the penalty trick, the bias in the central value is generally less severe, but the variance is now also biased.

A possible way out, alternative to the penalty trick of the previous section and based on a covariance matrix approach, was suggested by d'Agostini in Ref. [117], which is indicated in Fig. 2.3 as the self–consistent covariance matrix; however in Ref. [118] we showed that this method leads to results which are similar to those found using the penalty trick: for one experiment there is no bias, but for several experiments a bias arises. In the next section, we will present a new method which is free of multiple solutions, is unbiased when uncertainties are equal, but which also correctly weights the different experiments according to their normalisation uncertainties when these are unequal.

2.2.2 The t0 method

The biases found in the previous section come from the fact that the $\chi^2(t)$ function used in the fitting is no longer a quadratic function of the observable t being fitted, and thus the distribution of $\exp(-\chi^2_t(t)/2)$ is no longer Gaussian. The dependence of the covariances of the data on t^2 distorts the shape of the χ^2, and thus introduces a bias. The idea proposed in Ref. [118] in order to avoid this is to hold the covariance matrix fixed when performing the fitting. One can do this by evaluating the covariance matrix using some fixed value t_0 rather than t. The value of t_0 can then be tuned independently to be consistent with the value of t obtained from the fit. The basic idea is then to determine t_0 self–consistently in an iterative way.

I now show that this procedure gives unbiased results for equal normalisation uncertainties. Here I concentrate on the Hessian method; a more general discussion can be found in Ref. [118]. I also show that the iterative determination of t_0 converges very rapidly, and thus that the method is also practical.

In the case of a single experiment, in place of Eq. (2.12) the covariance matrix proposed in Ref. [118] is chosen to be

$$(\mathrm{cov}_{t_0})_{ij} = (\mathrm{cov})_{ij} + t_0^2 s^2, \tag{2.31}$$

where t_0 should be viewed as a guess for t, to be fixed beforehand. In a Hessian approach the χ^2 is then

$$\chi^2_{t_0}(t) = \sum_{i,j=1}^{n} (t - m_i)(\mathrm{cov}^{-1}_{t_0})_{ij}(t - m_j), \tag{2.32}$$

and minimisation is trivial:

$$t = \frac{\sum_{i,j=1}^{n} (\mathrm{cov}^{-1}_{t_0})_{ij} m_j}{\sum_{i,j=1}^{n} (\mathrm{cov}^{-1}_{t_0})_{ij}}, \tag{2.33}$$

while

$$V_{tt} = \frac{1}{\sum_{i,j=1}^{n} (\mathrm{cov}^{-1}_{t_0})_{ij}}. \tag{2.34}$$

For the special case in which the data have uncorrelated statistical errors only we recover $t = w$, independent of the value chosen for t_0, while

$$V_{tt} = \Sigma^2 + s^2 t_0^2. \tag{2.35}$$

This reduces to Eq. (2.24), but only if we tune $t_0 = t$.

When we have n independent experiments (each with one data point), the covariance matrix proposed in Ref. [118] is given by

$$(\mathrm{cov}_t)_{ij} = (\sigma_i^2 + s_i^2 t_0^2)\delta_{ij}, \tag{2.36}$$

so the χ^2 is now

$$\chi^2 = \sum_{i=1}^{n} \frac{(t - m_i)^2}{\sigma_i^2 + s_i^2 t_0^2}, \tag{2.37}$$

whence we have the single solution

$$t = \frac{\sum_{i=1}^{n} \frac{m_i}{\sigma_i^2 + s_i^2 t_0^2}}{\sum_{i=1}^{n} \frac{1}{\sigma_i^2 + s_i^2 t_0^2}}, \tag{2.38}$$

and

$$V_{tt} = \frac{1}{\sum_{i=1}^{n} \frac{1}{\sigma_i^2 + s_i^2 t_0^2}}. \tag{2.39}$$

For the special case $s_i = s$, $\sigma_i = \sigma$, these reduce to $t = \bar{m}$ and $V_{tt} = \frac{1}{n}(\sigma^2 + s^2 t_0^2)$ as they should: the fit is unbiased, and the variance is correctly estimated provided only that $t_0 = t$. The same can be shown in the case when normalisation uncertainties dominate. The result may be generalised to any general situation that one has to deal with when performing a parton fit [118].

For this reason the t_0-covariance matrix Eq. (2.31) gives a χ^2 function that can be used to obtain unbiased fits, with t_0 controlling the relative balance between statistical and normalisation errors, both in the Hessian and Monte Carlo method. The remaining difficulty with this approach is that t_0 is not determined self–consistently within the minimisation, but rather must be fixed beforehand. Clearly if the value chosen is incorrect, this may itself lead to an incorrect fit.

However, the dependence on t_0 is rather weak. That this is the case is qualitatively clear: firstly t_0 only determines the uncertainties, so an error we make in t_0 is a second order effect; furthermore all dependence on t_0 cancels when all the m_i are equal, when σ_i and s_i are equal, or when normalisation errors dominate over statistical, or indeed vice versa. In Ref. [118] this is proved in a quantitative way.

It follows that t_0 can be determined iteratively: a first determination of t is performed with a zeroth–order guess for t_0, such as, for example, $t_0 = 0$. The result for t thus obtained is used as t_0 for a second iteration, and so on. In Ref. [118] the method was applied to the NNPDF1.2 parton fit [70], showing that this procedure converges rapidly: two iterations were sufficient to get a stable result.

So far we have taken a naive approach, in which we constructed least squares estimators, and minimised them with respect to the theoretical prediction t. To conclude, it is interesting to consider instead how we might construct a likelihood function for the measurements and normalisations, and thus whether any of the above mentioned estimators are maximum likelihood estimators.

Consider for definiteness the case of several experiments, with measurements m_i and variances σ_i. In presenting the measurements in this form, there is an underlying assumption that the measurements are Gaussian. The likelihood is then simply de-

fined as the probability that the measurements take the observed values given a certain theoretical value t for such measurements:

$$P(m|t) = N \exp\Big(-\frac{1}{2} \sum_{i=1}^{n} \frac{(m_i - t)^2}{\sigma_i^2} \Big), \tag{2.40}$$

where N is some overall normalisation factor for the probability, dependent on σ_i but not on m_i or t. Of course this probability is just $N \exp(-\frac{1}{2}\chi^2(t))$, so the maximum likelihood estimator is found by minimising the χ^2, i.e. it is the same as the least squares estimator.

Now consider what happens when there are also normalisation uncertainties n_i with variances s_i. In the absence of further information it is natural to assume these are also Gaussian. Furthermore they are clearly entirely independent of the measurement uncertainties, since the physics involved in determining the normalisation is generally quite independent of that related to the measurements m_i or indeed the theoretical value t. Thus the total likelihood should factorise: $P(m, n|t) = P(m|t)P(n)$. The maximum likelihood estimators obtained from $P(m, n|t)$ and $P(m|t)$ should thus be the same.

The adoption of the fully–correlated χ^2-function, assumes for the likelihood

$$P_m(m|t) = N \exp\Big(-\frac{1}{2} \sum_{i=1}^{n} \frac{(m_i - t)^2}{\sigma_i^2 + s_i^2 m_i^2} \Big). \tag{2.41}$$

This is incorrect, because it is no longer Gaussian in m_i, and indeed not even properly normalised (to normalise it, N must depend on t, and then maximising $P_m(m|t)$ is no longer the same as minimising the χ^2). Thus the fully–correlated χ^2-function is not a maximum likelihood estimator, principally because the probability distribution it assumes is skewed by the normalisation uncertainties.

Similarly the Hessian method with penalty trick, by adopting the error function Eq.(2.25), assumes for the likelihood

$$P_n(m, n|t) = N \exp\Big(-\frac{1}{2} \sum_{i=1}^{n} \Big[\frac{(m_i - t/n_i)^2}{\sigma_i^2} + \frac{(n_i - 1)^2}{s_i^2} \Big] \Big). \tag{2.42}$$

This is also incorrect, because now the likelihood, while Gaussian in m_i, is not Gaussian in n_i, and furthermore cannot be factorised into a product $P(m|t)P(n)$. Once again it is not properly normalised: N must depend on t. So this too does not give us a maximum likelihood estimator. The main problem here is that, since the model for the likelihood does not factorise, the assumption of a common theoretical result t introduces artificial correlations between the normalisation measurements n_i; this

leads to biases since these measurements are in principle completely independent. This is why the penalty trick, while giving correct results for a single experiment with only one overall normalisation uncertainty, fails when applied to several independent experiments.

Finally consider the t_0–method discussed in this section, which takes as its starting point the χ^2-function Eq.(2.37). Here the assumption for the likelihood is

$$P_{t_0}(m|t) = N \exp\Big(-\frac{1}{2}\sum_{i=1}^{n} \frac{(m_i - t)^2}{\sigma_i^2 + t_0^2 s_i^2}\Big) \tag{2.43}$$

This is now correct: it is Gaussian in m_i, properly normalised, and all we have to do to make sure we get the correct answer is choose t_0 consistently. Note that it can alternatively be formulated as

$$P_{t_0}(m,n|t) = N \exp\Big(-\frac{1}{2}\sum_{i=1}^{n} \Big[\frac{(m_i - t/n_0)^2}{\sigma_i^2 + t_0^2 s_i^2} + \frac{(n_i - n_0)^2}{s_i^2}\Big]\Big). \tag{2.44}$$

This is rather like the penalty trick Eq.(2.42), but with the n_i in the first term replaced with its best estimate n_0, just as in going from m-cov Eq.(2.41) to t_0-cov Eq.(2.43) we replace m_i in the denominator with our best estimate t_0. Note of course that actually $n_0 = 1$: this is the natural choice for presenting the data that all experimentalists choose. Eq.(2.44) is also a good definition of the likelihood: it is Gaussian in both m_i and n_i, the normalisation is determined quite independently of t, and it factorises correctly into the product $P_{t_0}(m|t)P(n)$. Thus when we minimise with respect to t, $P_{t_0}(m|t)$ and $P_{t_0}(m,n|t)$ give the same maximum likelihood estimator for t, as they should.

Since the t_0-method yields the maximum likelihood estimator, it possesses all the nice asymptotic properties of maximum likelihood estimators: in particular it is consistent and unbiased.

To demonstrate the practical application of the t_0-method, in Ref. [70] the method was implemented it in the NNPDF1.2 parton fit [70]. This confirmed the rapid convergence of the technique, showed that the inclusion of normalisation uncertainties can lead to a small improvement in the quality of the fit through the resolution of tensions between datasets, and moreover that where these tensions are significant this can lead to a subsequent reduction in PDF uncertainties, as it is shown in Chap. 4.

2.3 Benchmark fits

In this section I present several benchmarks studies performed in the past years in order to clarify some of the aspects involved in the fitting of parton distribution functions. In order to disentangle them, common theoretical and model settings are adopted in order to better assess the differences between different procedures. Here I present some benchmarks whose aim is to comprehend the study of the statistical consistency of a parton determination under the inclusion of new data (HERA–LHC and H1–NNPDF benchmarks), the difference between Monte Carlo and Hessian approaches (H1 benchmark) and the impact of the parametrisation choice (H1–NNPDF benchmark).

2.3.1 HERA–LHC benchmarks

One of the main drawbacks of the analyses based on the fixed functional form parametrisation and the use of the Hessian method, is that sometimes the addition of new experimental data to the fit, therefore the increase of information, leads to an increase rather than to a decrease of the PDFs error bands because the new data require the use of a more general parametrisation. This makes a statistical interpretation of the uncertainty bands on parton distribution difficult. The issue was raised for the first time in Ref. [101] in the context of a benchmark analysis performed by the MRST [119] and the Alekhin [73] collaborations.

In order to compare the two parton determination procedures, the same datasets, the same cuts and theoretical prescriptions were used. To be conservative only the data in Tab. 2.2 were included in both fits and cuts of $Q^2 = 9\,\mathrm{GeV}^2$ and $W^2 = 15\,\mathrm{GeV}^2$ were applied in order to avoid the influence of higher twist. Both the Alekhin and MRST

dataset	Data points	Observable	Ref.
ZEUS97	206	F_2^p	[120]
H1lowx97	77	F_2^p	[121]
NMC	95	F_2^p	[79]
NMC_pd	73	F_2^d/F_2^p	[80]
BCDMS	322	F_2^p	[77]
Total	773		

Table 2.2: Data points used in the HERA–LHC benchmark after kinematic cuts of $Q^2 > 9\,\mathrm{GeV}^2$ and $W^2 > 15\,\mathrm{GeV}^2$ are applied.

benchmark partons are determined by using the Hessian method with $\Delta\chi^2 = 1$ and

parametrising PDFs at the starting scale $Q_0^2 = 1\ \mathrm{GeV}^2$ according to the following functional form:

$$x\, f_i(x, Q_0^2) = A_i(1-x)^{b_i}(1 + \epsilon_i x^{0.5} + \gamma_i x)x^{a_i}\ . \tag{2.45}$$

Four independent input PDFs (u and d valence, the sea and the gluon) are parametrised at the initial scale, the light sea asymmetry is set to zero ($\bar{u} = \bar{d}$) and there is no independent strange PDFs, which is rather set to be a fixed fraction of the non–strange sea ($s(x, Q_0^2) + \bar{s}(x, Q_0^2) = 0.5[\bar{u}(x, Q_0^2) + \bar{d}(x, Q_0^2)]$); finally ϵ_i and γ_i set to zero for the sea and gluon distributions. Hence, there is a total of 13 free PDF parameters plus $\alpha_s(M_Z)$ (which is fitted as a free parameter) after imposing sum rules. The common theoretical assumptions are

- NLO perturbative QCD in the $\overline{\mathrm{MS}}$ renormalisation and factorisation scheme;
- zero-mass variable flavour number scheme with quark masses $m_c = 1.5\ \mathrm{GeV}$ and $m_b = 4.5\ \mathrm{GeV}$;
- momentum and valence sum rules imposed;
- iterated solution of evolution equations.

The only differences lies in the treatment of the errors, since all details on correlations between errors is included for the Alekhin fit while in the MRST fit the correlations are only partially included.

From the plots in Fig. 2.4 the generally good agreement between the parton distributions derived in the two fits is apparent. However the comparison between the benchmark partons and the published partons from a global fit is more problematic. If the statistical analysis is correct, the benchmark partons should agree with the global partons within their uncertainties, which they do not, as one can see in Fig. 2.5. The disagreement is remarkable, not only in the extrapolation region, where no data constrain the PDFs behaviours, which are hence driven by theoretical prejudice in the absence of a flexible enough parametrisation, but also in the data region. It is also striking that the uncertainties in the two sets are rather similar despite the fact that the uncertainty on the benchmark partons is obtained from allowing $\Delta\chi^2 = 1$ in the fit while that for the MRST01 partons is obtained from $\Delta\chi^2 = 50$. Moreover the uncertainty in the benchmark gluon is much smaller than in the MRST01 gluon, despite the much smaller amount of low-x data in the fit for the benchmark partons. This comes about as a result of the lack of flexibility in the benchmark parametrisation of the gluon.

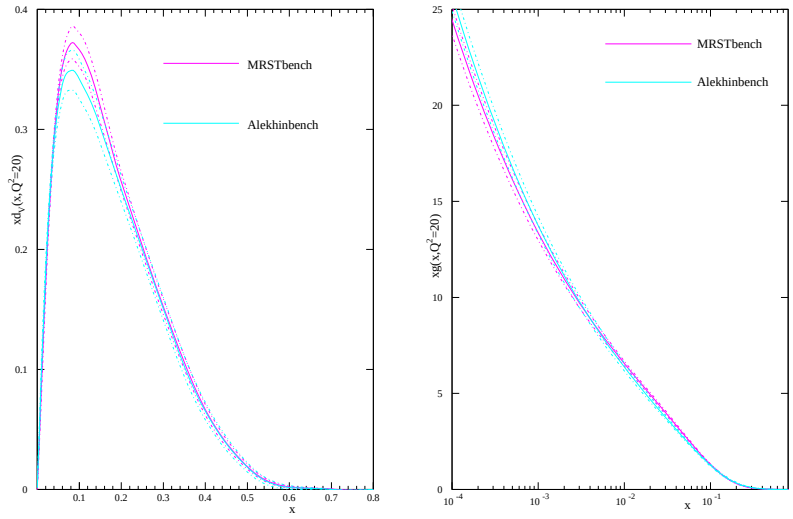


Figure 2.4: Left plot: $xd_V(x)$ at $Q^2 = 20$ GeV^2 from the MRST benchmark partons compared to the one from the Alekhin benchmark partons. Right plot: $xg(x)$ at $Q^2 = 20$ GeV^2 from the MRST benchmark partons compared to the one from the Alekhin benchmark partons. Taken from Ref. [101].

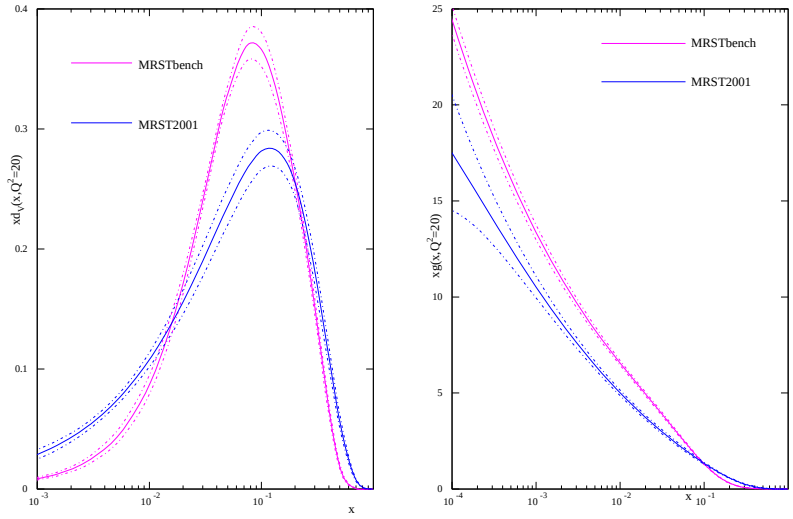


Figure 2.5: Left plot: $xd_V(x)$ at $Q^2 = 20$ GeV^2 from the MRST benchmark partons compared to that from the MRST01 partons. Right plot: $xg(x)$ at $Q^2 = 20$ GeV^2 from the MRST benchmark partons compared to that from the MRST01 partons. Taken from Ref. [101].

The fact that partons extracted using a very limited dataset are completely incompatible with those obtained from a global fit, even allowing for the uncertainties, implies that the inclusion of more data from a variety of different experiments moves the central values of the partons in a manner indicating either that the different experimental data are inconsistent with each other, or that the theoretical framework is inadequate for correctly describing the full range of data. This clearly illustrates the problems in determining the true uncertainty on parton distributions.

In order to understand such behaviour, a similar exercise was performed again in Ref. [122]. Here the comparison is extended to include a NNPDF fit and the recent MSTW 2008 fit. As far as the MSTW analysis is concerned, the benchmark analysis is much more closely aligned to the global analysis than was the case for the previous benchmark compared to the MRST global analysis. Indeed the MSTW08 analysis has several features which are different with respect to MRST2001. First of all the input parametrisation includes three additional free parameters associated with $(\bar{d} - \bar{u})$ and four additional free parameters associated with strangeness, giving a total of 20 eigenvectors, seven more than in MRST01. The choice of tolerance, $T = \sqrt{\Delta\chi^2}$, that in MRST01 was set to one for 1σ uncertainties and to $\sqrt{50}$ for a 90% confidence level (C.L.) uncertainty band, has been substituted by a new procedure described in Sect. 3.1.4. The NNPDF approach instead is going to be discussed in details in the next chapter. Here it is important to outline that, being based on the Monte Carlo determination of PDFs uncertainties and on the use of a redundant parametrisation provided by neural networks, it does not have to introduce any tolerance and the PDFs parametrisation used for the benchmark and the global fits are the same.

In Ref. [122] it is shown that the PDFs from the NNPDF and MSTW benchmark fits are compatible, especially in the data region, whereas in the extrapolation region there are some discrepancies. Differences are also sizeable in the estimation of uncertainties, probably due to the differences in parametrisation. Here we are mainly interested in the comparison betwen the benchmark and the global fits for each of the two collaborations. In Fig. 2.6 the NNPDF benchmark fit is compared to the NNPDF1.0 reference fit of Ref. [68] (NNPDF global), while in Fig. 2.7 the MSTW benchmark fit is compared to the MRST01 [119] (MRST global) and MSTW08 [43] global fits (MSTW global). For NNPDF, global and benchmark fits remain compatible within their respective error bands. The NNPDF benchmark fit has a sizeably larger error band than the reference, as one would expect from a fit based on a smaller set of (compatible) data. For MSTW, we first notice that the MSTW benchmark set has larger uncertainty bands than the MRST benchmark set and than each of the sets obtained from global fits. Consequently, the MSTW benchmark PDFs are generally more consistent with the MSTW global fit sets than the corresponding comparison

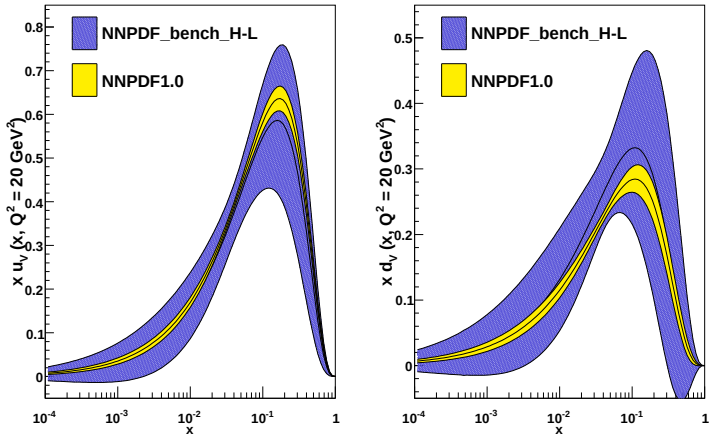


Figure 2.6: Comparison of the NNPDF benchmark and NNPDF1.0 reference fits for the u-valence (left) and d-valence (right) at $Q^2 = 20$ GeV^2.

between MRST benchmark PDFs and global fit PDFs shown in Ref. [101], largely due to the more realistic uncertainties in the MSTW benchmark. Unlike the NNPDF group, the MSTW group sees some degree of incompatibility between the benchmark PDFs and the global fit PDFs for the valence quarks, particularly in the case of the down valence, see Fig. 2.7.

2.3.2 H1 benchmark: determination of uncertainties

In this section we present another benchmark based on the H1–2000 parton fit [123]. It compares a new version of this fit, in which uncertainty bands are determined using a Monte Carlo method introduced in Ref. [110], to the reference fit, where uncertainty bands are obtained using the standard Hessian method. The main motivation of this benchmark is to study the impact of possible non-Gaussian behaviour of some experimental uncertainties and, more generally, the dependence on the error treatment.

As we discussed in Sect. 3.1, standard error estimation of PDFs relies on the assumption that all errors follow Gaussian statistics. However, this assumption may not always be correct. Some systematic uncertainties such as luminosity and detector acceptance follow rather a log-normal distribution [101]. Compared to the Gaussian case, the lognormal distribution which has the same mean and root mean square, is asymmetric and has a shifted peak. The non-Gaussian behaviour of the experimental uncertainties could lead to an additional uncertainty of the resulting PDFs.

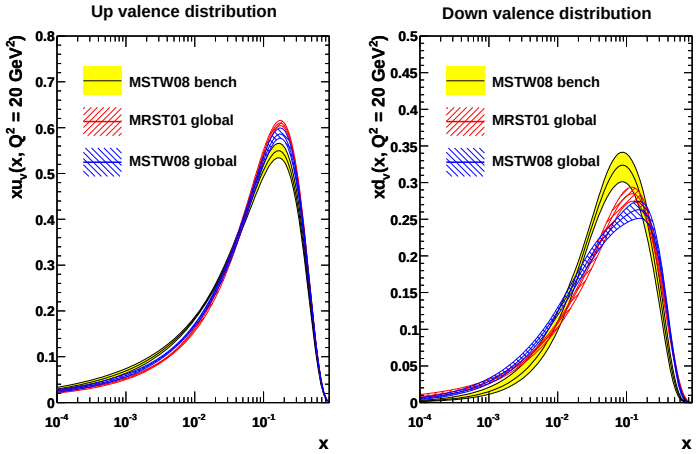


Figure 2.7: Comparison of the MSTW benchmark and MRST/MSTW global fits for the u-valence and d-valence at $Q^2 = 20\ \mathrm{GeV}^2$. All uncertainties shown correspond to 1σ bands.

In the Monte Carlo (MC) method presented earlier one does not have to rely on the Gaussian distribution of the uncertainties. The MC technique consists in generating replicas of the initial datasets which have the central value of the experimental observables, fluctuating within its systematic and statistical uncertainties taking into account all point to point correlations. Various assumptions can be considered for the error distributions. When dealing with the statistical and point to point uncorrelated errors, one could allow each data point to randomly fluctuate within its uncorrelated uncertainty assuming either Gaussian, lognormal, or any other desired form of the error distribution.

The MC method is tested by comparing the standard error estimation of the PDF uncertainties with the MC techniques by assuming that all the errors (statistical and systematic) follow Gaussian (normal) distribution. The good agreement between the methods is manifest in the left–hand side plot in Fig. 2.9. Assuming only a Gaussian distribution of all errors, the results agree well with the standard error estimation.

In the same analysis two cases are considered, which may represent most faithfully the error distributions: a lognormal distribution for the luminosity uncertainty and the rest of the errors are set to follow the Gaussian shape, versus a lognormal distributions for all the systematic errors and the statistical errors are set to follow the Gaussian distributions. The results of this comparison, shown in Fig. 2.8, shows that for the

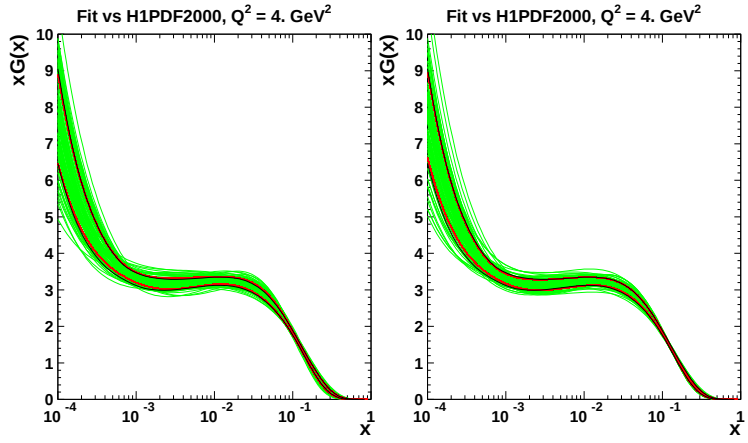


Figure 2.8: Comparison between errors on PDFs obtained via standard error calculation (black) where Gauss assumption is used, and errors obtained via Monte Carlo method (red) where luminosity uncertainty is allowed to fluctuate according to lognormal distributions and all the other uncertainties follow the Gaussian distribution (left), and where all the systematic uncertainties are allowed to fluctuate according to lognormal distributions (right). Only the gluon PDF is shown, where the errors are larger. The green lines show the spread of the N individual fits. Taken from Ref. [101]

precise H1 HERA-I data the effect of using a lognormal distribution is similar to using pure Gaussian distribution case.

2.3.3 H1-NNPDF benchmark: dependence on parametrisation

The third benchmark is a further elaboration on the benchmark presented in the previous subsection, extended to include the NNPDF fit, which also uses a Monte Carlo method for error esteem. The main purpose of this benchmark is to compare two fits (H1 and NNPDF) which have the same error treatment but different parton parametrisations. The inclusion in this benchmark of the NNPDF fit is also interesting because it allows a comparison of a fit based on a very consistent set of data coming from the H1 collaboration only, to fits which include all DIS datasets.

This analysis is based on all the DIS inclusive data by the H1 collaboration from the HERA-I run. A kinematic cut of $Q^2 > 3.5\ \mathrm{GeV}^2$ is applied to avoid any higher twist effect. The data points used in the analysis are summarised in Table 2.3.

dataset	Data points	Observable	Ref.
H197mb	35	$\tilde{\sigma}^{NC,+}$	[121]
H197lowQ2	80	$\tilde{\sigma}^{NC,+}$	[121]
H197NC	130	$\tilde{\sigma}^{NC,+}$	[124]
H197CC	25	$\tilde{\sigma}^{CC,+}$	[124]
H199NC	126	$\tilde{\sigma}^{NC,-}$	[125]
H199CC	28	$\tilde{\sigma}^{CC,-}$	[125]
H199NChy	13	$\tilde{\sigma}^{NC,-}$	[125]
H100NC	147	$\tilde{\sigma}^{NC,+}$	[123]
H100CC	28	$\tilde{\sigma}^{CC,+}$	[123]
Total	612		

Table 2.3: Data points used in the H1 benchmark after kinematic cuts of $Q^2 > 3.5\ \mathrm{GeV}^2$.

The theoretical assumptions are similar to those set in the HERA–LHC benchmark. Both the H1 and NNPDF methodologies are based on the Monte Carlo method to determine uncertainties. They differ in the way PDFs are parametrised: H1 parametrises PDFs according to a polynomial functional form which describes five independent PDFs with 10 free parameters after sum rules are imposed. NNPDF parametrises each independent parton distribution by a neural network characterised by a large number of parameters, 37 for each PDF, therefore $34 \cdot 4 = 136$ parameters in total.

In Ref. [122] the results of the NNPDF benchmark are compared to the NNPDF1.0 reference fit results. The general features of the benchmark are analogous to those of the HERA–LHC benchmark discussed in the previous section, with some effects being more pronounced because the benchmark dataset is now even smaller. However here we focus on the comparison between H1 and NNPDF benchmark fits with the purpose of understanding the impact of the respective methodologies.

The quality of the two fits is comparable, the differences in χ^2 being compatible with statistical fluctuations. In the region where experimental information is mostly concentrated, specifically for the u_v distribution over all the x-range and for the $\bar{d}$ and the d_v distributions in the small-x range, the results of the two fits are in good agreement, though the H1 uncertainty bands are generally smaller.

In the region where experimental information is scarce or missing, sizable differences are found. Specifically, in these regions NNPDF uncertainties are generally larger than H1 bands: the width of the uncertainty band for the H1 fit varies much less between the data and extrapolation regions than that of the NNPDF benchmark. Also, the H1 central value always falls within the NNPDF uncertainty band, but the NNPDF central value tends to fall outside the H1 uncertainty band whenever the central values differ

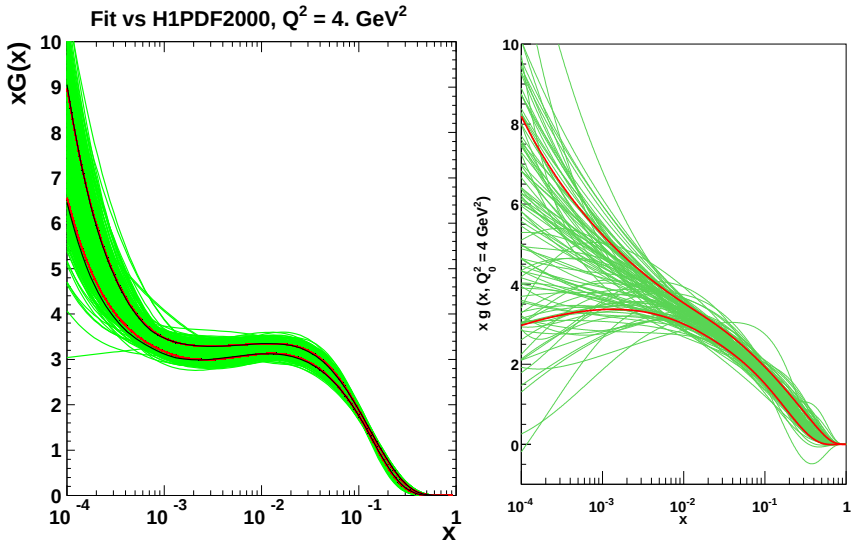


Figure 2.9: Left: The Monte Carlo set of gluon PDFs for the H1 benchmark. The comparison between the standard error calculations and the Gaussian error distribution for the gluon PDF shows a good agreement. Green lines represent the spread of Monte Carlo generated allowances for the errors, and the red lines are the RMS of this spread. The black lines correspond to the standard error calculations of the PDF errors. Right: The Monte Carlo set of gluon PDFs for the NNPDF benchmark. The red lines show the 1σ contour calculated from the Monte Carlo set. Taken from Ref. [122]

significantly. In Fig. 2.9 the respective full Monte Carlo PDF sets in the case of the gluon distribution is shown. It shows that the NNPDF parametrisation has a greater flexibility than the fixed functional form used by the HERAPDF collaboration.

Chapter 3

The NNPDF approach to parton fitting

In this chapter I present the determination of the nucleon parton distributions from the analyses performed within the NNPDF collaboration over the last three years. The general aim of the NNPDF approach is to determine objectively both the value and the uncertainty of a set of functions from a discrete set of many independent (and possibly incompatible) experimental measurements. This is achieved thanks to the combination of several ingredients which I describe in details in the first part of the chapter: the Monte Carlo sampling of the space of data, the redundant parametrisation provided by Neural Networks and the fitting strategy based on a cross–validation method.

I then turn to present the results obtained in the recent NNPDF analyses. On top of the results on the shapes and the uncertainties of PDFs, I discuss the compatibility between datasets, the statistical consistency of the results and their phenomenological relevance. All results are compared to other recent parton sets, differences and similarities are discussed. To conclude, I give an outlook of the upcoming analyses.

3.1 The NNPDF method

Even though over the last decade a huge progress has been made in the determination of sets of parton distributions with uncertainties [59, 41, 62, 63, 111, 75, 42, 43], many of the problems raised in Ref. [110] are still only partly solved. In particular, the benchmark comparisons presented in Sect. 3.3 have shown that the partonic uncertain-

ties are not easily interpreted in a statistical sense, given that they are to a significant amount determined by theoretical or phenomenological expectations. The results of the benchmark might be the consequences of incompatibilities between data or of inadequacy of the theory used to describe them or of both of them. The standard parton determination method based on fitting a particular functional form does not seem to be sufficiently flexible to ascertain whether this is the case, or whether the difficulties are due to an intrinsic limitation of the methodology.

Moreover, whereas uncertainty bands for parton determinations based on restricted data sets [75] are obtained by using standard error propagation of 1σ contours, those for global fits which include a large variety of data [59, 63] are obtained on the basis of a tolerance, determined by studying the compatibility of the data with each other and with the underlying theory. The question whether this factor is necessary and how it can be derived in a consistent way has been raised in several occasions [101, 126]. In particular, given that the effect of this tolerance is equivalent to multiplying experimental errors by a factor between four and six, the use of the tolerance might inflate the error of PDFs in regions where data do constrain them.

The drawbacks of the traditional method to extract PDFs have stimulated the formulation of alternative approaches. The one developed by the NNPDF collaboration has proved to be successful and to provide a statistically–sound determination of Parton Distribution Functions. The method is based on a Monte Carlo approach, with neural networks used as unbiased interpolants: the use of neural networks provides a robust and flexible parametrisation of the parton distributions at the initial scale, while the use of the Monte Carlo sampling allows to evaluate all quantities, such as the uncertainty or the correlation of PDFs, in a statistically-sound way.

A schematic representation of the NNPDF approach is given in Fig. 3.1. It involves two stages. First one generates a Monte Carlo ensemble of replicas of the original data points included into the fit. The ensemble is generated with the probability distribution of the data and contains all the available experimental information. It must be large enough that the statistical properties of the data are reproduced to the desired accuracy. Each element in the Monte Carlo set is a replica of the experimental data and contains as many data points as are originally available. Whether the given ensemble has the desired statistical features can be verified by means of statistical standard tests by comparing quantities calculated from it with the original properties of the data.

In the second stage, a set of parton distributions is constructed from each replica of the data. Each independent PDF at a given scale is parametrised by an individual neural network. Physical observables are computed from parton distributions by evolving the initial scale parton distributions to the scales of the experimental measurements by using the DGLAP evolution equations. Physical observables are computed by convo-

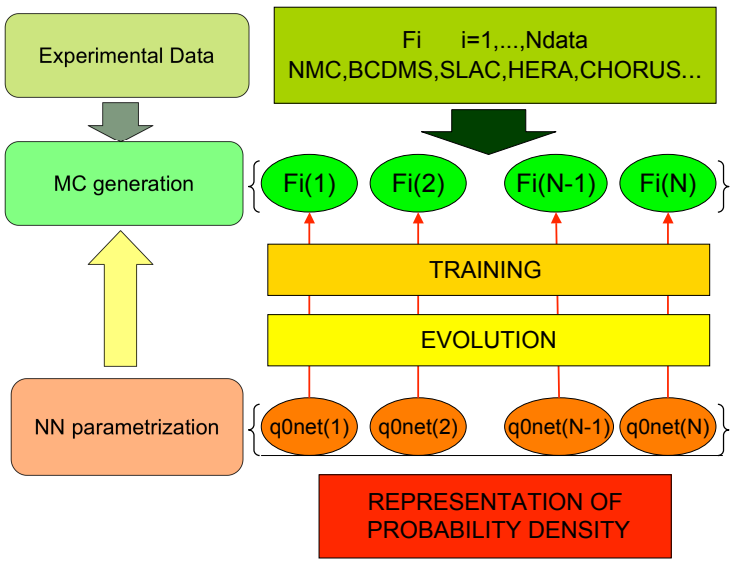


Figure 3.1: Schematic representation of the NNPDF approach.

luting the evolved parton distributions with hard partonic cross sections. The best fit set of parton distribution is determined by comparing the theoretical computation of the observable for a given PDF set with their replica experimental values by evaluating a suitable figure of merit. Both the minimisation and the determination of the best–fit in a big space of parameters such as the one spanned by the neural network parameters are delicate issues that I am going to describe in details in the next section.

The ensemble of these best fit PDFs, which contains as many elements as the number of replicas of the data that were generated, is the final result of the parton determination. The experimental values in each replica will fluctuate according to their distribution in the Monte Carlo ensemble and the best fit PDFs will fluctuate accordingly for each replica. Even though individual PDF replicas might fluctuate significantly, averaged quantities like central value and 1σ error bands are smooth inasmuch as the size of the ensemble increases. This is shown in Fig. 3.2 where the shape of individual replicas and the error band computed for two sets of 25 and 100 replicas of the gluon PDF in Ref. [68] are displayed.

An important feature of the approach is that many issues of parton determination can be addressed using standard statistical tools. For example, the stability of results upon

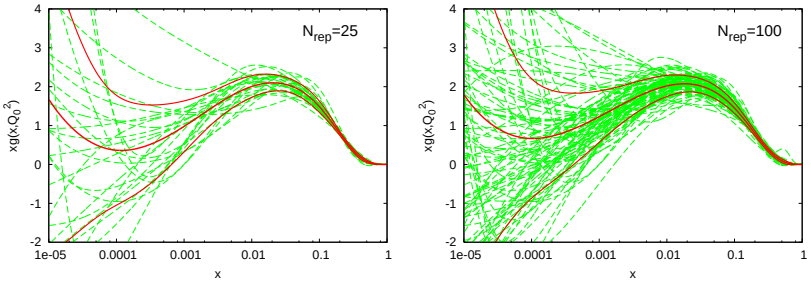


Figure 3.2: Sets of 25 replicas (left) and 100 replicas (right) of the gluon distribution at the initial scale $Q_0^2 = 2$ GeV^2 from Ref. [68]. The solid red (dark) lines show the average and 1σ intervals computed from the given sets.

a change of parametrisation can be verified by computing the distance between results in units of their standard deviation. An advantage of the Monte Carlo approach is that it is no more difficult to do this for uncertainties, correlation coefficients or even more indirect quantities, than it is for central values of physical observables. Likewise, it is possible to verify that fits performed by removing data from the set have wider error bands but remain compatible within these enlarged uncertainties, and so forth. The reliability of the results can thus be assessed directly.

In the next subsections I describe in detail each of the main ingredients of the NNPDF method and I show the improvements that have been performed along with the subsequent analyses.

3.1.1 The Monte Carlo sampling of the probability density

Given a PDF or a quantity depending on PDFs $\mathcal{O}(\{f_i\})$, its average is given by the integration in the functional space $V(\{f_i\})$ spanned by the parton distributions, weighted by a suitably defined probability measure of all possible functions describing PDFs at a reference scale:

$$\langle \mathcal{O} \rangle = \int_V d\mathcal{P}[\{f_i\}]\, \mathcal{O}(\{f_i\}). \tag{3.1}$$

In the NNPDF approach the probability measure is represented by a Monte Carlo sample in the space of PDFs defined in two steps: first the generation of an ensemble of replicas of the original data set, such that it reproduces the statistical distribution of the experimental data, followed by its projection into the space of PDFs through

the fitting procedure. Notice that all theoretical assumptions represent a prior for the determination of such probability measure.

The ensemble in the space of data has to contain the available experimental information. In practice, most data are given with multigaussian probability distributions of statistical and systematic errors, described by a covariance matrix and a normalisation error. In such cases this is the distribution that will be used to generate the pseudodata. However, any other probability distribution can be used if and when required by the experimental data[1]. The statistical sample in the space of data is obtained by generating $N_{\rm rep}$ artificial replicas of data points following a multi-Gaussian distribution centred on each data point with the variance given by the experimental uncertainty. More precisely, given a data point $F_{I,p}^{(\rm exp)} \equiv F_I(x_p, Q_p^2)$ we generate $k = 1, \ldots, N_{\rm rep}$ artificial points $F_{I,p}^{({\rm art})(k)}$ as follows

$$F_{I,p}^{({\rm art})(k)} = S_{p,N}^{(k)} F_{I,p}^{(\rm exp)} \left(1 + \sum_{l=1}^{N_c} r_{p,l}^{(k)} \sigma_{p,l} + r_p^{(k)} \sigma_{p,s} \right) \, , \; k = 1, \ldots, N_{\rm rep} \, , \quad (3.2)$$

where

$$S_{p,N}^{(k)} = \prod_{n=1}^{N_a} \left(1 + r_{p,n}^{(k)} \sigma_{p,n} \right) \prod_{n=1}^{N_r} \sqrt{1 + r_{p,n}^{(k)} \sigma_{p,n}} . \quad (3.3)$$

The variables $r_{p,l}^{(k)}$, $r_p^{(k)}$, $r_{p,n}^{(k)}$ are all univariate Gaussian random numbers that generate fluctuations of the artificial data around the central value given by the experiments. For each replica k, if two experimental points p and p' have correlated systematic uncertainties, then $r_{p,l}^{(k)} = r_{p',l}^{(k)}$, i. e. the fluctuations due to the correlated systematic uncertainties are the same for both points. A similar condition on $r_{p,n}^{(k)}$ ensures that correlations between normalisation uncertainties are properly taken into account.

It is possible to define appropriate statistical estimators able to quantify the accuracy of the statistical sampling obtained from a given ensemble of replicas [67]. They are defined in Appendix B. In Fig. 3.3, I show the scatter plot for mean values and errors evaluated and averaged over all the PDFs, for a sample of 10, 100 and 1000 replicas. It is clear that 10 replicas are enough to reproduce the central values within 1% accuracy and that one needs 100 replicas for reaching the same accuracy in the evaluation of uncertainties. Using these estimators, one may verify that a Monte Carlo sample of pseudo-data with $N_{\rm rep} = 1000$ is sufficient to reproduces also the correlations of experimental data with a 1% accuracy for all experiments. As an example, results

[1] For instance in Chap. 3 it was observed that for the precise H1 HERA-I data lognormal and Gaussian distributions produce similar results.

for the estimators computed from a sample of $N_{\rm rep} = 1000$ replicas are shown in Table 3.1 for some of the data sets included into the NNPDF fits.

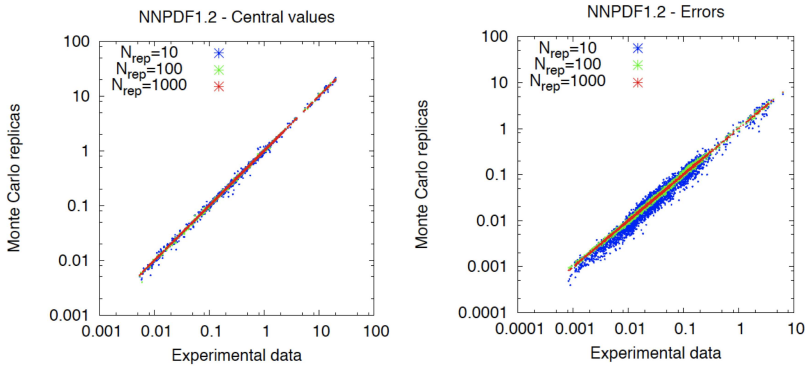


Figure 3.3: Scatter plot for mean values and errors evaluated and averaged over all the ensemble of MC replicas of the data according to the statistical estimators defined in Appendix B. Results are shown for a sample of 10, 100 and 1000 replicas in the NNPDF1.2 analysis.

3.1.2 The Neural Network parametrisation

The Monte Carlo technique adopted to propagate the experimental error into the space of PDFs is completely independent of the method used to parametrise parton distributions; it might well be used along with polynomial functional forms [101]. On the other hand, in order to get a faithful determination of parton distributions, one ought to make sure that the chosen functional form is redundant enough not to introduce a theoretical bias which would artificially reduce parton uncertainty in regions where data do not constrain enough PDFs. There are several ways for obtaining such a redundant parametrisation. One may use some clever polynomial basis, or more refined tools such as the self–organising maps [127]. Our choice was to adopt a Neural Networks parametrisation.

Neural Networks provide a redundant and minimally biased parametrisation of PDFs, the only theoretical assumption being smoothness, thanks to the flexibility and adaptability of their functional form. They represent one of the most successful and multidisciplinary subjects [128]. The birth of the artificial neural network goes back to a mathematical model formulated in 1943 for reproducing some of the characteristics

Experiment	NMC	NMC-pd	SLAC	BCDMS
$\left\langle PE\left[\left\langle F^{(\mathrm{art})}\right\rangle_{\mathrm{rep}}\right]\right\rangle_{\mathrm{dat}}$	$9.0\cdot10^{-5}$	$1.8\cdot10^{-5}$	$3.1\cdot10^{-4}$	$1.3\cdot10^{-3}$
$r\left[F^{(\mathrm{art})}\right]$	1.000	1.000	1.000	1.000
$\left\langle PE\left[\left\langle \sigma^{(\mathrm{art})}\right\rangle_{\mathrm{rep}}\right]\right\rangle_{\mathrm{dat}}$	$1.5\cdot10^{-3}$	$4.2\cdot10^{-3}$	$3.1\cdot10^{-3}$	$4.0\cdot10^{-3}$
$\left\langle \sigma^{(\mathrm{exp})}\right\rangle_{\mathrm{dat}}$	0.0147	0.0170	0.0104	0.0698
$\left\langle \sigma^{(\mathrm{art})}\right\rangle_{\mathrm{dat}}$	0.0146	0.0171	0.0104	0.0692
$r\left[\sigma^{(\mathrm{art})}\right]$	1.000	0.998	0.998	0.999
$\left\langle \rho^{(\mathrm{exp})}\right\rangle_{\mathrm{dat}}$	0.033	0.165	0.312	0.470
$\left\langle \rho^{(\mathrm{art})}\right\rangle_{\mathrm{dat}}$	0.033	0.176	0.311	0.463
$r\left[\rho^{(\mathrm{art})}\right]$	0.963	0.988	0.987	0.994
$\left\langle \mathrm{cov}^{(\mathrm{exp})}\right\rangle_{\mathrm{dat}}$	$6.52\cdot10^{-6}$	$4.39\cdot10^{-5}$	$3.07\cdot10^{-5}$	$2.90\cdot10^{-5}$
$\left\langle \mathrm{cov}^{(\mathrm{art})}\right\rangle_{\mathrm{dat}}$	$6.78\cdot10^{-6}$	$4.73\cdot10^{-5}$	$3.03\cdot10^{-5}$	$2.82\cdot10^{-5}$
$r\left[\mathrm{cov}^{(\mathrm{art})}\right]$	0.989	0.984	0.988	0.999

Table 3.1: Statistical estimators for the Monte Carlo artificial data generation with $N_{\mathrm{rep}} = 1000$. PE stands for percentage error and r is the scattering correlation. The definition of these statistical estimators is given in Appendix B.

of the synaptic connections of the brain [129]. Nowadays applications of artificial neural networks are widely used in many different contexts. An artificial neuron is defined as a processing element whose state ξ at the time t can assume two different values: $\xi(t) = 1$, if it is firing, or $\xi(t) = 0$, if it is at rest. The state of the i-th unit, $\xi_i(t)$, depends on the inputs coming from the other $N-1$ neurons through the discrete dynamical equation

$$\xi_i(t) = g\left(\sum_{j=1}^{N}\omega_{ij}\xi_j(t-1) - \theta_i\right) , \tag{3.4}$$

where the weights ω_{ij} represent the strength of the synaptic coupling between the j-th and the i-th neurons, θ_i is the threshold which must be overtaken to activate the signal, and g is the activation function. The latter is typically bounded either in the interval $[0,1]$ or $[-1,1]$. If it is of the form of the Θ step function

$$g(h) = \Theta(h) \equiv \begin{pmatrix} 0 \text{ if } h \leq 0\,, \\ 1 \text{ if } h > 1 \end{pmatrix},$$

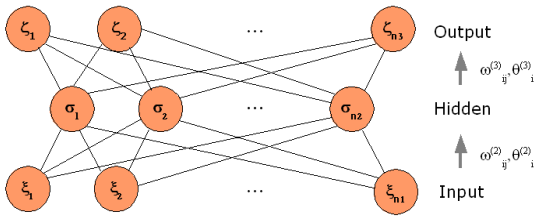


Figure 3.4: Schematic structure of a multi–layer feed-forward neural network.

the activation is said to be discrete. If the functions is for instance the sigmoid function

$$g(h) = \frac{1}{1+e^{-h}}\,, \tag{3.5}$$

which satisfies

$$\lim_{h\to\infty} g(h) = \Theta(h) \tag{3.6}$$

the activation function is said to be continuous. It has been shown that any continuous function can be uniformly approximated by a continuous neural network having only one infinitely large internal layer, and with an arbitrary continuous sigmoid nonlinearity [130].

The particular kind of neural network that we use in our analysis is the so-called multilayer feed-forward neural networks or perceptron, shown in Fig. 3.4. The first layer is the input one, where the input patterns are introduced into the rest of the network. In the NNPDF fits, the input parameters are x and $\log(x)$. The latter has been introduced for optimising the computation. Between the input layer and the output one, proportional to the initial scale PDF, there are one or more hidden layers of neurons evaluating the function g of the weighted sum of their inputs, which, in turn, are sent forward to the following layer and so on, until the output level is reached. For illustrating purposes, we consider a three–layer neural network. When an input vector ξ is introduced to the network, the states of the hidden neurons acquire the values

$$\sigma_i = g\left(\sum_{k=1}^{n_1} \omega_{ij}^{(2)}\xi_k - \theta_j^{(2)}\right)\,, j = 1,\ldots,n_2\,; \tag{3.7}$$

the output of the network is the vector ζ whose components are given by

$$\zeta_i = g\left(\sum_{j=1}^{n_1}\omega_{ij}^{(3)}\sigma_j - \theta_i^{(3)}\right), i = 1,\ldots,n_3\,. \tag{3.8}$$

Generally, if one has L layers with $n_1,\ldots,n_L$ units respectively, the state of the multilayer perceptron is established by recursive relations

$$\xi_i^{(l)} = g\left(\sum_{j=1}^{n_l-1}\omega_{ij}^{(l-1)}\xi_j^{(l-1)} - \theta_i^{(l)}\right), i = 1,\ldots,n_l\,, l = 2,\ldots,L\,, \tag{3.9}$$

where $\xi^{(l)}$ represents the state of the neurons in the l^{th} layer, $\omega_{ij}^{(l)}$ the weights between units in the $(l-1)^{th}$ and the l^{th} layers, and $\theta_i^{(l)}$ the threshold of the i^{th} unit in the l^{th} layer. Then the input is the vector $\xi^{(1)}$ and the output the vector $\xi^{(L)}$. The number of parameters for a given architecture, being l the number of layers and $n(l)$ the number of units in each layer, is

$$N_{\rm par} = \sum_{j=1}^{l-1} n(j+1)\left[1+n(j)\right]. \tag{3.10}$$

As an example, one can write explicitly the functional form for a 1–2–1 network, having one single input, one output and one intermediate layer with only two neurones, thus determined by 7 parameters, as

$$\xi_1^{(3)} = \frac{1}{1+e^{\theta_1^{(3)} - \frac{\omega_{11}^{(2)}}{1+e^{\theta_1^{(2)}-\xi_1^{(1)}\omega_{11}^{(1)}}} - \frac{\omega_{12}^{(2)}}{1+e^{\theta_2^{(2)}-\xi_1^{(1)}\omega_{21}^{(1)}}}}}.$$

The explicit functional form may be written for any other architecture, just yielding a more complicated expression.

In the NNPDF approach, each of the independent PDFs in the evolution basis introduced in Eq. (1.58) is parametrised using a multi-layer feed-forward neural network supplemented with a polynomial preprocessing. In Tab. 3.2, the independent combinations of PDFs and their parametrisations are shown. In the first parton analysis [68], only five out of thirteen partonic distribution were considered to be independent of each others, basically the light up and down quarks and the corresponding anti–quarks and the gluon. The other PDFs were determined by mean of some flavor assumptions: the strange sea was assumed to be proportional to the non–strange sea and to

	$f(x)$	Parametrisation	Architecture	$N_{\rm par}$
Singlet	$\Sigma(x)$	$\rm NN_\Sigma$	2-5-3-1	37
Gluon	$g(x)$	$\rm NN_g$	2-5-3-1	37
Total Valence	$V(x) \equiv \sum_i f_i^v(x)$	$\rm NN_V$	2-5-3-1	37
Triplet	$T_3(x) \equiv u^+(x) - d^+(x)$	$\rm NN_{T_3}$	2-5-3-1	37
Sea Asymmetry	$\Delta_S(x) \equiv \bar{d}(x) - \bar{u}(x)$	$\rm NN_\Delta$	2-5-3-1	37
Total strangeness	$s^+(x) \equiv (s(x) + \bar{s}(x))/2$	$\rm NN_{s+}$	2-5-3-1	37
Strange valence	$s^-(x) \equiv (s(x) - \bar{s}(x))/2$	$\rm NN_{s-}$	2-5-3-1	37

Table 3.2: Neural Network parametrisation of parton distributions in the NNPDF analyses. In Ref. [68] only the first five parton distributions were considered to be independent, the last two were added in Ref. [69] and in all subsequent analyses.

be symmetric. These assumptions were a source of bias in the determination of parton densities. Since the second analysis of Ref. [69], an independent parametrisation for the strange and anti–strange distribution was added. In all the analyses that we published until now, the zero mass variable flavor number (ZM-VFN) scheme to incorporate the effects of the heavy quarks. Details about treatment of the heavy quarks in the past and future analyses is discussed in the next section. The neural networks we use are chosen to have all the same architecture, namely 2–5–3–1. This corresponds to 37 free parameters for each PDF, i.e. a total of 185 free parameters in Ref. [68] and 259 free parameters in Ref. [69] and therein. This number is to be compared to less than a total of 30 free parameters for parton fits based on standard functional parametrisations [59, 63, 75]. The use of a redundant architecture reduces a priory the possibility of a functional bias. Lack of bias will be checked a posteriori. In the next section I show that results are independent of the choice of architecture.

Neural networks can accommodate any functional form, provided they are made of a large number of layers and sufficient time is used to train them. Nevertheless, it is customary to use preprocessing of data to subtract some dominant functional dependence. Then, smaller neural networks can be trained in a shorter time to deal with the deviations with respect to the dominant function. In the NNPDF fits, we use preprocessing to divide out some of the asymptotic small– and large–x behaviour of PDFs.

PDF	$[m_{\min}, m_{\max}]$	$[n_{\min}, n_{\max}]$	$r\,\vert\chi^2, m\vert$	$r\,\vert\chi^2, n\vert$
$\Sigma(x,Q_0^2)$	$[2.55, 3.45]$	$[1.05, 1.35]$	-0.018	0.131
$g(x,Q_0^2)$	$[3.55, 4.45]$	$[1.05, 1.35]$	-0.002	0.050
$T_3(x,Q_0^2)$	$[2.55, 3.45]$	$[0, 0.5]$	-0.023	-0.130
$V_T(x,Q_0^2)$	$[2.55, 3.45]$	$[0, 0.5]$	0.003	-0.068
$\Delta_S(x,Q_0^2)$	$[12, 14]$	$[-0.95, -0.65]$	0.000	-0.069
$s^+(x,Q_0^2)$	$[2.55, 3.45]$	$[1.05, 1.35]$	0.021	-0.055
$s^-(x,Q_0^2)$	$[2.55, 3.45]$	$[0, 0.5]$	-0.027	-0.015

Table 3.3: The range of random variation of the large-x and small-x preprocessing exponents m and n used in the NNPDF20 analysis [71]. The last two columns give the correlation coefficient defined in Eq. (3.13) between the χ^2 and respectively the large– and small–x preprocessing exponents.

The input PDF basis can be thus written in terms of neural networks as

$$\begin{aligned}
\Sigma(x,Q_0^2) &= (1-x)^{m_\Sigma} x^{-n_\Sigma} \mathrm{NN}_\Sigma(x)\,, \\
V(x,Q_0^2) &= A_V (1-x)^{m_V} x^{-n_V} \mathrm{NN}_V(x)\,, \\
T_3(x,Q_0^2) &= (1-x)^{m_{T_3}} x^{-n_{T_3}} \mathrm{NN}_{T_3}(x)\,, \\
\Delta_S(x,Q_0^2) &= A_{\Delta_S} (1-x)^{m_{\Delta_S}} x^{-n_{\Delta_S}} \mathrm{NN}_{\Delta_S}(x)\,, \\
g(x,Q_0^2) &= A_g (1-x)^{m_g} x^{-n_g} \mathrm{NN}_g(x) \\
s^+(x,Q_0^2) &= (1-x)^{m_{s^+}}\, x^{-n_{s^+}} \mathrm{NN}_{s^+}(x), \\
s^-(x,Q_0^2) &= (1-x)^{m_{s^-}}\, x^{-n_{s^-}} \mathrm{NN}_{s^-}(x) - s_{\mathrm{aux}}(x,Q_0^2),
\end{aligned} \tag{3.11}$$

where

$$s_{\mathrm{aux}}(x,Q_0^2) = A_{s^-} \left[x^{r_{s^-}}\, (1-x)^{t_{s^-}} \right]. \tag{3.12}$$

The relative normalisation of the neural networks is set by imposing the momentum sum rules as it is explained in Sect. 3.1.6.

In Ref. [68] the exponents of the preprocessing functions were kept fixed to a given value and the independence of the choice of the pre–processing was verified a posteriori. However, in Ref. [69] and henceforth we avoided possible bias related to this choice. We reached a greater stability by exploring a large space of preprocessing functions: a randomised range of variation of preprocessing exponents was introduced in the fit. The range of preprocessing exponents used in the most recent fit [71] is shown in Table 3.3.

A way for verifying the explicit independence of results on preprocessing exponents within the ranges defined in Table 3.3 was introduced in Ref. [70] by evaluating the correlation between the value of a given preprocessing exponent and the associated value of the χ^2 computed between the k–th net and experimental data. The correlation coefficient is defined as follows: considering for definiteness the large–x preprocessing exponent of the singlet PDF $\Sigma(x,Q^2)$, one has

$$r\left[\chi^2, m_\Sigma\right] \equiv \frac{\left\langle \chi^2 m_\Sigma \right\rangle_{\rm rep} - \left\langle \chi^2 \right\rangle_{\rm rep} \left\langle m_\Sigma \right\rangle_{\rm rep}}{\sigma^2_{m_\Sigma}} \ . \tag{3.13}$$

This provides the variation $\delta\chi^2$ as the large–x exponent δm_Σ is varied around its mean value. The correlations was found to be very weak as it is shown in the last two columns of Table 3.3. The $\chi^{2(k)}$ for the individual replicas is only marginally affected. This validates quantitatively the stability of the results with respect to the preprocessing exponents.

3.1.3 Figure of merit and t0 algorithm

The figure of merit for fitting of the neural networks on the individual replicas is the error function, defined in Ref. [67] as

$$E^{(k)} = \frac{1}{N_{\rm dat}} \sum_{i,j=1}^{N_{\rm dat}} \left(F_i^{({\rm art})(k)} - F_i^{({\rm net})(k)}\right) \left(\left(\widetilde{\rm cov}^{(k)}\right)^{-1}\right)_{ij} \left(F_j^{({\rm art})(k)} - F_j^{({\rm net})(k)}\right) \ , \tag{3.14}$$

where the value $F_i^{({\rm net})}$ of the observable corresponding to the $i-$th data point is computed from the PDFs by evolving PDFs to the scale of the measurements and convoluting them with the coefficient functions. The error function is similar to the χ^2 per degree of freedom defined in Eq. (2.4). However, the error function measures the quality of the fit of each set to its corresponding replica, while the χ^2 measures the quality of the fit of each set to the experimental data. Moreover the covariance matrix employed $\widetilde{\rm cov}^{(k)}$ is different from the one defined in Eq. (2.3).

Indeed, as it was discussed in Chap. 3, the treatment of normalisation uncertainties needs some care: including normalisation uncertainties in the covariance matrix would lead to a fit that is systematically biased to lie below the data [117]. The way this problem was handled up to the NNPDF1.2 fit [68, 70], consisted in including normalisation uncertainties by rescaling all uncertainties, i.e. by constructing for each replica

the modified covariance matrix $\widetilde{\mathrm{cov}}_{ij}^{(k)}$ appearing in Eq. (3.14) as

$$\widetilde{\mathrm{cov}}_{ij}^{(k)} = \overline{\mathrm{cov}}^{(k)}{}_{ij} \equiv \left(\sum_{l=1}^{N_c} \overline{\sigma}^{(k)}{}_{i,l} \overline{\sigma}^{(k)}{}_{j,l} + \delta_{ij} \left(\overline{\sigma}^{(k)}{}_{i,s} \right)^2 \right) F_i F_j \,, \tag{3.15}$$

with the statistical uncertainties $\sigma_{p,s}$ and each systematic uncertainty $\sigma_{p,l}$ being rescaled by a factor which depends on the considered replica (k), according to

$$\overline{\sigma}_{i,s}^{(k)} = S_{i,N}^{(k)} \sigma_{i,s} \,, \qquad \overline{\sigma}_{i,l}^{(k)} = S_{i,N}^{(k)} \sigma_{i,l} \,, \qquad l = 1, \ldots, N_c \,. \tag{3.16}$$

and thus

$$\overline{\mathrm{cov}}^{(k)}{}_{ij} = \overline{\mathrm{cov}}^{(\mathrm{exp})}{}_{ij} S_{i,N}^{(k)} S_{j,N}^{(k)} \,. \tag{3.17}$$

Therefore the experimental correlation matrix without normalisation uncertainties needs to be evaluated only once, given that it does not depend on the replica, while $\overline{\mathrm{cov}}^{(k)}{}_{ij}$ is obtained by multiplying by the normalisation factors $S_{i,N}^{(k)}$ and $S_{j,N}^{(k)}$ for each replica. If within an experiment all sets have only a common global normalisation uncertainty, the rescaling is an overall multiplicative factor.

However we figured out that this method is only accurate when all normalisation uncertainties have a similar size, as it is discussed in Chap. 3. In Ref. [71] an improved treatment of normalisation uncertainties was implemented. Following Ref. [118], the covariance matrix for each experiment is computed from the knowledge of statistical, systematic and normalisation uncertainties as follows:

$$\begin{aligned} \widetilde{\mathrm{cov}}_{ij} \quad = \quad & (\mathrm{cov}_{\mathrm{t}_0})_{ij} \equiv \left(\sum_{l=1}^{N_c} \sigma_{i,l} \sigma_{j,l} + \delta_{ij} \sigma_{i,s}^2 \right) F_i F_j \\ & + \left(\sum_{n=1}^{N_a} \sigma_{i,n} \sigma_{j,n} + \sum_{n=1}^{N_r} \sigma_{i,n} \sigma_{j,n} \right) F_i^{(0)} F_j^{(0)} \,, \end{aligned} \tag{3.18}$$

where now the matrix does not depend on the replica k, i and j run over the experimental points and $F_i^{(0)}$, $F_j^{(0)}$ are the corresponding observables iteratively determined from some previous fit. As it was shown in Chap. 3, the convergence of the iterative procedure is very fast and the final values of $F_i^{(0)}$ used in Eq. (3.18) do not differ significantly from the final fit results.

3.1.4 Genetic algorithm minimisation

The error function Eq. (3.14) can be minimised with a variety of techniques, including standard steepest–descent in the space of parameters. Due to the non–local nature of the error function and the complex structure of the parameter space, genetic algorithms [131] turn out to be the most efficient method. The main advantage of genetic algorithms is that they work on a population of solutions, rather than tracing the progress of one point through parameter space. Thus, many regions of parameter space are explored simultaneously, thereby lowering the possibility of getting trapped in local minima.

The state of the neural network is represented by the vector

$$\mathbf{nn} = \left(\mathrm{nn}_1, \mathrm{nn}_2, \ldots, \mathrm{nn}_{N_{\mathrm{par}}}\right) \ . \tag{3.19}$$

where each element nn_i corresponds either to a weight $\omega_{ij}^{(l)}$ or to a threshold θ_i, and N_{par} indicates the total number of parameters determining the output of the networks. The initial set of parameters is chosen at random. At each iteration of the minimisation, usually referred as generation, a set of N_{mut} copies of the vector **nn** is generated. The N_{mut} mutants are obtained by replacing one or more randomly chosen elements of the state vector $\mathbf{nn}$ by a new value according to the rule

$$\mathrm{nn}_k \to \mathrm{nn}_k + \eta\left(r - \frac{1}{2}\right) \ , \tag{3.20}$$

where r is a uniform random number between 0 and 1. This step is referred as mutation and η, the mutation rate, is a free parameter of the algorithm.

Once the set of different mutants is generated, the vector or the set of vectors with lowest values of the error Eq.(3.14) are selected out of the total population of N_{mut} individuals, and used to replace the original vector. This choice might be replaced by other methods based on probabilistic selection. However we have found no advantage in using probabilistic methods for the selection of the best mutant. The selection of the copy with the lowest figure of merit is best suited for our strategy. With this choice the genetic algorithm produces a monotonically decreasing profile of the figure of merit. The procedure is iterated until the vector with the smallest value of the error function meets a suitable convergence criterion, to be discussed in the next subsection. However, in order to avoid unacceptably long fits, when a very large number of iterations $N_{\mathrm{gen}}^{\mathrm{max}}$ is reached, training is stopped anyway. This leads to a small loss of accuracy of the corresponding fits which is acceptable provided it only happens for a small fraction of replicas.

Genetic algorithms are controlled by some parameters, like the mutation rate, that may be tuned in order to optimise the efficiency of the whole minimisation procedure. In order to avoid the local minima and increase the training speed, the most obvious improvement consists in introducing multiple mutations $\eta \to \eta_i$ with different probabilities ρ_i with $i = 1, ..., N_{\rm multmut}$. We verified that this produces a significant improvement of the convergence rate [67].

Further improvements were introduced in Refs. [68, 71]. The first of these was to allow, for each PDF $j = 1, \ldots, N_{\rm pdf}$, different values of the mutation rates $\eta_{i,j}$. This is motivated by the fact that each PDF functionality is different, and thus best approached using a specific learning rate. Furthermore, as training proceeds, all mutation rates are adjusted dynamically as a function of the number of iterations $N_{\rm ite}$

$$\eta_{i,j} = \eta_{i,j}^{(0)} / N_{\rm ite}^{r_\eta} \, . \tag{3.21}$$

In order to optimally span the range of all possible beneficial mutations the exponent r_η was randomised between 0 and 1 at each iteration of the genetic algorithm. An analysis of the values of r_η for which mutations are accepted in each generation reveals a flat profile: both large and small mutations are beneficial at all stages of the minimisation. We also tuned the number of mutants depending on the stage of the training. When the number of generations is smaller than $N_{\rm gen}^{\rm mut}$, we use a large population of mutants $N_{\rm mut}^a \gg 1$, while afterwards we use a much reduced population $N_{\rm mut}^b \ll N_{\rm mut}^a$. The reason for this procedure is that at early stages of the minimisation it is beneficial to explore as large a parameter space as possible, thus we need a large population. Once we are closer to a minimum, a reduced population helps in propagating the beneficial mutations to further improve the fitness of the best candidates.

Finally, in order to deal more efficiently with the needs of fitting data from a wide variety of different experiments and different data sets within an experiment, we adopt a weighted fitting technique, following an earlier study in Ref. [67]. The aim of the technique is to let the minimisation procedure converge rapidly towards a configuration for which the final χ^2 is even among all the experimental sets. Weighted fitting consists of adjusting the weights of the data sets in the determination of the error function during the minimisation procedure according to their individual figure of merit: data sets that yield a large contribution to the error function get a larger weight in the total figure of merit. The function which is minimised is thus

$$E_{\rm wt}^{(k)} = \frac{1}{N_{\rm dat}} \sum_{j=1}^{N_{\rm sets}} p_j^{(k)} N_{{\rm dat},j} E_j^{(k)} \, , \tag{3.22}$$

$\eta^{(0)}_{i,\Sigma}$	$\eta^{(0)}_{i,\mathrm{g}}$	$\eta^{(0)}_{i,T_3}$	$\eta^{(0)}_{i,\mathrm{V_T}}$	$\eta^{(0)}_{i,\Delta_\mathrm{S}}$	$\eta^{(0)}_{i,S^+}$	$\eta^{(0)}_{i,S^-}$
$[10,1]$	$[10,1]$	$[1,0.1]$	$[1,0.1]$	$[1,0.1]$	$[5,0.5]$	$[1,0.1]$

$N^{\mathrm{wt}}_{\mathrm{gen}}$	$N^{\mathrm{mut}}_{\mathrm{gen}}$	$N^{\mathrm{max}}_{\mathrm{gen}}$	E^{sw}	N^{a}_{mut}	N^{b}_{mut}	N_{update}
10000	2500	30000	2.6	80	10	10

Table 3.4: Parameter values for the genetic algorithm adopted in Ref. [71].

where $N_{\mathrm{dat},j}$ is the number of data points of the $j-$th set and $E_j^{(k)}$ the error function defined in Eq. (3.14) but restricted to the points of the $j-$th dataset. The weights $p_j^{(k)}$ are determined as

$$p_j^{(k)} = \left(\frac{E_j^{(k)}}{E_{\mathrm{max}}^{(k)}} \right)^2 \quad , \tag{3.23}$$

with $E_{\mathrm{max}}^{(k)}$ being the highest among the $E_j^{(k)}$ at the given GA generation. Their values are updated every N_{update} generations, with default $N_{\mathrm{update}} = 10$.

In Ref. [71] and therein, a slightly different way of determining the weights $p_j^{(k)}$ was adopted. In Eq. (3.23) $E_{\mathrm{max}}^{(k)}$ is replaced by E_j^{targ}. The idea is the following: in the beginning of the fit, target values E_j^{targ} for the figure of merit of each experiment are chosen. Then, at each generation of the minimisation, the weights of individual sets are updated using the conditions

1. If $E_j^{(k)} \geq E_j^{\mathrm{targ}}$, then $p_j^{(k)} = \left(E_j^{(k)} / E_j^{\mathrm{targ}} \right)^2$,
2. If $E_j^{(k)} < E_j^{\mathrm{targ}}$, then $p_j^{(k)} = 0$.

Hence, sets which are far above their target value will get a larger weight in the figure of merit. On the other hand, sets which are below their target are likely to be already learnt properly and thus are removed from the figure of merit which is being minimised. However, every N_{ite} iterations, the figure of merit associated to the sets removed from the minimisation is computed and, if the latter has deteriorated, these sets are put back into the global figure of merit. The determination of the target values E_j^{targ} for all the sets which enter into the fit is an iterative procedure that works as follows. We started with all $E_j^{\mathrm{targ}} = 1$ and proceeded to a first very long fit. Then, we use the outcome of the fit to produce a first nontrivial set of E_j^{targ} values. This procedure is iterated until convergence. In practice, convergence is very fast. This implementation of targeted weighted training is such that the error function of each

dataset tends smoothly to its "natural" value, that is, $p_j^{(k)} \to 1$ as the minimisation progresses. Those sets which are harder to fit are given more weight than the experiments that are learnt faster.

As an illustration of the procedure, I show in Fig. 3.5 the $p_j^{(k)}$ weight profiles as a function of the number of genetic algorithm generations for some sets of a given replica. Note how, at the early stages of the minimisation, sets which are harder to learn, such as BCDMSp or NMC-pd are given more weight than the rest, while at the end of the weighted training epoch all weights are either $p_j^{(k)} \sim 1$ or oscillate between 0 and 1, a sign that these sets have been properly learnt.

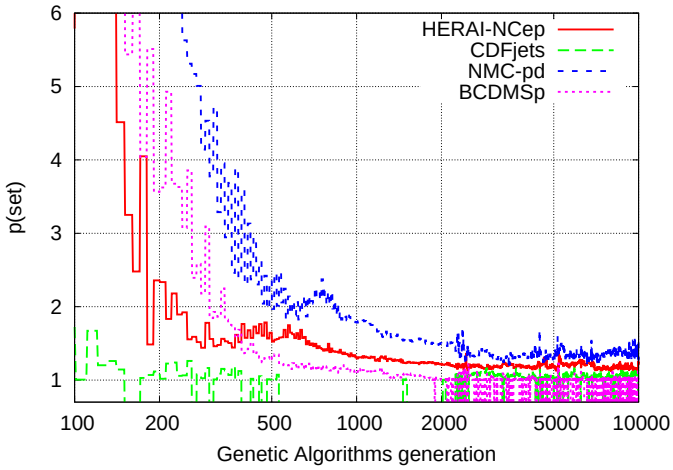


Figure 3.5: Illustration of the weighted training in one particular replica. Individual weights for each dataset converge to a value of p_i which is close to 1 as the training progresses. Only the behaviour of representative datasets is shown here.

An important feature of weighted training is that weights are given to individual datasets and not just to experiments. This is motivated by the fact that typically each dataset covers a distinct, restricted kinematic region. Hence, the weighting takes care of the fact that the data in different kinematic regions carry different amounts of information and thus require unequal amounts of training. This procedure poses the problem that different sets coming from the same experiment are correlated with each others and these correlations are neglected in the evaluation of Eq. (3.22). For this reason, the targeted weighted training epoch lasts for $N_{\rm gen}^{\rm wt}$ generations, unless the total error function Eq. (3.14) is above some threshold $E^{(k)} \geq E^{\rm sw}$. If it is, weighted training continues until $E^{(k)}$ falls below the threshold value. Afterwards, the error function is just the unweighted error function Eq. (3.14) computed on experiments.

It is in this final training epoch that a dynamical stopping of the minimisation is activated, as I discuss in the next section. Going through a final training epoch with the unweighted error function is in principle important in order to eliminate any possible residual bias from the choice of $E_i^{\rm targ}$ values in the previous epoch. However, in practice this safeguard has little effect, as it turns out that all weights tend to unity at the end of the targeted weighted training epoch as they ought to [71]. The whole procedure ensures that a uniform quality of the fit for all datasets is achieved, and that the fit is refined using the correct figure of merit which includes all the information on correlated systematics.

The final choices of the parameters of the genetic algorithm discussed in this section which have been adopted in the NNPDF2.0 parton determination are summarised in Table 3.4.

3.1.5 Determination of the optimal fit

The very large and redundant parametrisation of the initial scale PDFs requires a detailed analysis of the fitting strategy. On top of devising an algorithm to fit neural network PDFs, one has to deal with the fact that any redundant parametrisation may accommodate not only the smooth shape of the true underlying PDFs, but also fluctuations of the experimental data. The best fit then is not given by the absolute minimum of the likelihood. When dealing with quantities with some built–in smoothness, such as physical cross sections, this procedure does not produce the optimal fit. Indeed, even for fully compatible data, independent measurements of the same quantity at the same point will fluctuate within the uncertainty of the measurement. If fitted by maximum likelihood, such independent measurements will automatically be combined into their weighted average [132]. However, if two independent measurements are performed of the same observable, but measured at very close values of the underlying kinematic variables: for example the structure function $F(x,Q^2)$ at the same Q^2 and two different but close values x. Then a fit which goes through the central values of both measurements might be possible, but in the limit in which the two measurement are performed at infinitesimally close points this would correspond to a discontinuous behaviour of the observable, which is unphysical. This problem is exacerbated in the case of incompatible measurements. Instead, the best fit should be characterised by a value of the χ^2 which is equal to the value expected on the basis of the fluctuations of the data, beyond which the figure of merit improves only because one is fitting the statistical noise in the data.

In order to determine this value, a strategy was developed in Ref. [67], based on the cross-validation method used quite generally in neural network studies [133]. Namely, PDFs are trained on a fraction of the data and validated on the rest of the data. Training

is stopped when the quality of the fit to validation data (i.e. the data which are not used in the training) deteriorates while the quality of the fit to training data keeps improving. This corresponds to the overlearning regime where neural networks start to fit random fluctuations rather than the underlying physics. The method is made possible by the availability of a very large and mostly compatible set of data, and it guarantees that the best fit does not attempt to reproduce random fluctuations of the data. The method also handles incompatible data, by automatically tolerating fluctuations in the data even when they are larger than the nominal uncertainty, whenever fitting these fluctuations would not lead to an improvement of the global quality of the fit.

The application of the cross–validation method in the NNPDF fits has been described in details in Refs. [68, 67]. For each replica the data set is partitioned into training and validation subsets with fraction $f_{\rm tr}^{(j)}$ and $f_{\rm val}^{(j)} = 1 - f_{\rm tr}^{(j)}$ of the data points respectively. The values of the fractions can in general be different for each dataset j. The points in each set are chosen randomly out of the total dataset. This random partitioning of the data is different for each replica. The same value of the training fraction for all data sets is taken, $f_{\rm tr}^{j} = f_{\rm val}^{j} = \frac{1}{2}$, i.e. randomly for each replica, half of the points of each set belong to the training set and half to the validation set. The division is performed set by set in order to make sure that all data sets (and thus essentially all kinematic regions) are represented in the training and validation sets for each replica.

The fit is then performed on the data in the training set, and the figure of merit in both sets is monitored. Whereas usually, when using a genetic algorithm, the figure of merit cannot increase during the minimisation, with the weighted training algorithm discussed earlier the value of the figure of merit during minimisation oscillates due to the updating of the weights. In order to avoid spurious stopping induced by these oscillations, the monitored error functions are computed as moving averages over a given number of iterations, namely

$$\langle E_{\rm tr,val}(j)\rangle \equiv \frac{1}{N_{\rm sm}} \sum_{l=j-N_{\rm sm}+1}^{j} E_{\rm tr,val}(l)\,. \tag{3.24}$$

Minimisation is stopped when $\langle E_{\rm val}(j)\rangle$ in the validation set (not used for fitting) stops decreasing. This is illustrated in Fig. 3.6, where the moving averaged training and validation error functions are plotted as a function of the number of generations for one particular replica of the NNPDF2.0 reference fit. Overlearning is apparent as beyond the stopping point the training figure of merit keeps decreasing steadily while the validation flattens out and actually rises by a small amount. The stopping criteria

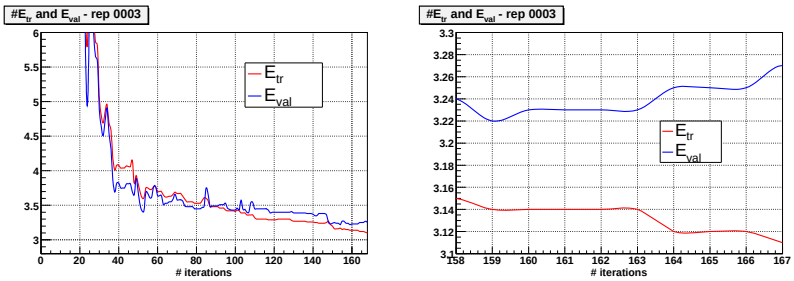


Figure 3.6: Training and validation error functions as a function of the number of iterations for one of the replicas in the reference fit. The stopping region is zoomed on the right–hand side plot.

are then satisfied if the averaged training error function is decreasing

$$r_{\rm tr} = \frac{\langle E_{\rm tr}(j)\rangle}{\langle E_{\rm tr}(j-\Delta_{\rm smear})\rangle} < 1-\delta_{\rm tr}\,, \tag{3.25}$$

while the averaged validation error function increases

$$r_{\rm val} = \frac{\langle E_{\rm val}(j)\rangle}{\langle E_{\rm val}(j-\Delta_{\rm smear})\rangle} > 1+\delta_{\rm val}\,. \tag{3.26}$$

The parameters $\delta_{\rm tr}$, $\delta_{\rm val}$ set the accuracy to which the increase and decrease is required in order to be significant. The values of the stopping parameters must be determined by analysing the behaviour of the fit for the particular dataset which is being used for neural network training. As an illustration of how this is done in practice, we show in Fig. 3.7 the averaged training and validation $E_{\rm tr/val}$ ratios Eqs. (3.25, 3.26) for a given replica and different values of the smearing length $N_{\rm smear}$. For this particular replica the training has been artificially prolonged beyond its stopping point. From Fig. 3.7 it is apparent that while the training ratio satisfies $r_{\rm tr} < 1$ always, i.e. that $E_{\rm tr}$ continues to decrease, after a given number of generations we have $r_{\rm val} > 1$, which then oscillates above and below 1: this is the sign that we have entered an 'overlearning' regime and minimisation needs to be stopped.

The optimal values of the stopping parameters are chosen to be small enough that overlearning is avoided, but large enough that the fit does not stop on statistical fluctuations. The latter condition can be met only if the value of $N_{\rm smear}$ is large enough, but if $N_{\rm smear}$ is too large stopping becomes very difficult and the first condition cannot be met. In practice, as explained in Ref. [71], a set of 100 replicas with very long training is generated, and for each value of $N_{\rm smear}$ a range of values of $\delta_{\rm tr}$ and $\delta_{\rm val}$ has been

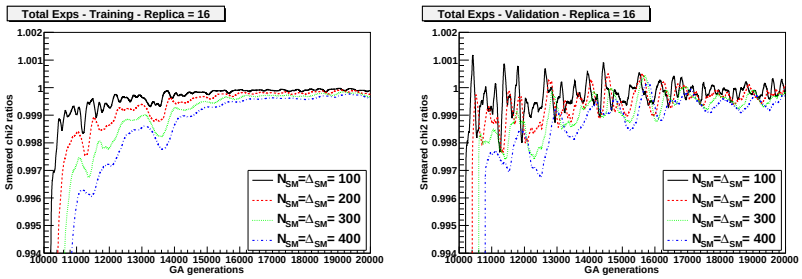


Figure 3.7: The training (left) and validation (right) ratios (defined in Eqs.(3.25, 3.26)) for a particular replica, as a function of the number of genetic algorithms generations, for various choices of the smearing parameter $N_{\rm smear} = \Delta_{\rm smear}$. The value $N_{\rm smear} = \Delta_{\rm smear} = 200$ is used in the reference fit (see Table 3.5).

$N_{\rm smear}$	$\Delta_{\rm smear}$	$\delta_{\rm tr}$	$\delta_{\rm val}$	$E_{\rm thres}$	$N_{\rm gen}^{\rm max}$
200	200	10^{-4}	$3\,10^{-4}$	6	30000

Table 3.5: Parameter values for the stopping criterion in the NNPDF2.0 analysis [71].

tried out until an optimal set of values which satisfies all the above criteria has been found. The final values of the parameters used in Ref. [71] are listed in Table 3.5.

In order to check the consistency of the whole procedure, in Ref. [71] a set of 100 replicas from a fit with the same settings as the final reference fit but with no stopping and a large maximum number of genetic algorithm generations $N_{\rm gen}^{\rm max} = 50000$ was produced. This set of 100 replicas allowed us to verify that the targeted weighted training and stopping criterion do not bias the fitting procedure and that the stopping criterion does not introduce underlearning by stopping the fit at a time when the quality of the fit is still improving. We verified that, while the average χ^2 for this fit is only marginally better than that of the reference fit, some experiments do show signs of overlearning, with an accordingly lower value of the contribution to the χ^2 . This is illustrated in Fig. 3.8, where the $E_i^{(k)}$ profiles for two particular experiments (NMC-pd and E605) and replicas taken from this fit without stopping is shown. In the first training epoch, in which the weighted training Eq. (3.22) is activated, one can see oscillations, but the downwards trend is clearly visible. Once targeted weighted training is switched off, minimisation proceeds smoothly, and we see in the two cases that after a given number of genetic algorithms generations we enter in overlearning. For the two experiments the typical overlearning behaviour, characterised by the fact that the validation $E_{\rm tr}^{(k)}$ is rising while the training $E_{\rm val}^{(k)}$ is still decreasing, sets in at about 15000 generations. This is the point where dynamical stopping avoids overlearning.

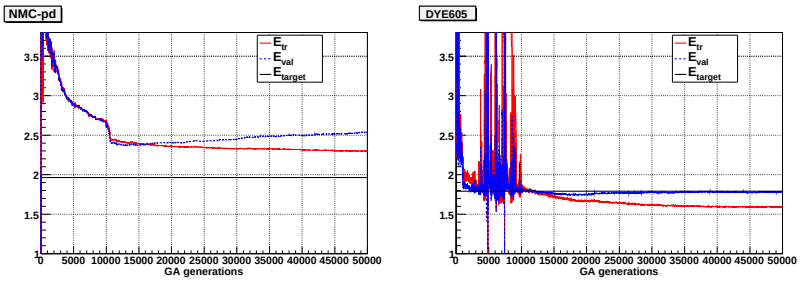


Figure 3.8: Two typical examples of overlearning behaviour, extracted from a fit with the same settings as the final reference fit of the NNPDF2.0 analysis [71] but with no stopping and a large maximum number of genetic algorithm generations $N_{\rm gen}^{\rm max} = 50000$. The right plot shows the overlearning of the E605 experiment observed in one particular replica, and the left plot corresponds to the NMC-pd experiment. Note that in these fits weighted training is switched off at $N_{\rm gen}^{\rm wt} = 10000$.

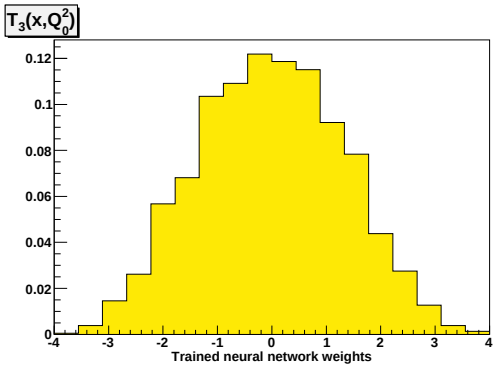


Figure 3.9: Distribution of the parameters of the neural network describing the T_3 parton densities at the initial scale $Q_0^2 = 2$ GeV^2 over the final ensemble of 1000 fitted initial parametrisations in the NNPDF2.0 analysis [71].

A final remark. The set of neural nets at stopping provides our best fit, but it is otherwise impossible to endow the best fit values of their parameters with a physical interpretation. In fact, since the nets are redundant, the values of most of these parameters is unconstrained or zero. This is visible in Fig. 3.9, where the distribution of neural network weights at stopping for 100 replicas of the triplet neural network $T_3(x, Q_0^2)$ is displayed. The well-balanced distribution of weights around zero shows that the individual neurons in the neural network operate in their natural range of sensitivity.

3.1.6 Positivity constraints and sum rules

In parton fits general theoretical constraints can be imposed during the minimisation procedure, thereby guaranteeing that the fitting procedure only explores the subspace of acceptable physical solutions.

Valence and momentum sum rules are enforced in the NNPDF approach [68] by constraining the relative normalisation of PDFs. For four of the basis PDF parametrisations defined in Eq. (3.11), namely g, V, Δ_S and s_+, an overall normalisation constant is factored out. The value of this constant is determined by requiring that the valence and momentum sum rules be satisfied. The valence sum rules

$$\int_0^1 dx\, u^-(x) = 2 \qquad \int_0^1 dx\, d^-(x) = 1 \tag{3.27}$$

fix the value of the total valence and sea asymmetry normalisations to be

$$\begin{aligned} A_V &= \frac{3}{\int_0^1 dx\ [(1-x)^{m_V}\mathrm{NN}_V(x)/x^{n_V}]}\,, \\ A_{\Delta_S} &= \frac{1-\int_0^1 dx\ [(1-x)^{m_{T_3}}\mathrm{NN}_{T_3}(x)/x^{n_{T_3}}]}{2\int_0^1 dx\ [(1-x)^{m_{\Delta_S}}\mathrm{NN}_{\Delta_S}(x)/x^{n_{\Delta_S}}]}\,, \end{aligned} \tag{3.28}$$

while the momentum sum rule

$$\int_0^1 dx\, x\,(\Sigma(x)+g(x)) = 1, \tag{3.29}$$

constrains the normalisation of the gluon density

$$A_g = \frac{1-\int_0^1 dx\, x\,[(1-x)^{m_\Sigma}\mathrm{NN}_\Sigma(x)/x^{n_\Sigma}]}{\int_0^1 dx\, x\,[(1-x)^{m_g}\mathrm{NN}_g(x)/x^{n_g}]}\,. \tag{3.30}$$

In the same way, the contribution $s_{\rm aux}(x,Q_0^2)$ is introduced in order to enforce the strange valence sum rule; the constant A_{s^-} is fixed by requiring

$$\int_0^1 dx\, s^-(x) = 0, \tag{3.31}$$

which gives the condition

$$A_{s^-} = \frac{\Gamma\left(r_{s^-}+t_{s^-}+2\right)}{\Gamma\left(r_{s^-}+1\right)\Gamma\left(t_{s^-}+1\right)}\int_0^1 dx\ (1-x)^{m_{s^-}}\ x^{-n_{s^-}}\,\mathrm{NN}_{s^-}(x). \tag{3.32}$$

The sum rules requires s^- to change sign at least once. This way of implementing the sum rule is designed in order to ensure that this crossing happens naturally in the valence region, rather than in some contrived way outside the data region where the shape of s^- is completely unconstrained. To this purpose, the exponents r_{s^-}, t_{s^-} are chosen in such a way that $s_{\rm aux}(x,Q_0^2)$ peaks in the valence region, and that the small–x and large–x behaviour of $s^-(x,Q_0^2)$ are not controlled by the $s_{\rm aux}(x,Q_0^2)$ contribution. In practice the latter condition is enforced by requiring $r_{s^-} \geq -n_{s^-}$ and $t_{s^-} \geq m_{s^-}$, while the former is enforced by letting $r_{s^-} = t_{s^-}/k$, which sets the maximum of $s_{\rm aux}(x,Q_0^2)$ at $x = \frac{1}{k+1}$. We then choose $t_{s^-} = 3.5$, and take k as a uniformly distributed random number in the range $k \in [1,3]$. The consequences of this very flexible implementation of the strangeness valence sum rule will be discussed in Chap. 6.

Besides direct experimental information and momentum sum rules, a further constraint on the input PDFs comes from the requirement of positivity. Indeed, even though PDFs are not positive-definite beyond LO, cross sections must remain positive, and this constrains the set of admissible PDFs [134]. The implementation of positivity constraints is nontrivial, because in principle one should require positivity of all observables, regardless of the fact that they are measurable in a realistic experiment. In practice, in Ref. [71] we imposed positivity of $F_L(x,Q^2)$, which constrains the gluon and the singlet PDFs at small–x, as well as that of the dimuon cross section $d^2\sigma^{\nu,c}/dxdy$ [70], which constrains the strange PDFs. Positivity of $F_L(x,Q^2)$ is implemented in the range $10^{-9} \leq x \leq 0.005$ and that of the dimuon cross section in $10^{-9} \leq x \leq 0.5$, in both cases at the initial evolution scale $Q^2 = 2$ GeV^2. This is done because, if positivity is enforced at low scales, it will be preserved by DGLAP evolution.

In the NNPDF approach, positivity constraints on relevant physical observables are imposed during the genetic algorithm minimisation using a Lagrange multiplier, which strongly penalises those PDF configurations which lead to negative observables. Therefore, due to positivity constraints, the minimised error function Eq. (3.14) (or Eq. (3.22) in the weighted training epoch) is modified as follows

$$E^{(k)} \to E^{(k)} - \lambda_{\rm pos} \sum_{i=1}^{N_{\rm dat,pos}} \Theta\left(-F_i^{({\rm net})(k)}\right) F_i^{({\rm net})(k)}\,, \tag{3.33}$$

where $N_{\rm dat,pos}$ is the number of pseudodata points used to implement the positivity constraints and we choose $\lambda_{\rm pos} \sim 10^{10}$ as its associate Lagrange multiplier. The impact of the positivity constraints is going to be quantified in Sect. 4.2.

3.1.7 Distances

An important feature of the NNPDF approach is that many issues of parton determination can be addressed using standard statistical tools. For example, the stability of results upon a change of parametrisation or upon the preprocessing range, as well as the difference introduced by a change in the fitting strategy, or by the addition of new data, can be verified by computing the distance between results in units of their standard deviation.

Given two sets of $N_{\rm rep}^{(1)}$ and $N_{\rm rep}^{(2)}$ replicas, one is often interested in knowing whether they correspond to different instances of the same underlying probability distribution, or whether instead they come from different underlying distributions. For finite $N_{\rm rep}^{(i)}$ this question can only be answered in a statistical sense. To this purpose, one may define the square distance between two estimators based on the given samples as the square difference between the estimators divided by its expectation value, i.e. divided by the corresponding standard deviation. By construction, the expectation value of the distance is one. Note that asking whether two PDF determinations come from the same underlying distribution is much more restrictive than asking whether they are consistent within uncertainties. For instance in the case of a pair of PDF determinations, such that the data–set on which one of the two is based is a subset of the dataset of the other, and such that all data are consistent with each other, these two determinations will clearly not come from the same underlying distribution, because the distribution of PDFs obtained from the wider data–set will have smaller uncertainty. However, if the data are consistent they will remain nevertheless consistent within uncertainties.

Given a set of $N_{\rm rep}^{(k)}$ replicas $q_i^{(k)}$ of some quantity q, the estimator for the expected (true) value of q is the mean

$$\langle q^{(k)} \rangle_{(i)} = \frac{1}{N_{\rm rep}^{(i)}} \sum_{i=1}^{N_{\rm rep}^{(i)}} q_i^{(k)}. \tag{3.34}$$

The square distance between the two estimates of the expected value obtained from sets $q_i^{(1)}$, $q_i^{(2)}$ is then

$$d^2 \left(\langle q^{(1)} \rangle, \langle q^{(2)} \rangle \right) = \frac{\left(\langle q^{(1)} \rangle_{(1)} - \langle q^{(2)} \rangle_{(2)} \right)^2}{\sigma^2_{(1)}[\langle q^{(1)} \rangle] + \sigma^2_{(2)}[\langle q^{(2)} \rangle]} \tag{3.35}$$

where the variance of the mean is given by

$$\sigma^2_{(i)}[\langle q^{(i)} \rangle] = \frac{1}{N_{\rm rep}^{(i)}} \sigma^2_{(i)}[q^{(i)}] \tag{3.36}$$

in terms of the variance $\sigma^2_{(i)}[q^{(i)}]$ of the variables $q^{(i)}$ (which a priori could come from two distinct probability distributions). The expressions that we use for estimating the variance of the mean and the distance between variances are provided in Appendix B.

By construction, the probability distribution for the distance coincides with the χ^2 distribution with one degree of freedom, and thus it has mean $\langle d\rangle = 1$, and $d \lesssim 2.3$ at 90% confidence level. However the accuracy in the determination of the expectation value scales as $1/\sqrt{N_{\rm rep}}$, so if the underlying probability distributions are different the distance will grow as $\sqrt{N_{\rm rep}}$ in the large $N_{\rm rep}$ limit. In this limit (in which the central values of the underlying distribution are accurately estimated by mean over the replica sample) the distance between central values is given by the distance rescaled by $\sqrt{N_{\rm rep}}$: if $N^{(1)}_{\rm rep} = N^{(2)}_{\rm rep} = N_{\rm rep}$, then

$$\delta(\sigma^2_{(1)}, \sigma^2_{(2)}) \equiv \frac{1}{\sqrt{N_{\rm rep}}} d(\sigma^2_{(1)}, \sigma^2_{(2)}) \tag{3.37}$$

provides (in the large $N_{\rm rep}$) limit, the difference between central values in units of the standard deviation. It follows that because of the halving of the size of the sample required for averaging as discussed above, for all distances shown in the next section, and computed with $N_{\rm rep} = 100$ replicas, one sigma corresponds to $d = \sqrt{50} \approx 7$.

3.2 Results

The viability of the method discussed in this chapter was originally demonstrated in the determination of the structure function $F_2(x, Q^2)$ of the proton and neutron from its direct measurement [135, 136]. It was then applied to problems of increasing difficulty. In Ref. [67] it was used to provide the determination of a single parton distribution (the $T_3(x)$ distribution), thereby addressing the issue of determining a quantity which is not measured directly, but rather related through theory to an experimental observable. Eventually a full analysis of the DIS data was published [68], followed by the study of the strange content of the proton [69, 70].

However, the requirements of precision physics are such that it is mandatory to exploit all the available information in PDFs determinations. 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. For this reason in the recent NNPDF2.0 global fit [71], on top of all the data used in the previous analyses, Drell–Yan and inclusive jet data have been included. Furthermore the separate ZEUS and H1 datasets have been replaced with the HERA-I combined dataset [84]. The dataset used in this parton determination is thus comparable in variety and size to that used by the current state–of–the–art PDFs determination, namely

CTEQ6.6 [42] and MSTW08 [43]. All NNPDF parton sets are available through the LHAPDF interface [137].

In this section, rather than reporting all results of Refs. [68, 69, 70, 71], I select some results and collect them in thematic subsections. The aim is to show that PDFs determined using the NNPDF methodology enjoy several desirable features: the Monte Carlo behaves in a statistically consistent way, given that uncertainties scale as expected with the size of the sample; results are demonstrably independent of the parton parametrisation; PDFs behave as expected upon the addition of new data, i. e. uncertainties expand when data are removed and shrink when they are added unless the new data is incompatible with the old; results are even stable upon the addition of new independent PDF parametrisations. To conclude I discuss the role of the theoretical uncertainty and give an overview of the next sets which we are going to be made available, namely a refined analysis where the effect of the mass of the heavy quarks is included in a general mass scheme and the NNLO analysis.

3.2.1 Experimental data and physical observables

In this section I present the datasets included in the NNPDF analyses [68, 70, 71] and the corresponding observables. The kinematical region covered by these data in shown in Fig. 3.10. The covariance matrix of each experiment included in the fit is computed from knowledge of statistical, systematic and normalisation uncertainties provided by the experimental collaboration. Whenever the correlated systematics are not provided, the statistical and systematic errors are summed in quadrature and the covariance matrix is diagonal.

The first set of PDFs determined with the NNPDF approach [68] was based on a comprehensive set of experimental data from deep-inelastic scattering with various lepton beams and nucleon targets. To keep higher–twist corrections under control, only data with $Q^2 > 2\,\mathrm{GeV}^2$ and $W^2 > 12.5\,\mathrm{GeV}^2$ are retained. The DIS data shown in Fig. 3.10 are those actually used in the analysis. Since the kinematic cuts we use are not too conservative, we supplemented our fits with target mass corrections, as it is discussed in Chapter 5.

These data include the proton and deuteron structure functions, defined in Eq. (1.21), determined in fixed-target experiments by the BCDMS [77, 78] and NMC [80, 79] collaborations. They provide information on the valence region of parton distributions and help in disentangling isospin triplet and isospin singlet contribution. They are supplemented with data on the structure functions from SLAC [76] which, though rather older and less precise, improve the kinematic coverage in the large–x region. When available, the ratio F_2^d/F_2^p is included, which benefits from cancellations in the

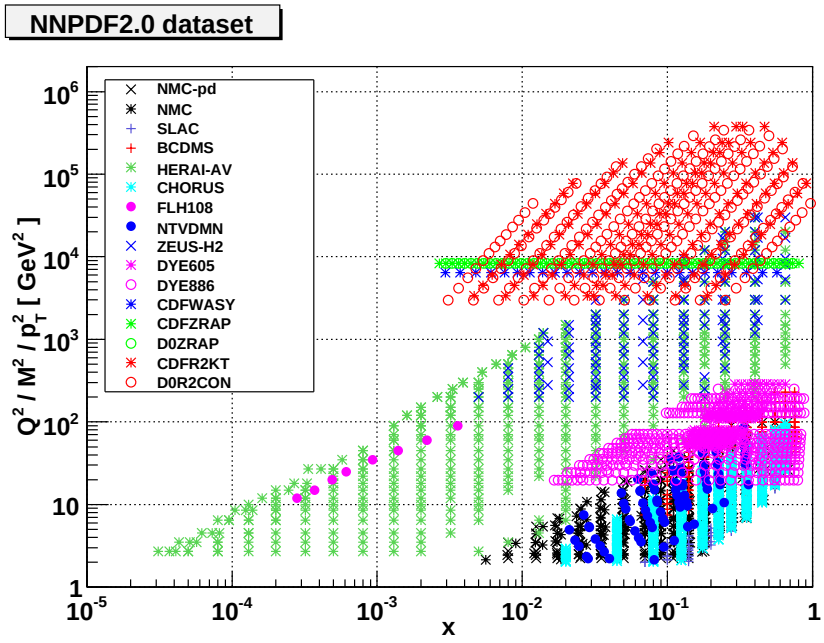


Figure 3.10: Experimental data which enter the NNPDF2.0 analysis. For hadronic data, the values of x_1 and x_2 determined by leading order partonic kinematics (Eqs. (3.40), (3.41) and (3.48)) are plotted.

correlated systematic uncertainties. Altogether these data cover the middle- to large-x and smaller Q^2 region of the kinematical range, corresponding to the lower-right corner in Fig. 3.10.

Collider experiments have explored a larger kinematical range in great detail. Neutral and charged current reduced cross sections, defined in Eq. (1.15), from the H1 [121, 124, 125, 123] and ZEUS [120, 138, 139, 140, 141, 142] collaborations were used in the first analyses [68, 69, 70]. Both neutral and charged current scattering data from charged lepton beams and neutrino scattering data enable us to disentangle the quark and antiquark distributions. In the most recent NNPDF2.0 analysis [71] these data were replaced by the combined HERA–I data [84]. They have a better accuracy than the one expected on purely statistical grounds from the combination of previous H1 and ZEUS data, because of the reduction of systematics from the cross-calibration of the two experiments. These data are given with 110 correlated systematics and three correlated procedural uncertainties, which we fully include in the covariance matrix. These HERA data sets yield information in a much wider region of the (x, Q^2) plane, in both the small-x and the large-Q^2 directions. The data for F_L that have recently

appeared in Ref. [83] and the measurements of neutral current and charged current deep-inelastic cross sections by the ZEUS experiment based on HERA-II data [85, 143] are also included. The F_L measurement is a rather small data set, but it provides an important direct measurement of F_L.

In order to control the valence–sea (or quark–antiquark) separation, we also included neutrino DIS data. Specifically, we use the large, up-to-date, and consistent set of neutrino and antineutrino scattering data by the CHORUS collaboration [88]. These data have a similar kinematic coverage to the fixed target charged lepton DIS data[2]. In the NNPDF1.2 fit [70] the analysis supplemented by data on deep-inelastic neutrino production of charm from NuTeV [89, 144] (dimuon data, henceforth) which give us a handle on the strange distribution, whose determination was the main goal of the paper [70]. NuTeV dimuon data overlap with the rest of fixed target experiments, providing information of the proton strangeness for $x \gtrsim 10^{-2}$. The charm production cross section is obtained from the published NuTeV neutrino dimuon production cross sections [144] as

$$\frac{1}{E_\nu}\frac{d^2\sigma^{\nu(\bar{\nu}),c}}{dx\,dy}(x,y,Q^2) = \frac{1}{\langle \mathrm{Br}\,(D\to\mu)\rangle \cdot \mathcal{A}\,(x,y,E_\nu)}\frac{1}{E_\nu}\frac{d^2\sigma^{\nu(\bar{\nu}),2\mu}}{dx\,dy}(x,y,Q^2), \quad (3.38)$$

where $\langle \mathrm{Br}\,(D\to\mu)\rangle$ is the average branching ratio of charmed hadrons into muons and $\mathcal{A}\,(x,y,E_\nu)$ is a bin-dependent experimental acceptance correction. More details are provided in Sect. 6.1, where the study of the strange content of the proton is discussed.

In the NNPDF2.0 PDF determination three classes of hadronic processes were included into the fit: Drell–Yan production in fixed target experiments, collider weak vector boson production, and collider inclusive jet production.

The fixed–target Drell-Yan data included in the NNPDF2.0 fit are E605 and E866. The former provides the absolute cross section for DY production from a proton beam on a copper target [90]. The double differential distribution in the rapidity y and the invariant mass of the Drell-Yan lepton pair, M^2, is given. E886, also known as NuSea, is based on the experimental set-up of E605. The absolute cross section measurements on a proton target is described in Refs. [92, 93], while the cross section ratio between deuteron and proton targets can be found in Ref. [94]. Double differential distributions in Feynman x_F and M are provided. For both experiments systematic and statistical errors are added in quadrature, since the correlation matrix is not provided by experimentalists. Fixed target Drell-Yan data from the E772 experiment [91] and from the deuteron data of E866 [92, 93] are not included. They have been shown to have poor compatibility with other Drell-Yan measurements [75] and thus do not add additional

[2]The CHORUS data, as well as the NuTeV, the E605 and the E886 data, refer to a nuclear target rather than a proton target, therefore they might be affected by the effect of nuclear corrections. However we do not include them in our default fit, since the study of their effect in the analysis that we performed in Ref. [70] reveals that their effect is smaller than the PDF uncertainty, about 1-2%

information to the global PDF analysis. The double–differential distribution in M and either the rapidity of the pair y or Feynman x_F is defined in terms of the hadronic kinematics as

$$y \equiv \frac{1}{2}\ln\frac{q_0+q_z}{q_0-q_z}; \quad x_F \equiv \frac{2q_z}{\sqrt{s}}, \tag{3.39}$$

where $\sqrt{s}$ is the hadron–hadron centre-of-mass energy, q is the four-vector of the Drell-Yan pair and q_z is its projection on the longitudinal axis. At leading order, the parton kinematics is entirely fixed in terms of hadronic variables by

$$x_1^0 = \sqrt{\tau}e^y = \frac{M}{\sqrt{s}}e^y\ , \qquad x_2^0 = \sqrt{\tau}e^{-y} = \frac{M}{\sqrt{s}}e^{-y}\ , \tag{3.40}$$

or equivalently

$$x_1^0 = \frac{1}{2}\left(x_F + \sqrt{x_F^2+4\tau}\right), \qquad x_2^0 = \frac{1}{2}\left(-x_F + \sqrt{x_F^2+4\tau}\right)\ . \tag{3.41}$$

The corresponding inverse relations are

$$\tau = x_1^0 x_2^0; \quad M^2 = s x_1^0 x_2^0 \tag{3.42}$$

and

$$y = \frac{1}{2}\ln\frac{x_1^0}{x_2^0}; \qquad x_F \equiv x_1^0 - x_2^0. \tag{3.43}$$

At leading order, the y or x_F Drell-Yan differential distribution is given by

$$\frac{d\sigma}{dM^2dy}(y) = \frac{4\pi\alpha^2}{9M^2s}\sum_i e_i^2\left[q_i(x_1)\bar{q}_i(x_2)+\bar{q}_i(x_1)q_i(x_2)\right]\ , \tag{3.44}$$

$$\frac{d\sigma}{dM^2dx_F}(y) = \frac{1}{x_1^0+x_2^0}\frac{d\sigma}{dM^2dy}(y), \tag{3.45}$$

where the dependence on M^2 is made implicit, α is the fine–structure constant and e_i the quark electric charges.

The weak boson production data included in the NNPDF2.0 fit are the D0 and CDF Z rapidity distribution and the CDF W boson asymmetry. The D0 Z rapidity distribution measurement was performed at Tevatron Run II and is described in Ref. [96]. It gives the Z/γ^* rapidity distribution in the range $71 \leq M_{ee} \leq 111$ GeV[3]. The

[3]The contribution from the Z^0/γ^* interference terms is well below the experimental uncertainties and it is neglected

CDF Z rapidity distribution is analogous to its D0 counterpart, and it is described in Ref. [97]. At leading order, the parton kinematics is given in Eqs. (3.39, 3.43), and the differential distribution is given by

$$\frac{d\sigma}{dy} = \frac{\pi G_F M_V^2 \sqrt{2}}{3s} \sum_{i,j} c_{ij} \left[q_i(x_1)\bar{q}_j(x_2) + \bar{q}_i(x_1) q_j(x_2) \right], \tag{3.46}$$

where the PDFs are evaluated at M_V, which denotes either M_W or M_Z; the electroweak couplings are defined in Tab. 1.1.

The CDF W boson asymmetry measurement, also performed at Tevatron Run II, is described in Ref. [95]. The physical observable is the rapidity asymmetry

$$A\left(y_W\right) \equiv \frac{d\sigma^{W^+}/dy_W - d\sigma^{W^-}/dy_W}{d\sigma^{W^+}/dy_W + d\sigma^{W^-}/dy_W} \,. \tag{3.47}$$

Because of the lack of a fast analytic implementation, we do not include lepton–level data, such as the Tevatron W asymmetries described in Refs. [145, 146, 147], which have been included in recent parton fits [43, 148] using K–factors. However a study on the inclusion of these data through reweighting is presented in Chap. 6.

Finally also the inclusive jet production cross section as a function of the transverse momentum p_T of the jet for fixed rapidity bins $\Delta\eta$ are included. The leading–order parton kinematics is fixed by

$$x_1^0 = \frac{p_T}{\sqrt{s}} e^{\eta} \,, \quad x_2^0 = \frac{p_T}{\sqrt{s}} e^{-\eta}, \tag{3.48}$$

while a simple leading–order expression for the cross–section is not available because of the need to provide a jet algorithm. Both the CDF Run II — k_T algorithm data and D0 Run II — midpoint algorithm data are included. The former are obtained using the k_T algorithm with $R = 0.7$. The dataset and the various sources of systematic uncertainties have been described in Ref. [99]. We choose to use the k_T algorithm measurements rather than the cone algorithm measurements [149], since the latter are not infrared safe. Data at $R = 0.7$ are preferable to available measurements at $R = 0.5$ or $R = 1$ since at Tevatron energies $R = 0.7$ optimises the interplay between sensitivity to perturbative radiation and impact of non-perturbative effects like Underlying Event [150, 151].

The data is provided in bins of rapidity $\Delta\eta$ and transverse momentum p_T. The kinematical coverage can be seen in Fig. 3.10. The D0 data is obtained using the MidPoint algorithm with $R = 0.7$. The dataset and the various sources of systematic uncertainties have been described in Ref. [100]. While the MidPoint algorithm is infrared unsafe, the effects of such unsafety in inclusive distributions are smaller than typical

$\chi^2_{\rm tot}$	1.21
$\langle E\rangle \pm \sigma_E$	2.32 ± 0.10
$\langle E_{\rm tr}\rangle \pm \sigma_{E_{\rm tr}}$	2.29 ± 0.11
$\langle E_{\rm val}\rangle \pm \sigma_{E_{\rm val}}$	2.35 ± 0.12
$\langle {\rm TL}\rangle \pm \sigma_{\rm TL}$	16175 ± 6257
$\left\langle \chi^{2(k)}\right\rangle \pm \sigma_{\chi^2}$	1.29 ± 0.09
$\left\langle \sigma^{(\rm exp)}\right\rangle_{\rm dat}$(%)	11.4
$\left\langle \sigma^{(\rm net)}\right\rangle_{\rm dat}$(%)	6.0
$\left\langle \rho^{(\rm exp)}\right\rangle_{\rm dat}$	0.18
$\left\langle \rho^{(\rm net)}\right\rangle_{\rm dat}$	0.54

Table 3.6: Table of statistical estimators for NNPDF2.0 reference fit [71] with $N_{\rm rep} = 1000$ replicas. The total average uncertainty is given in percentage. All statistical estimators are defined in Appendix B.

uncertainties [152] and thus it is safe to include this dataset into the analysis. The data is provided in bins of rapidity $\Delta\eta$ and transverse momentum p_T.

3.2.2 Statistical features and data compatibility

The set of fitted parton distribution functions at the initial scale provides an ensemble of parton distributions from which we can study the quality of the fit and the compatibility between different data–sets. As an example, I summarise the statistical features of the NNPDF2.0 analysis [71] in Table 3.6. The $\chi^2_{\rm tot}$ estimates the quality of the fit and refers to the best fit PDF set, given by the average over the $N_{\rm rep}$ replicas. The quality of the fit has improved in comparison to NNPDF1.2 [70] despite the widening of the data–set to also include hadronic data. This is mainly due to the improvement in the minimisation and stopping algorithm previously discussed. The value $\chi^2_{\rm tot} = 1.21$ has a small Gaussian probability and it is quite unlikely as a statistical fluctuation. It suggests experimental uncertainties might be underestimated at the 10% level, or that there might be theoretical uncertainties of the same order. This appears consistent with the expected accuracy of a NLO treatment of QCD, and the typical accuracy with which experimental uncertainties are estimated.

The distribution of the $\chi^{2(k)}$ and of the error function $E^{(k)}$ computed for each replica are displayed in Fig. 3.11. Note that $E^{(k)}$ and $\chi^{2(k)}$ differ because in the former each PDF replica is compared to the data replica it is fitted to, while in the latter it is compared to the actual data, and also because of the different treatment of normalisation uncertainties. The means of the χ^2 and the E distributions differ by about one unit, consistently with the expectation that the best fit correctly reproduces the underlying

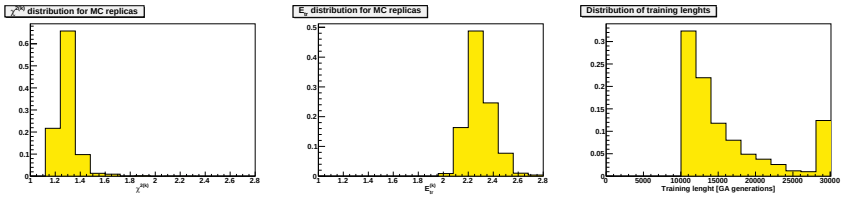


Figure 3.11: Distribution of $\chi^{2(k)}$ (left), $E_{\rm tr}^{(k)}$ of the training set (centre) and of the training lengths (right) over the sample of $N_{rep} = 1000$ replicas.

true behaviour about which data fluctuate, with replicas further fluctuating about data. The average training length $\langle TL \rangle$ (expressed as a number of generations of the genetic algorithm) given in Table 3.6 and its distribution displayed in Fig. 3.11 show that, while most of the replicas fulfil the stopping criterion, a small fraction ($\sim 12\%$) of them stop at the maximum training length $N_{\rm gen}^{\rm max}$, thereby causing some loss of accuracy in outlying fits. We have checked that, as $N_{\rm gen}^{\rm max}$ is raised, more and more of these replicas would stop, and that the loss of accuracy due to our choice of value of $N_{\rm gen}^{\rm max}$ is very small. Finally, as in the previous NNPDF determinations, the uncertainty of the fit, as measured by the average standard deviation $\langle \sigma \rangle$ is rather smaller than that of the data: 6.0% vs. 11.4%. The uncertainty reduction shows that the PDF determination is combining the information contained in the data into a determination of an underlying physical law.

To study the consistency between data sets, one has to look at the statistical estimators shown in Tab. 3.7, and at the histogram of χ^2 values for each experimental data–set shown in Fig. 3.12, where the unweighted average $\left\langle \chi^2 \right\rangle_{\rm sets} \equiv \frac{1}{N_{\rm set}} \sum_{j=1}^{N_{\rm set}} \chi^2_{\rm set,j}$ and standard deviation over datasets are also shown. From the histogram, no evidence of any specific dataset being clearly inconsistent with the other is seen, and the distribution of values looks broadly consistent with statistical expectations, with about five datasets with χ^2 at more than one but less than two sigma from the average. Also, we see no obvious difference or tension between hadronic and DIS datasets. Clearly, the χ^2 values for some experiments if taken at face value have low Gaussian probabilities (though only one, namely NMC, has a probability less than 0.01%). However, they appear to be stable upon the inclusion of new data, thus suggesting a lack of tension between different datasets. For instance, the χ^2 value of the NMC data is very close to that of Refs. [68, 70]: this value thus appears to reflect the internal consistency of these data, not their consistency with other data.

The χ^2 of the HERA-I combined data is $\chi^2 = 1.14$, somewhat larger that the value found when fitting the separate ZEUS and H1 data. The value comes from averaging the relatively large $\chi^2 \sim 1.3$ for the very precise NC positron dataset, with a low value

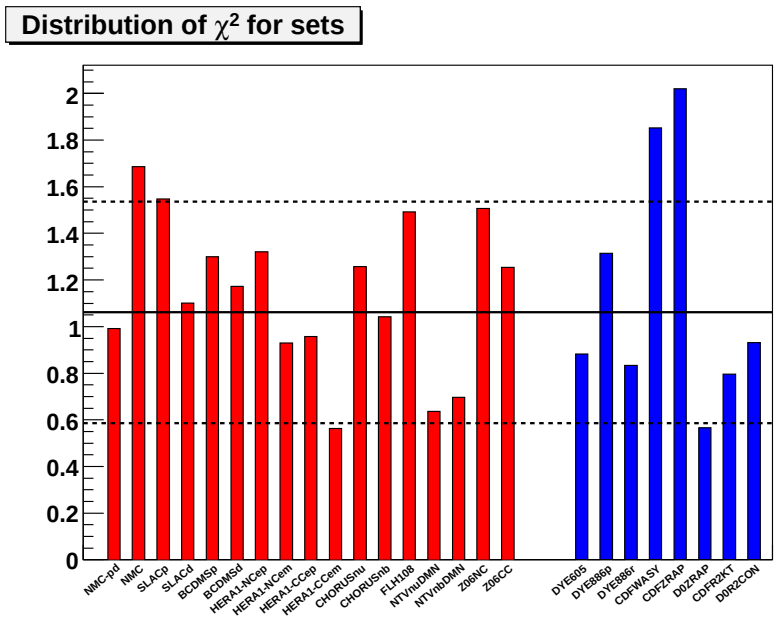


Figure 3.12: Values of the χ^2 for the datasets included in the NNPDF2.0 reference fit [71], listed in Table 3.7. The horizontal line corresponds to the unweighted average of χ^2 over datasets and and the black dashed line to the 1σ interval about it: $\langle\chi^2\rangle_{\rm sets} = 1.06$, $\sigma_{\chi^2} = 0.40$; DIS and hadronic datasets are grouped respectively to the left and right of the histogram and distinguished by different colors.

$\chi^2 \sim 0.6$ for CC electron data. The reasons for this distribution of values are unclear, however, we note that also in NNPDF1.2 [70] the χ^2 of the CC datasets was typically smaller than the average as well. We note also that the same pattern of χ^2 among the different datasets has been obtained within the framework of the HERAPDF1.0 analysis of these combined HERA-I dataset [84]. The CDF direct $W-$asymmetry measurements have $\chi^2 = 1.85$. The poor compatibility of these data with the rest of the global fit data was also noted in the global analysis of Ref. [153]. The quality of the fit to Z rapidity distribution data at the Tevatron differs widely between experiments: while an excellent fit is obtained for D0 data, CDF data are not so well described. This suggests that there might be problems of internal consistency between the two experiments. A similar pattern was observed in the MSTW08 global fit [43]. Note that these datasets have a very moderate impact on the global fit, as the χ^2 of these data is essentially the same in NNPDF2.0 and in NNPDF1.2, where they are not fitted.

Experiment	χ^2	$\langle E\rangle$	$\left\langle \sigma^{(\exp)}\right\rangle_{\rm dat}^{(\%)}$	$\left\langle \sigma^{(\rm net)}\right\rangle_{\rm dat}^{(\%)}$	$\left\langle \rho^{(\exp)}\right\rangle_{\rm dat}$	$\left\langle \rho^{(\rm net)}\right\rangle_{\rm dat}$
NMC-pd	0.99	2.05	1.8	0.5	0.03	0.36
NMC	1.69	2.79	4.9	1.7	0.16	0.77
SLAC	1.34	2.42	4.2	1.9	0.31	0.84
BCDMS	1.27	2.40	5.7	2.6	0.47	0.55
HERAI-AV	1.14	2.25	7.5	1.3	0.06	0.44
CHORUS	1.18	2.32	14.8	12.8	0.09	0.38
FLH108	1.49	2.51	71.9	3.3	0.65	0.68
NTVDMN	0.67	1.90	21.1	14.6	0.03	0.63
ZEUS-H2	1.51	2.66	13.6	1.2	0.29	0.58
DYE605	0.88	1.85	22.6	8.3	0.47	0.75
DYE866	1.28	2.35	20.8	9.1	0.20	0.45
CDFWASY	1.85	3.09	6.0	4.3	0.52	0.72
CDFZRAP	2.02	2.96	11.5	3.5	0.83	0.65
D0ZRAP	0.57	1.65	10.2	3.0	0.53	0.69
CDFR2KT	0.80	2.22	23.0	5.2	0.78	0.67
D0R2CON	0.93	1.92	16.2	6.0	0.78	0.64

Table 3.7: Same as Table 3.6 for individual individual experiments. Note that experimental uncertainties are always given in percentage. All statistical estimators are defined in Appendix B.

Finally, in Ref. [71] we have checked that if we run a very long fit without dynamical stopping, the χ^2 of the experiments whose values exceed the average by more than one sigma does not improve significantly. This shows that the deviation of these χ^2 values from the average is not due to underlearning.

3.2.3 Parton Distributions

In this section I show the results obtained in the various NNPDF analyses relative to the shape and the error of the parton distribution functions. I compare the results of the NNPDF fits to each others and to the MSTW and CTEQ parton fits. All PDF combinations are defined as in Eq. (1.58). The uncertainty bands shown are 1σ.

In Fig. 3.13 the predictions for some of the parton densities extracted in the subsequent analyses are shown. The statistical consistency of the NNPDF approach is apparent. The inclusion of new data reduces the uncertainty of the PDFs in regions were the new data provide further constraints, however leaving them compatible with the PDFs previously determined. For instance the gluon at large–x is much more constrained by the jet data, however at small–x it remains basically unconstrained by data and the green and blue bands look the same. The uncertainty of the valence–like PDFs is also greatly reduced by the inclusion of the Drell-Yan data. Notice that the sea asymmetry, whose sign was not determined in absence of these data, is naturally constrained to be positive just due to their inclusion. The consistency is facilitated by the fact that in

the subsequent analysis the same parametrisation and the same statistical features are employed.

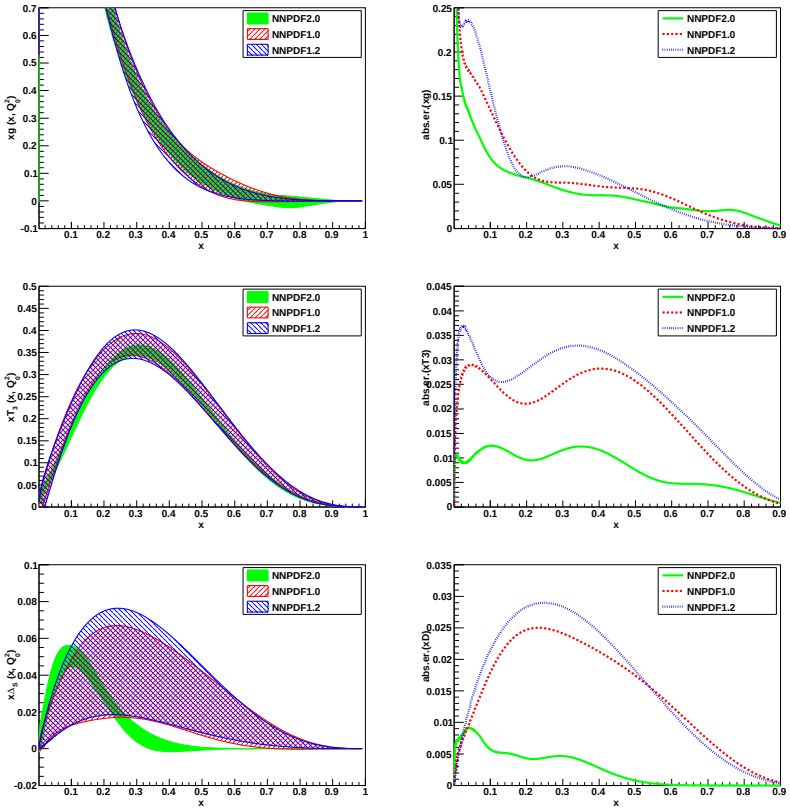


Figure 3.13: The gluon $xg(x)$ (top), the triplet $xT_3(x)$ (middle) and the sea asymmetry $x\Delta_S(x)$ (bottom) at the initial scale $Q_0^2 = 2$ GeV^2 from the NNPDF2.0 analysis. Both the PDFs (left) and their absolute uncertainties (right) are shown, compared to the previous NNPDF releases NNPDF1.0 [68] and NNPDF 1.2 [71].

The NNPDF2.0 PDFs are also compared to those extracted in the CTEQ6.6 [42] and MSTW08 [43] analyses. These next–to–leading order parton fits contain the same amount of experimental information, the only difference being in the different treatment of the heavy quark masses. In Fig. 3.14 we see that most NNPDF2.0 uncertainties are comparable to the CTEQ6.6 and MSTW08 ones; there are however some interesting exceptions. The uncertainty on strangeness, which NNPDF2.0 parametrises

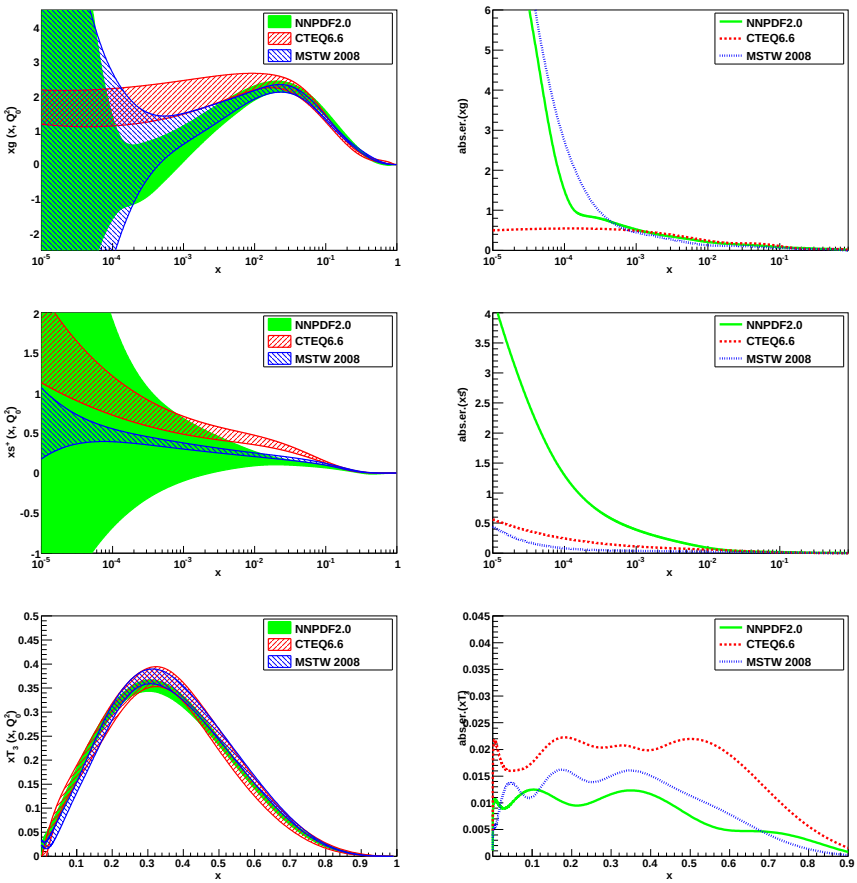


Figure 3.14: The gluon $xg(x)$ (top), the total strangeness $xS^+(x)$ (middle) and the triplet $T_3(x)$ (bottom) at the initial scale $Q_0^2 = 2$ GeV^2 from the NNPDF2.0 analysis. Both the PDFs (left) and their absolute uncertainties (right) are shown, compared to MSTW08 [43] and CTEQ6.6 [42] PDFs.

with as many parameters as any other PDF, is rather larger than those of MSTW08 and CTEQ6.6, in which these PDFs are parametrised with a very small number of parameters. The NNPDF2.0 uncertainty on total quark singlet (which contains a sizable strange contribution) is also larger. The uncertainty on the small–x gluon is significantly larger than that found by CTEQ6.6, but comparable to that of MSTW08, which has an extra parameter to describe the small–x gluon in comparison to CTEQ6.6. The uncertainty on the triplet combination is rather smaller in NNPDF2.0 than either

MSTW08 or CTEQ6.6, largely due to the impact of Drell-Yan data, which are found to be completely consistent with DIS data within our NLO treatment.

An important advantage of the Monte Carlo method used in the NNPDF approach to determine PDF uncertainties is that, unlike in a Hessian approach, one does not have to rely on linear error propagation. For instance it is possible to test for non-Gaussian distribution of the fitted PDFs even though our starting data and data replicas are Gaussian distributed. A simple way for doing that is to compute a 68% confidence level (C.L.) for it (which is straightforwardly done in a Monte Carlo approach), and compare the result to the standard deviation. In Fig. 3.15 this comparison is shown for some NNPDF2.0 PDFs at the initial scale as a function of x. The plots show that, in the extrapolation region for most PDFs deviations from Gaussian behaviour are sizable. This is especially noticeable for the gluon at small–x, and for the quark singlet and total strangeness both at small and large–x. However, in Ref. [71] is shown that in the regions in which the PDFs are constrained by experimental data the standard deviation and the 68% confidence levels coincide to good approximation, thus suggesting a Gaussian behaviour.

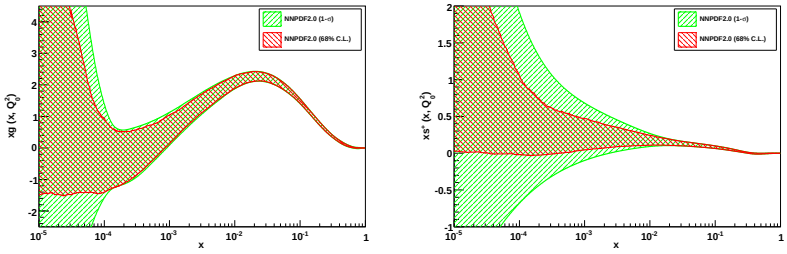


Figure 3.15: Comparison of 68% confidence level and 1σ intervals for the gluon (left) and the total strangeness (right). The PDFs have been determined in the NNPDF2.0 analysis [71] at the initial scale $Q_0^2 = 2$ GeV^2.

Deviations from Gaussian behaviour are sometimes related to positivity constraints Eq. (3.33): for instance positivity of F_L and the dimuon cross–section limits the possibility for the small–x gluon and strange sea PDFs respectively to go negative, thereby leading to an asymmetric uncertainty band. In order to assess quantitatively the effect of the positivity constraints, in Fig. 3.16 PDFs with uncertainties determined as 68% C.L. with and without positivity constraints are compared. We see that positivity of $F_L(x, Q^2)$ leads to substantial uncertainty reduction in the small-x gluon. Note that there is nevertheless a kinematic region in which the gluon goes negative by a small amount, though F_L remains positive. Also, removing positivity of the dimuon cross

section would lead to a much softer strange sea at small-x with rather larger uncertainties. This in turn leads to a softer small-x singlet, also with larger uncertainties. This is due to the fact that below $x \lesssim 0.01$, where no neutrino data are available, positivity is the only constraint on the total strangeness s^+. It is also interesting to observe that positivity also has the effect of stabilising the replica sample: indeed, the 68% confidence levels computed without positivity display some visible fluctuations which would only be smoothed out by using a significantly wider replica sample. These fluctuations are absent when positivity is imposed, meaning that such wide fluctuations in individual replicas are removed by the constraint.

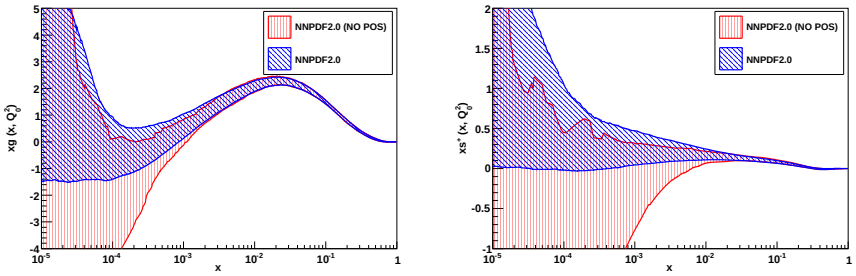


Figure 3.16: Gluon (left) and total strangeness (right) PDFs determined in the NNPDF2.0 analysis [71] at the initial scale $Q_0^2 = 2$ GeV^2 with and without positivity constraints. All uncertainty bands are determined as 68% confidence levels. PDFs not shown here are not affected by the positivity constraints.

3.2.4 Stability

An advantage of the NNPDF approach is that various features of the PDF set can be assessed using standard statistical tools. In this section, we assess the stability of the fit and the reliability of the error estimate by discussing three examples

1. the stability upon the choice of parametrisations originally presented in Ref. [68].
2. the stability upon the addition of unconstrained parametrisations presented in Ref. [69].
3. the detailed comparison between the NNPDF1.2 and NNPDF2.0 fits, by varying one by one each of the procedural aspects and computing the distances presented in Ref. [71].

In the first analysis, I show the dependence of results on the architecture of the neural networks. Specifically, we reduced the architecture from 2–5–3–1 to 2–4–3–1, thereby decreasing the number of parameters of each PDF from 37 to 31, the total number from 185 to 155. In Fig. 3.17 the best–fits and the error bands for the gluon

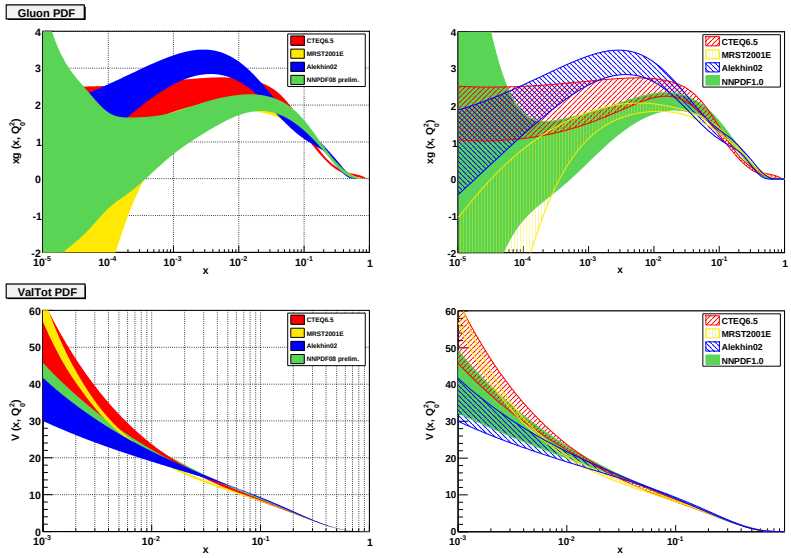


Figure 3.17: Comparison of best–fit gluon (top) and total valence (bottom) at the initial scale $Q_0^2 = 2$ GeV^2 obtained in the NNPDF1.0 analysis [68] with two different architectures: 2–4–3–1 (left) and 2–5–3–1 (right).

and the total valence distributions obtained with the the two different architectures are shown. The plots suggest a remarkable independence of the parametrisation. In order to quantify this statement, the distance between the two ensembles obtained with different parametrisations has to be evaluated. Results are given in Table 3.8. The distances confirm the result: fluctuations are at most at the 2σ level in poorly controlled quantities, such as the value of the light quark sea asymmetry in the extrapolation region or the uncertainty on the triplet combination in the extrapolation region, which in the NNPDF1.0 analysis [68] is poorly constrained due to the lack of Drell–Yan data. Results are indeed independent of the number of parameters.

In the second analysis I show the results obtained in the NNPDF1.1 fit [69]. The latter is not much useful for practical purposes, however it provides a proof of the consistency of the whole NNPDF procedure. In the NNPDF1.1 analysis, the strange parton distributions $s^{\pm} = s \pm \bar{s}$ were parametrised by two independent neural net-

	Data	Extrapolation
$\Sigma(x,Q_0^2)$	$5\ 10^{-4} \leq x \leq 0.1$	$10^{-5} \leq x \leq 10^{-4}$
$\langle d[q]\rangle$	0.98	1.25
$\langle d[\sigma]\rangle$	1.14	1.34
$g(x,Q_0^2)$	$5\ 10^{-4} \leq x \leq 0.1$	$10^{-5} \leq x \leq 10^{-4}$
$\langle d[q]\rangle$	1.52	1.15
$\langle d[\sigma]\rangle$	1.16	1.07
$T_3(x,Q_0^2)$	$0.05 \leq x \leq 0.75$	$10^{-3} \leq x \leq 10^{-2}$
$\langle d[q]\rangle$	1.00	1.11
$\langle d[\sigma]\rangle$	1.76	2.27
$V(x,Q_0^2)$	$0.1 \leq x \leq 0.6$	$3\ 10^{-3} \leq x \leq 3\ 10^{-2}$
$\langle d[q]\rangle$	1.30	0.90
$\langle d[\sigma]\rangle$	1.10	0.98
$\Delta_S(x,Q_0^2)$	$0.1 \leq x \leq 0.6$	$3\ 10^{-3} \leq x \leq 3\ 10^{-2}$
$\langle d[q]\rangle$	1.04	1.91
$\langle d[\sigma]\rangle$	1.44	1.80

Table 3.8: Distance between results obtained from a sets of 100 PDFs with neural network architecture 2-5-3-1 and a sets of 100 PDFs with neural network architecture 2-4-3-1.

works, instead of being taken to be proportional to the light antiquark distribution as in NNPDF1.0. However, the dataset is the same as for NNPDF1.0: so the s^+ distribution is only very weakly constrained, and the s^- is essentially unconstrained by the the data. The only weak constraint to the strangeness is given by the CHORUS and the HERA charged–current data as well as by the strange valence sum rule. Nevertheless, when results of this pair of fits are compared, they show remarkable stability, despite the fact that each neural network is parametrised by a very redundant set of parameters (the addition of two neural nets results in the addition of 74 extra free parameters in the fit). In Fig. 3.18 I show the results from the NNPDF1.1 analysis for the $\Sigma(x)$, $g(x)$, $s_+(x)$ and $s_-(x)$ distributions compared to NNPDF1.0. We can see that the central values for both PDFs are reasonably close between NNPDF1.0 and NNPDF1.1, thus ensuring the validity of the flavour assumptions in the former case. Moreover, the parton distributions which are unaffected by the addition of independent strange degrees of freedom (such as the gluon) are unchanged, and the only marked effect of the independent parametrisation of strangeness is an increase, by about a factor two, of the uncertainty on the total valence quark distribution ($V = u - \bar{u} + d - \bar{d} + s - \bar{s}$). Remarkably, statistical analysis of the NNPDF1.0 set alone was already sufficient to show that the uncertainty on this combination was underestimated [68]. The other PDFs are fairly stable, which is an important result since both two new input PDFs and a randomisation of the preprocessing have been incorporated in the new analysis. A comparable increase in uncertainty is observed in the extrapolation region of

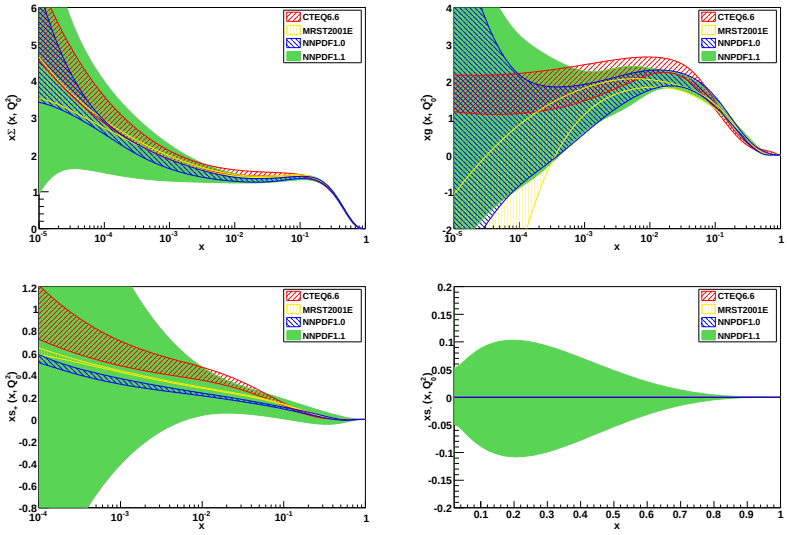


Figure 3.18: The NNPDF1.1 [69] singlet (top left), gluon(top right), total strangeness (bottom left) and strange valence (bottom right) compared to the NNPDF1.0, CTEQ66 and MRST01 sets.

$\Sigma(x)$, which can be attributed to the extra flexibility induced by the presence of the independent $s_+(x)$ PDF.

In the third part of this section, starting from NNPDF1.2, we perform a series of fits in which the procedural aspects are varied in succession, followed by another series of fit where new datasets are introduced one by one. At each step in this procedure we examine the general features of the fit by evaluating the quality of the fit, which is shown in Table 3.9, and computing the distances with respect to the precedent fit.

- *Effect of the improved genetic algorithm and stopping criterion (IGA).*
  The improvement in neural network training described in Section 4.1 leads to a significant improvement in fit quality, as one can see in Table 3.9: each replica fits better the corresponding data replica (lower $\langle E\rangle$), and also each replica neural network is more efficient in subtracting the statistical noise from data (lower $\langle\chi^{2\,(k)}\rangle$), thereby leading to a better global fit (lower $\chi^2_{\rm tot}$). The improvement is due to the improvement in fit quality of fixed–target DIS experiments (NMC, BCDMS and CHORUS) are known to have a certain amount of data inconsistency [136, 68, 154], without change in fit quality for other experiments: this means that the new algorithm is more efficient in leading to a balanced fit qual-

Fit NNPDF_	1.2	1.2+IGA	1.2+IGA+t_0	2.0 DIS	2.0 DIS+JET	2.0
$\chi^2_{\rm tot}$	1.32	1.16	1.12	1.20	1.18	1.21
$\langle E\rangle$	2.79	2.41	2.24	2.31	2.28	2.32
$\langle E_{\rm tr}\rangle$	2.75	2.39	2.20	2.28	2.24	2.29
$\langle E_{\rm val}\rangle$	2.80	2.46	2.27	2.34	2.32	2.35
$\left\langle \chi^{2(k)} \right\rangle$	1.60	1.28	1.21	1.29	1.27	1.29
NMC-pd	1.48	0.97	0.87	0.85	0.86	0.99
NMC	1.68	1.72	1.65	1.69	1.66	1.69
SLAC	1.20	1.42	1.33	1.37	1.31	1.34
BCDMS	1.59	1.33	1.25	1.26	1.27	1.27
HERAI	1.05	0.98	0.96	1.13	1.13	1.14
CHORUS	1.39	1.13	1.12	1.13	1.11	1.18
FLH108	1.70	1.53	1.53	1.51	1.49	1.49
NTVDMN	0.64	0.81	0.71	0.71	0.75	0.67
ZEUS-H2	1.52	1.51	1.49	1.50	1.49	1.51
DYE605	*11.19*	*22.89*	*8.21*	*7.32*	*10.35*	0.88
DYE866	*53.20*	*4.81*	*2.46*	*2.24*	*2.59*	1.28
CDFWASY	*26.76*	*28.22*	*20.32*	*13.06*	*14.13*	1.85
CDFZRAP	*1.65*	*4.61*	*3.13*	*3.12*	*3.31*	2.02
D0ZRAP	*0.56*	*0.80*	*0.65*	*0.65*	*0.68*	0.47
CDFR2KT	*1.10*	*0.95*	*0.78*	*0.91*	0.79	0.80
D0R2CON	*1.18*	*1.07*	*0.94*	*1.00*	0.93	0.93

Table 3.9: Statistical estimators for the sequence of fits that take from NNPDF1.2 to NNPDF2.0. The estimators shown for NNPDF1.2 are as in Tab. 5-6 of Ref. [70] and those for NNPDF2.0 are as in Tab. 3.6–3.7. Estimators are shown for the total datasets in the upper part of the table, while the lower part of the table shows the χ^2 for each individual experimental dataset. Values of the χ^2 for data not included in any given fit are shown in italic; the total $\chi^2_{\rm tot}$ shown in the first line does not include the contribution from these data. The value of the χ^2 in the HERAI line refers in the first three columns of the table to the weighted sum of the H1 and ZEUS data, and in the latter three columns to the combined dataset, according to which data has been included in the fit.

ity between experiments, without some data being underlearnt while others are overlearnt. The IGA affects essentially all PDFs by reducing their uncertainties, the two fits are always consistent at the 1σ level.

- *Impact of the treatment of normalisation uncertainties.*
  The previous fit is modified by implementing the improved treatment of the normalisation uncertainties described in Chap. 3. A small but noticeable improvement in the quality of the fit is observed (the total χ^2 going from 1.16 to 1.12). The latter is mostly due to the improved description of the fixed–target DIS experiments. The distances between this fit and the previous one, which only differ in the treatment of normalisation uncertainties, are displayed in Fig. 3.19. The PDFs which are mostly affected are the small–x singlet and gluon and the triplet distributions.

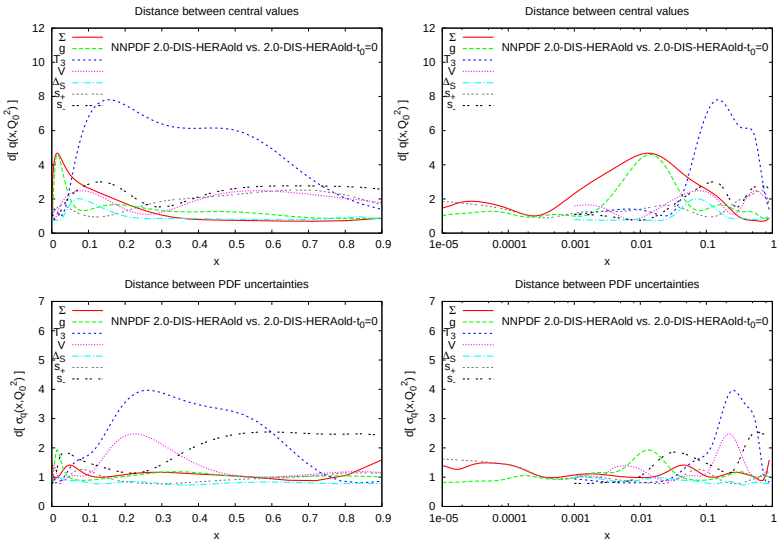


Figure 3.19: Distances between the PDFs obtained in a fit with improved genetic algorithm (IGA) and a fit with IGA + improved treatment of normalisation uncertainties (IGA+t_0). The distances are shown in linear (left) and logarithmic (right) scales for the central values (top) and the 1σ uncertainties.

- *Impact of the combined HERA-I data.*
  The previous IGA+t_0 fit is modified by replacing the ZEUS and H1 data with the new combined HERA-I [84]. This fit is now identical to the NNPDF2.0 fit, but with only DIS data included (NNPDF2.0-DIS). The inclusion of the

very precise HERA-I data leads to a slight deterioration of fit quality (the total χ^2 going from 1.12 to 1.20) which remains however still better than that of NNPDF1.2. This deterioration is concentrated in the HERA data themselves, with the quality of the fit to all other data unchanged. This suggests good consistency of the HERA and fixed target data, but with the accuracy of the combined HERA-I data now exceeding the accuracy of the theory used to describe them in NNPDF2.0: for instance the lack of inclusion of charm mass corrections, but also possibly deviations from NLO DGLAP at small–x, or possible evidence for NNLO corrections at larger x. A particularly interesting aspect of this fit is that the quality of the fit to Drell-Yan data (not fitted), which was poor in all previous fits, improves considerably, especially for the W asymmetry. This suggests that the accuracy of the charged–current data in the HERA-I combined set is now sufficient to provide some handle on the flavour decomposition of the sea at large–x which is only weakly constrained by neutral current DIS data, and strongly constrained by DY data. The distances between these fits is shown in Fig. 3.20: the impact of the combined HERA data is a moderate but generalised improvement in accuracy at small–x.

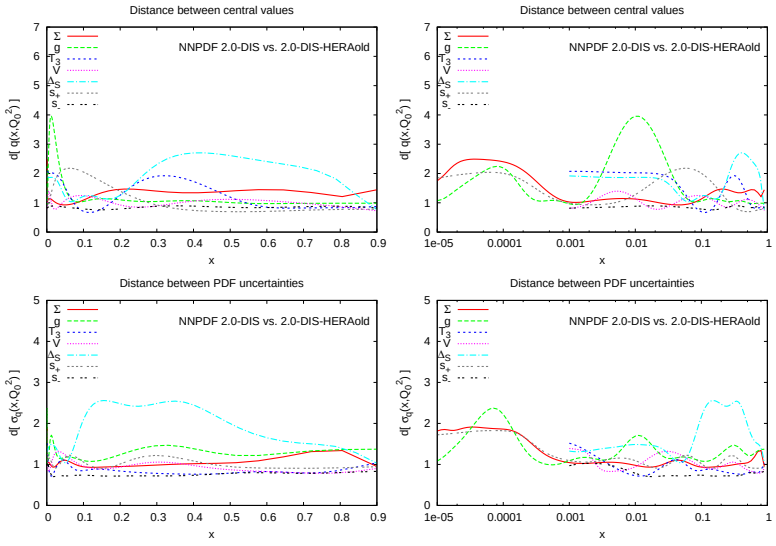


Figure 3.20: Distances between the PDFs obtained with a IGA+t_0 fit of Fig. 3.19 and a fit in which the separate H1 and ZEUS data are replaced by the combined HERA-I DIS data (NNPDF2.0 DIS). The distances are shown in linear (left) and logarithmic (right) scales for the central values (top) and the 1σ uncertainties.

- *Impact of jet data.*
  The addition of jet data to the 2.0-DIS fit leaves the quality of the global fit unchanged. This demonstrates the perfect compatibility of jet data with DIS data: in fact, the quality of the fit to jet data was quite good even in all previous fits, in which they were not included in the fitted data–set. The distance between the 2.0-DIS and 2.0-DIS+JET fits, displayed in Fig. 3.21, shows that these data affect almost only the gluon, as one would expect [155], leading to a better determination at medium– and large–x. This is shown in Fig. 3.22, where the gluons of 2.0-DIS and 2.0-DIS+JET are compared.

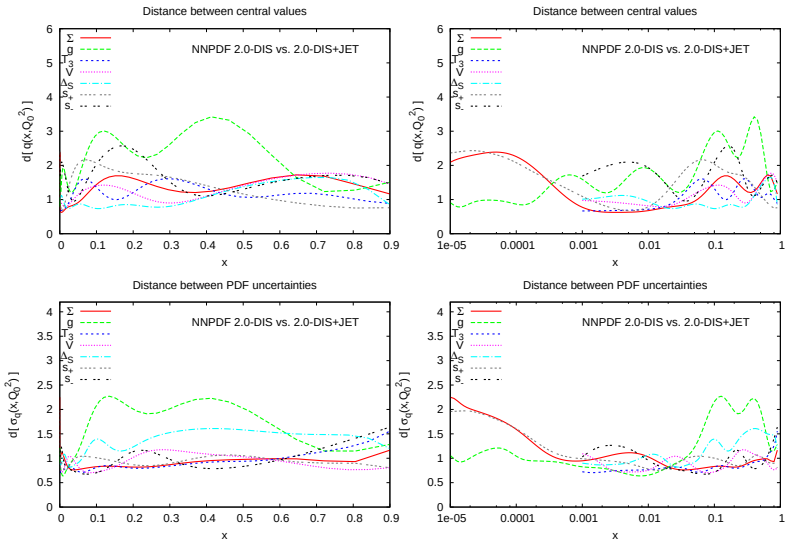


Figure 3.21: Distances between the PDFs obtained from a NNPDF2.0 DIS fit of Fig. 3.20 and those obtained from a fit in which jet data are also included (NNPDF2.0 DIS+JET). The distances are shown in linear (left) and logarithmic (right) scales for the central values (top) and the 1σ uncertainties.

- *Impact of Drell-Yan data.*
  The addition of Drell-Yan data to the 2.0-DIS+JET fit leaves the quality of the global fit unchanged. Taken together with the previous comparison of the 2.0-DIS and 2.0-DIS+JET data, this shows that DIS data and hadronic data are fully compatible, and furthermore the two classes of hadronic data included here, DY and inclusive jets, are compatible with each other. Minor incompatibilities only appear within each dataset (typically due to some subset of data points or, in the case of Drell-Yan to the CDF W asymmetry and Z rapidity distribution data).

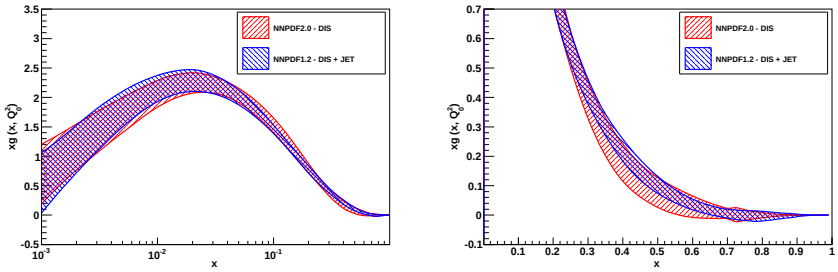


Figure 3.22: Comparison between the gluon density at the initial scale $Q_0^2 = 2\,\mathrm{GeV}^2$ at small (left) and large (right) x obtained from NNPDF2.0 DIS fit of Fig. 3.20 and a fit in which jet data are also included (NNPDF2.0 DIS+JET) (the distances are shown in Fig. 3.21).

However, the quality of the fit to Drell-Yan data was generally poor when they were not included in the fit, due to the fact that they are sensitive to the separation of individual flavours at large–x which is only very weakly constrained by other data. The distances between the 2.0-DIS+JET and the full NNPDF2.0 fits, displayed in Fig. 3.23, show the sizable impact of the Drell-Yan data on all valence–like PDF combinations at medium and large-x: the triplet, the valence, the sea asymmetry and the strangeness asymmetry. The significant improvement in accuracy on all these PDFs is apparent in Fig. 3.22. The remarkable improvement in the accuracy of the determination of the strangeness asymmetry $s^-(x)$ will turn out to have relevant phenomenological implications for the so–called NuTeV anomaly, as I discuss in Chap. 6.

3.2.5 Theoretical uncertainty and outlook

In the determination of PDFs, all systematic uncertainties in the data have been accounted for in the Monte Carlo data generation: they are then propagated through the fitting procedure onto the ensemble of fitted PDFs. Therefore, error bands already include both statistical and systematic uncertainties of the data. On top of these, however, there are the theoretical uncertainties mentioned in Chap. 3, which might cause systematic shifts in PDFs central values and uncertainties.

The two main sources of theoretical uncertainties in the NNPDF analyses are related to the fact that they are performed at NLO, and thus they neglect effects at NNLO and beyond, and to the approximate zero–mass (ZM) scheme used to deal with the finite masses of the heavy quarks.

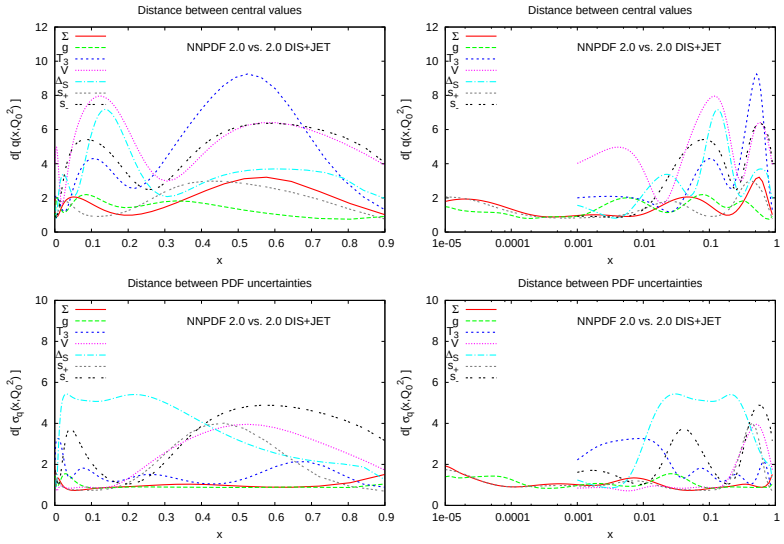


Figure 3.23: Distances between the PDFs obtained from the NNPDF2.0 DIS+JET fit of Fig. 3.21 and the reference NNPDF2.0 fit (Drell-Yan data also included). The distances are shown in linear (left) and logarithmic (right) scales for the central values (top) and the 1σ uncertainties.

A way for estimating the theoretical error due to higher–order contributions in perturbation theory consists in varying the factorisation and renormalisation scales. In order to combine PDFs and scale variation uncertainties we could produce several sets with varying μ_F and μ_R and combine the obtained ensembles according to the method proposed in Ref. [156] to combine PDFs and α_s uncertainties. The latter is going to be described in details in Sect. 6.2. Another possible method would consist in taking these scales as random variables distributed between μ/k and $k\mu$, where k sets their variation range, and letting them fluctuate between replicas. Such analysis has not been performed yet, but it represents an interesting possibility to be explored in the near future. Notice that the same procedure might be applied to estimate the uncertainty related to all parameters which enters into a PDF determination, such as the values of the masses of the heavy quarks m_c, m_b and m_t.

A less accurate way to assess the uncertainties related to higher perturbative orders has been adopted in Ref. [68]. The reference NNPDF1.0 NLO fit was compared to the fit repeated at LO. The comparison is displayed in Fig. 3.25. It shows that the size of the uncertainties remains essentially unchanged. Furthermore, all central values of the LO fit vary by an amount which is compatible with statistical fluctuations, with the

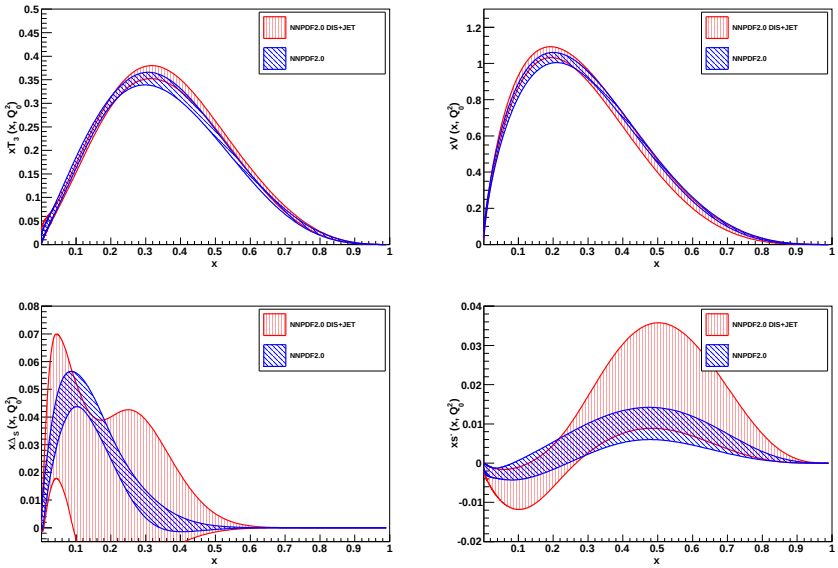


Figure 3.24: Comparison between PDFs at the initial scale $Q_0^2 = 2\,\mathrm{GeV}^2$ from the NNPDF2.0 DIS+JET fit of Fig. 3.21 and the reference NNPDF2.0 fit which included also Drell–Yan data (the distances are shown in Fig. 3.23). From left to right and from top to bottom the following distributions are show: triplet, valence, sea asymmetry and strange valence.

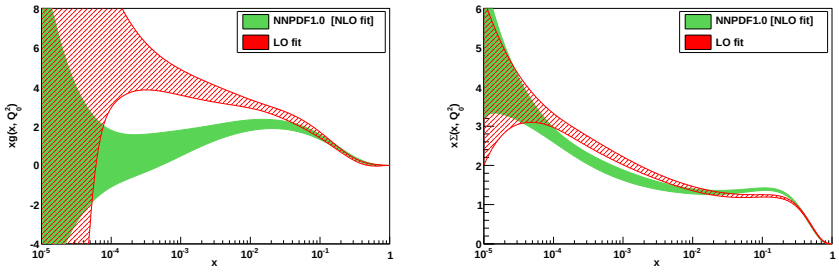


Figure 3.25: Comparison of the NLO NNPDF1.0 fit [68] and LO fit results for the gluon (left) and quark singlet PDFs (right).

exception of the singlet and gluon shown in figure, which vary by about 2-3σ. This is expected due to the fact that the gluon contribution to DIS coefficient functions only starts at NLO. On the other hand, the quality of the fit deteriorates significantly when using LO theory, as we might also expect. This comparison suggests that in the NNPDF1.0 analysis the theoretical uncertainties due to lack of inclusion of NNLO

corrections are negligible on the scale of statistical uncertainties. This conclusion is based on the observation that NNLO corrections to all PDFs are known to be smaller than the typical NLO corrections (i.e. the NLO-LO difference) in the nonsinglet sector [157, 66] and the observation that the latter corrections are already smaller than the statistical uncertainty by a factor of two. A more accurate analysis, which is in preparation in the context of the definition of a suitable LO parton fit able to match the requirements of the LO Monte Carlo event generators, refers to the most up–to–date NNPDF2.0 reference fit [71] and will enable us to draw more detailed conclusions.

Another source of theoretical uncertainty is the treatment of the heavy quark masses. The theoretical framework has been extensively discussed in Chap. 2. All publicly available NNPDF analyses are performed according to the ZM–VFN scheme, with the exception of the dimuon data, for which the prediction is evaluated in the improved ZM scheme. This approximate treatment of the heavy quark masses introduces a shift in theoretical predictions that ought to be quantified. Indeed the latter might be responsible for the differences in the predictions obtained for some LHC standard candle processes when compared to those obtained with other parton sets including a GM–VFNS in their analyses[4]. To improve over the zero–mass approximation, in the upcoming NNPDF2.1 release [105], a general mass scheme is implemented, in particular the FONLL scheme [54]. For the time being, only preliminary results based on a fit to the NNPDF2.0 dataset [71], supplemented by charm structure function data $F_2^c(x, Q^2)$ are available. In Fig. 3.26 a preliminary comparison of the singlet and gluon PDFs at the initial evolution scale $Q_0^2 = 2$ GeV^2 in NNPDF2.1 and NNPDF2.0, normalised to the NNPDF2.0 central values is displayed. It shows that the inclusion of heavy quark mass effects leads to a increase in the singlet at medium and small-x, as well as to a marked increase in the small-x gluon. However differences appear to be within the PDF uncertainty bands. This preliminary analysis implies that heavy quark mass effects should modify the NNPDF2.0 predictions for LHC observables by 1σ or so at most. The results still need to be validated and are going to be made publicly available soon.

Further possible sources of theoretical uncertainty, which would be intersting to assess in future studies, include effects related to large– and small–x resummation of the perturbative expansion, higher twist corrections, and nuclear effects.

[4]For a comparison of standard candle predictions, see Chap. 6.

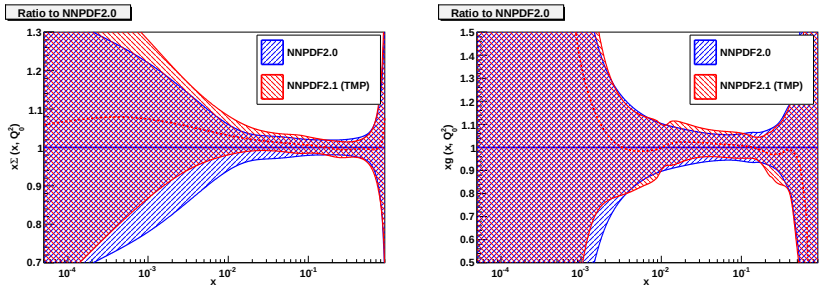


Figure 3.26: The ratio of the preliminary NNPDF2.1 Singlet (left plot) and gluon PDFs (right plot) to the respective NNPDF2.0 ones, at the initial evolution scale $Q_0^2 = 2\,\mathrm{GeV}^2$. Error bands are normalised to the NNPDF2.0 central value.

Chapter 4

The FastKernel Method

A global analysis of Parton Distribution Functions involves the significant complication of dealing with the coupled evolution of a full set of parton distributions. Being PDFs extracted from many different experiments, one has to combine the evolved parton distributions with their coefficient functions for a wide set of different observables. In parton fits, this computation aims to be as fast and accurate as possible.

The hybrid x–N space solution of the DGLAP evolution equations was introduced in Ref. [67], for the evolution of a single non–singlet parton density. Here I describe the method elaborated for extending this formalism to the case of coupled evolution of a full set of PDFs. Furthermore, the FastKernel method is formulated. The latter, using higher order polynomial interpolations, yields a fast and accurate evolution of PDFs and a fast computation of Drell–Yan observables. The latter allows us to deal with the computation of hadronic observables at next–to–leading order without relying on any K–factor approximation. The accuracy of the partonic evolution is tested against other evolution codes and the precision of the computation of the NLO order Drell–Yan distributions is assessed.

4.1 Hybrid x-N space solution of DGLAP equations

The perturbative computation of physical observables involves the evolution of PDFs up to the scale of the measurements and their convolution with the hard cross-sections. One way of speeding up the computation of the observables during a PDF fit was

introduced in Ref. [67] and adopted since the earliest NNPDF fits. The main advantage is that each of the two computations can be optimised separately from a numerical point of view. In particular, one may use a Mellin-space approach to solve evolution equations, but adopt x-space parametrisation of PDFs. Moreover evolution kernels may be pre-computed, benchmarked, and stored once and for all before the fitting procedure.

In this section I discuss details of the method and its application to the computation of DIS observables in Refs. [67, 68]. In the next section I introduce an improved faster method which enabled us to broaden the analysis by including hadronic observables.

4.1.1 Leading-twist factorisation and evolution

Deep inelastic observables $F_I(x,Q^2)$ may always be expressed at leading twist as a convolution of parton distributions $f_j(x,Q^2)$ and hard coefficient functions $C_{Ij}(x,\alpha_s(Q^2))$, computed in perturbation theory, as it was shown in Eq. (1.44):

$$F_I(x,Q^2) = \sum_j C_{Ij}(x,\alpha_s(Q^2)) \otimes f_j(x,Q^2), \tag{4.1}$$

where $\otimes$ denotes the usual convolution product. The label I refers to the considered observable, which may be a structure function or a reduced cross-section, while j runs over parton distribution functions. The scale dependence of the parton distribution functions is given by the renormalisation group, or DGLAP equations

$$Q^2\frac{\partial}{\partial Q^2} f_i(x,Q^2) = \sum_j P_{ij}(x,\alpha_s(Q^2)) \otimes f_j(x,Q^2), \tag{4.2}$$

where P_{ij} are the Altarelli-Parisi splitting functions, explicitly given at leading–order in Eq. (1.51). The structure of the solution of these coupled integro-differential equations may be written as

$$f_i(x,Q^2) = \sum_j \Gamma_{ij}(x,\alpha_s,\alpha_s^0) \otimes f_j(x,Q_0^2), \tag{4.3}$$

where $f_j(x,Q_0^2)$ are the input PDFs and $\Gamma_{ij}(x,\alpha_s,\alpha_s^0)$ are the evolution factors. From now on, I use the shorthand notation introduced in Eq. (1.54), $\alpha_s \equiv \alpha_s(Q^2)$ $\alpha_s^0 \equiv$

$\alpha_s(Q_0^2)$. Substituting Eq. (4.3) into Eq. (4.1)

$$\begin{aligned} F_I(x,Q^2) &= \sum_{jk} C_{Ij}(x,\alpha_s) \otimes \Gamma_{jk}(x,\alpha_s,\alpha_s^0) \otimes f_k(x,Q_0^2) \\ &= \sum_{j} K_{Ij}(x,\alpha_s,\alpha_s^0) \otimes f_j(x,Q_0^2), \end{aligned} \tag{4.4}$$

where the hard kernel

$$K_{Ij}(x,\alpha_s,\alpha_s^0) = \sum_k C_{Ik}(x,\alpha_s) \otimes \Gamma_{kj}(x,\alpha_s,\alpha_s^0), \tag{4.5}$$

is the convolution between evolution factors and coefficient functions and may be computed in perturbation theory.

Performing many nested convolutions is numerically rather time consuming. However the hard kernels Eq. (4.5) are independent of the particular set of input PDFs adopted, and may thus be calculated only once and stored.

4.1.2 Calculating the evolved x–space PDFs

The QCD evolution equations are most easily solved using Mellin moments [32], as all the convolutions become simple products, and the equations may be solved in a closed form. In Sect. 2.1.3 the N-space solution of the DGLAP evolution equation is explicitly written up to next–to–leading order and may be easily generalised to higher orders. Given the explicit expression for the N–space evolution kernels and for the hard kernels, the problem is reduced to the computation of the Mellin inversion integral.

The x–space evolution factors are obtained by taking the inverse Mellin transforms of the solutions obtained in Eqs. (1.71, 1.77):

$$\Gamma(x,\alpha_s,\alpha_s^0) = \int_{\mathcal{C}} \frac{dN}{2\pi i} x^{-N} \Gamma(N,\alpha_s,\alpha_s^0). \tag{4.6}$$

On the other hand, the x–space hard kernels are obtained by taking the inverse Mellin transforms of the product between the N–space evolution kernels and the Mellin moments of the coefficient functions, Eq. (1.45)

$$\Gamma(x,\alpha_s,\alpha_s^0) = \int_{\mathcal{C}} \frac{dN}{2\pi i} x^{-N} \Gamma(N,\alpha_s,\alpha_s^0)\, C(N,\alpha_s). \tag{4.7}$$

In what follows, I explicitly consider the inversion of Γ evolution factors, but the same considerations hold for the hard kernels K.

The numerical computation of the Mellin inversion integral is delicate, because the oscillatory behaviour of the integrand at large-x is not damped by multiplication by an initial PDF. This problem is mitigated by a suitable choice of the integration path $\mathcal{C}$. One possible choice [67, 68, 70, 71] is the Talbot path, drawn in Fig. 4.1, which goes around the singularities at $N = 0, -1, -2, ...$ It is defined by the condition

$$N(\theta) \;=\; r\theta\left(\frac{1}{\tan\theta}+1\right) \qquad -\pi \leq \theta \leq \pi, \tag{4.8}$$

where r is a constant corresponding to the intercept of the curve on the real axis. To

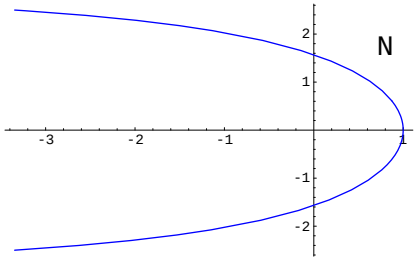


Figure 4.1: The path in the complex N-space followed by the Talbot integration path, Eq. (4.8) for r = 1.

further improve the numerical efficiency the Fixed Talbot algorithm can be used [158], where the integral is replaced by the sum

$$\Gamma(x) = \frac{r}{M}\left[\frac{1}{2}\Gamma(N=r)\,x^{-r} + \sum_{k=1}^{M-1}\mathrm{Re}\left[x^{-N(\theta_k)}\Gamma\left(N\left(\theta_k\right)\right)\left(1+i\sigma\left(\theta_k\right)\right)\right]\right], \tag{4.9}$$

where

$$\sigma(\theta) \equiv \theta + (\theta/\tan\theta - 1)/\tan\theta,$$

$\theta_k = k\pi/M$, and $r = 2M/(5\ln 1/x)$. As shown in Ref. [158], M yields the relative accuracy, i.e. the number of accurate digits. Sixteen digits are more than enough to achieve an accuracy of $\mathcal{O}(10^{-6})$ in the inversion. For the singlet evolution the Talbot

path must displaced by one unit to the right

$$\mathbf{\Gamma}_{\rm S}(x,\alpha_s,\alpha_s^0) = \int_{C+1} \frac{dN}{2\pi i} x^{-N} \mathbf{\Gamma}_{\rm S}(N,\alpha_s,\alpha_s^0) = x \int_C \frac{dN}{2\pi i} x^{-N} \mathbf{\Gamma}_{\rm S}(N-1,\alpha_s,\alpha_s^0), \tag{4.10}$$

since now the singularities are at $N = 1, 0, -1, \ldots$.
However all splitting functions, except the off-diagonal entries of the singlet matrix, diverge when $x \to 1$; this implies that the evolution kernels $\Gamma(x,\alpha_s,\alpha_s^0)$ will likewise be divergent as $x \to 1$, and must thus be interpreted as distributions. Specifically, one may define

$$\Gamma_{\rm NS}^{(+)}(x,\alpha_s,\alpha_s^0) = \Gamma_{\rm NS}(x,\alpha_s,\alpha_s^0) - G_{NS}(\alpha_s,\alpha_s^0)\delta(1-x)\,, \tag{4.11}$$

$$\mathbf{\Gamma}_{\rm S}^{(+)}(x,\alpha_s,\alpha_s^0) = \mathbf{\Gamma}_{\rm S}(x,\alpha_s,\alpha_s^0) - \mathbf{G}_S(\alpha_s,\alpha_s^0)\, x^{-1}\delta(1-x)\,, \tag{4.12}$$

where

$$G_{\rm NS}(\alpha_s,\alpha_s^0) = \int_0^1 dx\, \Gamma_{\rm NS}(x,\alpha_s,\alpha_s^0) = \Gamma_{\rm NS}(N,\alpha_s,\alpha_s^0)|_{N=1}, \tag{4.13}$$

$$\mathbf{G}_{\rm S}(\alpha_s,\alpha_s^0) = \int_0^1 dx\, x\, \mathbf{\Gamma}_{\rm S}(x,\alpha_s,\alpha_s^0) = \mathbf{\Gamma}_{\rm S}(N,\alpha_s,\alpha_s^0)|_{N=2}, \tag{4.14}$$

are all finite constants. The convolutions Eq. (4.3) may then be evaluated as

$$\begin{aligned} f_i(x,Q^2) &= G_{\rm NS}(\alpha_s,\alpha_s^0) f_i(x,Q_0^2) + \int_x^1 \frac{dy}{y}\, \Gamma_{\rm NS}^{(+)}(y,\alpha_s,\alpha_s^0) f_i\left(\frac{x}{y},Q_0^2\right) \\ &= \left(G_{\rm NS}(\alpha_s,\alpha_s^0) - \int_0^x dy\, \Gamma_{\rm NS}(y,\alpha_s,\alpha_s^0)\right) f_i(x,Q_0^2) \\ &+ \int_x^1 \frac{dy}{y}\, \Gamma_{\rm NS}(y,\alpha_s,\alpha_s^0) \left(f_i\left(\frac{x}{y},Q_0^2\right) - y f_i(x,Q_0^2),\right). \end{aligned} \tag{4.15}$$

for nonsinglet distributions f_i. Similarly

$$\begin{aligned} \mathbf{f}_{\rm S}(x,Q^2) &= \mathbf{G}_{\rm S}(\alpha_s,\alpha_s^0) \mathbf{f}_{\rm S}(x,Q_0^2) + \int_x^1 \frac{dy}{y}\, \mathbf{\Gamma}_{\rm S}^{(+)}(y,\alpha_s,\alpha_s^0) \mathbf{f}_{\rm S}\left(\frac{x}{y},Q_0^2\right) \\ &= \left(\mathbf{G}_{\rm S}(\alpha_s,\alpha_s^0) - \int_0^x dy\, y \mathbf{\Gamma}_{\rm S}(y,\alpha_s,\alpha_s^0)\right) \mathbf{f}_{\rm S}(x,Q_0^2) \\ &+ \int_x^1 \frac{dy}{y}\, \mathbf{\Gamma}_{\rm S}(y,\alpha_s,\alpha_s^0) \left(\mathbf{f}_{\rm S}\left(\frac{x}{y},Q_0^2\right) - y^2 \mathbf{f}_{\rm S}(x,Q_0^2)\right), \end{aligned} \tag{4.16}$$

for singlet distributions $\mathbf{f}_{\rm S}$, where now all integrals converge and can be computed numerically.

The solution of the evolution equations through the determination of $x-$space evolution factors, Eqs. (4.15, 4.16), is particularly efficient because of the universality of the evolution factors, i.e., their independence of the specific boundary condition which is being evolved. In this way the evolution factors can be pre-computed and stored, and then used during the process of parton fitting without having to recompute them each time.

4.1.3 LH benchmark

During the PDF fitting procedure, a given PDF set must be evolved many times up to the fixed values of (x, Q^2) at which data are available. For each (x, Q^2) the numerical determination of the right–hand side of Eqs. (4.15, 4.16) involves the evaluation of two contributions: the first requires the multiplication of the PDF by a predetermined constant $(G(\alpha_s, \alpha_s^0) - \int_0^x dy\, \Gamma(y, \alpha_s, \alpha_s^0))$; while the second requires a convolution of the predetermined evolution factor $\Gamma(y, \alpha_s, \alpha_s^0)$ with the subtracted PDF, and thus the numerical evaluation of the integral over y.

In Refs. [67, 68], the numerical integration is performed using a $N_{\rm quad}-$point Gaussian integration in each of the $2^{N_{\rm iter}+1} - 1$ intervals in which the integration range $(x, 1)$ of y is divided. The total number of points used to perform the convolutions in y is given by

$$N_{\rm pt} = N_{\rm quad} \left(2^{N_{\rm iter}+1} - 1\right) \ , \tag{4.17}$$

and the values of y are determined accordingly, for each given value of x.

The accuracy of the PDF evolution code has been cross-checked against the Les Houches PDF evolution benchmark tables [126, 101]. Those tables were obtained from a comparison of the HOPPET[159] and PEGASUS [32] evolution codes, which are $x-$space and $N-$space codes respectively. In order to perform a meaningful comparison, the iterated solution of the $N-$space evolution equations is used, Eqs. (1.71, 1.77), and the same initial PDFs and the same running coupling are used, following the procedure described in detail in Refs. [101, 126].

In Table 4.1 the relative difference $\epsilon_{\rm rel}$ for various combinations of PDFs between the NNPDF evolution and the benchmark tables of Refs. [101, 126] at NLO in the ZM-VFNS, for two different values of $N_{\rm iter}$, is shown. In the upper part of the table $N_{\rm iter} = 6$, that is, each convolution integral is performed with approximately 500 points. This choice leads to an accuracy which is enough to reproduce the Les

x	$\epsilon_{\rm rel}\,(u_v)$	$\epsilon_{\rm rel}\,(d_v)$	$\epsilon_{\rm rel}\,(\Sigma)$	$\epsilon_{\rm rel}\,\left(d+\bar{u}\right)$	$\epsilon_{\rm rel}\,(s+\bar{s})$	$\epsilon_{\rm rel}\,(g)$
$N_{\rm iter}=6$						
10^{-7}	$2.2\ 10^{-5}$	$8.1\ 10^{-6}$	$4.9\ 10^{-6}$	$1.5\ 10^{-5}$	$1.2\ 10^{-6}$	$2.2\ 10^{-5}$
10^{-6}	$6.3\ 10^{-6}$	$3.2\ 10^{-6}$	$9.8\ 10^{-6}$	$1.1\ 10^{-5}$	$5.4\ 10^{-6}$	$3.0\ 10^{-6}$
10^{-5}	$1.8\ 10^{-5}$	$1.4\ 10^{-5}$	$8.3\ 10^{-6}$	$3.0\ 10^{-6}$	$3.6\ 10^{-6}$	$1.4\ 10^{-6}$
10^{-4}	$3.1\ 10^{-5}$	$1.6\ 10^{-5}$	$3.6\ 10^{-5}$	$4.3\ 10^{-5}$	$3.3\ 10^{-5}$	$3.2\ 10^{-5}$
10^{-3}	$1.8\ 10^{-6}$	$1.2\ 10^{-5}$	$5.9\ 10^{-6}$	$5.8\ 10^{-6}$	$8.9\ 10^{-6}$	$3.6\ 10^{-6}$
10^{-2}	$2.8\ 10^{-5}$	$1.5\ 10^{-5}$	$4.7\ 10^{-5}$	$4.3\ 10^{-5}$	$4.6\ 10^{-5}$	$8.2\ 10^{-5}$
0.1	$3.2\ 10^{-6}$	$1.3\ 10^{-5}$	$3.0\ 10^{-6}$	$9.4\ 10^{-6}$	$2.1\ 10^{-5}$	$5.1\ 10^{-7}$
0.3	$1.9\ 10^{-6}$	$2.4\ 10^{-5}$	$6.5\ 10^{-6}$	$1.0\ 10^{-5}$	$3.2\ 10^{-6}$	$2.6\ 10^{-6}$
0.5	$1.70\ 10^{-5}$	$1.3\ 10^{-5}$	$1.5\ 10^{-5}$	$1.3\ 10^{-5}$	$3.0\ 10^{-6}$	$3.5\ 10^{-6}$
0.7	$7.0\ 10^{-5}$	$8.0\ 10^{-6}$	$5.9\ 10^{-5}$	$8.9\ 10^{-6}$	$2.4\ 10^{-5}$	$9.9\ 10^{-6}$
0.9	$1.4\ 10^{-5}$	$6.2\ 10^{-6}$	$1.3\ 10^{-5}$	$7.4\ 10^{-4}$	$1.8\ 10^{-3}$	$5.1\ 10^{-5}$
$N_{\rm iter}=4$						
10^{-7}	$4.2\ 10^{-2}$	$4.5\ 10^{-2}$	$5.1\ 10^{-2}$	$5.1\ 10^{-2}$	$5.1\ 10^{-2}$	$5.1\ 10^{-2}$
10^{-6}	$1.6\ 10^{-2}$	$1.8\ 10^{-2}$	$2.4\ 10^{-2}$	$2.3\ 10^{-2}$	$2.4\ 10^{-2}$	$2.5\ 10^{-2}$
10^{-5}	$4.9\ 10^{-3}$	$4.4\ 10^{-3}$	$8.7\ 10^{-3}$	$8.3\ 10^{-3}$	$8.7\ 10^{-3}$	$9.6\ 10^{-3}$
10^{-4}	$2.3\ 10^{-3}$	$2.2\ 10^{-3}$	$3.9\ 10^{-3}$	$3.7\ 10^{-3}$	$3.9\ 10^{-3}$	$4.4\ 10^{-3}$
10^{-3}	$1.1\ 10^{-3}$	$6.7\ 10^{-4}$	$3.5\ 10^{-3}$	$3.0\ 10^{-3}$	$3.4\ 10^{-3}$	$4.6\ 10^{-3}$
10^{-2}	$1.5\ 10^{-3}$	$8.5\ 10^{-4}$	$3.4\ 10^{-3}$	$2.7\ 10^{-3}$	$3.7\ 10^{-3}$	$5.5\ 10^{-3}$
0.1	$3.9\ 10^{-6}$	$1.3\ 10^{-5}$	$4.3\ 10^{-6}$	$1.1\ 10^{-5}$	$2.4\ 10^{-5}$	$1.0\ 10^{-4}$
0.3	$1.9\ 10^{-6}$	$2.6\ 10^{-5}$	$6.6\ 10^{-6}$	1.610^{-5}	$5.9\ 10^{-6}$	$7.1\ 10^{-7}$
0.5	$1.6\ 10^{-5}$	$1.1\ 10^{-5}$	$1.4\ 10^{-5}$	$2.0\ 10^{-5}$	$5.8\ 10^{-6}$	$3.3\ 10^{-5}$
0.7	$6.8\ 10^{-5}$	$1.2\ 10^{-5}$	$5.7\ 10^{-5}$	$6.5\ 10^{-6}$	$4.6\ 10^{-5}$	$3.4\ 10^{-5}$
0.9	$1.4\ 10^{-5}$	$5.1\ 10^{-5}$	$1.6\ 10^{-5}$	$6.4\ 10^{-4}$	$1.7\ 10^{-3}$	$1.2\ 10^{-4}$

Table 4.1: Comparison of the accuracy of our PDF evolution with respect to the Les Houches benchmark tables for different PDF combinations at NLO in the ZM-VFNS. We show results for two values of $N_{\rm iter}$, which define the number of points over which the Gaussian integrations are performed, as discussed in the text. The number of Gaussian points in each interval is set to $N_{\rm pts}=4$.

Houches tables with $\mathcal{O}\left(10^{-5}\right)$ precision for all values of x, which is the nominal precision of the agreement between HOPPET and PEGASUS. In the lower part of Table 4.1 the accuracy results for the actual parameters which are used in the NNPDF1.0 fit [68] are shown. An integration with 128 points, corresponding to $N_{\rm iter}=4$, provides a sufficient accuracy, considering the typical sizes of both experimental and theoretical uncertainties.

The main point to keep in mind is that in this approach, for each measurement evaluated at a different value of x, the grid determined by the distribution of the Gaussian points is different. Therefore, the computational cost increases linearly with the number of data points included in the analysis. This slows down the fitting procedure and makes it practically impossible if a double convolution is involved, as in the com-

putation of Drell–Yan or Jet observables. This problem is solved by the FastKernel method, discussed in Sect. 5.2, where instead the same grid is used independently of x.

4.1.4 Hard cross-sections and physical observables

To determine the input PDFs, we must compare the predictions obtained with them to the experimental data. This involves convolution of the evolved PDFs with the hard coefficient functions defined in Eq. (4.1). This may be done most efficiently by pre-computing the hard kernels Eq. (4.5), which may then be convoluted with the initial PDFs as in Eq. (4.4). In Mellin space

$$K_{Ij}(N,\alpha_s,\alpha_s^0) = \sum_k C_{Ik}(N,\alpha_s)\Gamma_{kj}(N,\alpha_s,\alpha_s^0), \tag{4.18}$$

so the most efficient procedure is to compute $K_{Ij}(N,\alpha_s,\alpha_s^0)$, and invert the Mellin transform using a formula corresponding to Eq. (4.10), i.e.

$$K_{Ij}(x,\alpha_s,\alpha_s^0) = \begin{cases} \int_C \frac{dN}{2\pi i} x^{-N} K_{Ij}(N,\alpha_s,\alpha_s^0), & \text{if} \quad j = T,V, \\ x\int_C \frac{dN}{2\pi i} x^{-N} K_{Ij}(N-1,\alpha_s,\alpha_s^0) & \text{if} \quad j = \Sigma,g. \end{cases} \tag{4.19}$$

The convolutions may then be performed using formulae similar to those used in Eqs. (4.15, 4.16), writing $F_I(x,Q^2) = \sum_j F_{Ij}(x,Q^2)$, for the nonsinglet contributions (i.e. $j = T,V$)

$$\begin{aligned} F_{Ij}(x,Q^2) &= \left(\kappa_{Ij}(\alpha_s,\alpha_s^0) - \int_0^x dy\, K_{Ij}(y,\alpha_s,\alpha_s^0)\right) f_j(x,Q_0^2) \\ &+ \int_x^1 \frac{dy}{y} K_{Ij}(y,\alpha_s,\alpha_s^0)\left(f_j\left(\frac{x}{y},Q_0^2\right) - y f_j(x,Q_0^2),\right) \end{aligned} \tag{4.20}$$

while for the singlets (i.e. $j = \Sigma, g$)

$$\begin{aligned} F_{Ij}(x,Q^2) &= \left(\kappa_{Ij}(\alpha_s,\alpha_s^0) - \int_0^x dy\, y K_{Ij}(y,\alpha_s,\alpha_s^0)\right) f_{Ij}(x,Q_0^2) \\ &+ \int_x^1 \frac{dy}{y} K_{Ij}(y,\alpha_s,\alpha_s^0)\left(f_j\left(\frac{x}{y},Q_0^2\right) - y^2 f_j(x,Q_0^2)\right) \end{aligned} \tag{4.21}$$

where

$$\kappa_{Ij}(\alpha_s,\alpha_s^0) = \begin{cases} \int_0^1 dx\, K_{Ij}(x,\alpha_s,\alpha_s^0) = K_{Ij}(N,\alpha_s,\alpha_s^0)|_{N=1}, & \text{if} \quad j = T,V, \\ \int_0^1 dx\, x\, K_{Ij}(x,\alpha_s,\alpha_s^0) = K_{Ij}(N,\alpha_s,\alpha_s^0)|_{N=2}, & \text{if} \quad j = \Sigma,g, \end{cases} \tag{4.22}$$

are all finite constants. The convolutions in Eqs. (4.20, 4.21) are evaluated in the same way as those in Eqs. (4.15, 4.16), i.e. with all the kernels pre-computed.
As an illustration, let us consider one of the observables fitted, the neutral current DIS reduced cross section, given by

$$\widetilde{\sigma}^{\mathrm{NC,e}^\pm} = F_2^{NC}(x,Q^2) \mp \frac{Y_-}{Y_+} x F_3^{NC} - \frac{y^2}{Y_+} F_L^{NC}, \tag{4.23}$$

with $Y_\pm = 1 \pm (1-y)^2$. In the above equation F_i^{NC} is given by the sum of the F_i^γ and F_i^Z, whose leading–order expression are defined in Eqs. (1.26, 1.29) respectively. In terms of the PDF evolution eigenstates Eq. (1.58)

$$F_2^{\mathrm{NC}} = x\{E_{\mathrm{S}}^+\Sigma + E_{\mathrm{NS}}^+(T_3 + \frac{1}{3}(T_8 - T_{15}) + \frac{1}{5}(T_{24} - T_{35}))\}, \tag{4.24}$$

$$F_3^{\mathrm{NC}} = E_{\mathrm{S}}^- V + E_{\mathrm{NS}}^-(V_3 + \frac{1}{3}(V_8 - V_{15}) + \frac{1}{5}(V_{24} - V_{35})), \tag{4.25}$$

where the charge coefficients

$$\begin{aligned} E_{\mathrm{S}}^+ &= \frac{5}{18} + \tfrac{1}{2}(B_u + B_d), \qquad E_{\mathrm{NS}}^+ = \frac{1}{6} + \tfrac{1}{2}(B_u - B_d), \\ E_{\mathrm{S}}^- &= \tfrac{1}{2}(D_u + D_d), \qquad E_{\mathrm{NS}}^- = \tfrac{1}{2}(D_u - D_d), \end{aligned} \tag{4.26}$$

are combination of the coefficients B_i and D_i defined in Eq. (1.30). In perturbative QCD the reduced cross sections are written as

$$\begin{aligned} \widetilde{\sigma}^{\mathrm{NC,e}^\pm} = x\{&(C_{2,q}^s - \tfrac{y^2}{Y_+} C_{L,q}^s) \otimes E_{\mathrm{S}}^+\Sigma + E_g(C_{2,g} - \tfrac{y^2}{Y_+} C_{L,g}) \otimes g \\ &+ (C_{2,q} - \tfrac{y^2}{Y_+} C_{L,q}) \otimes (E_{\mathrm{NS}}^+(T_3 + \tfrac{1}{3}(T_8 - T_{15}) + \tfrac{1}{5}(T_{24} - T_{35}))), \\ &\mp \tfrac{Y_-}{Y_+} C_{3,q} \otimes (E_{\mathrm{S}}^- V + E_{\mathrm{NS}}^-(V_3 + \tfrac{1}{3}(V_8 - V_{15}))),\} \end{aligned} \tag{4.27}$$

where $V_{24} = V_{35} = V$, and

$$E_g = \langle e_q^2 \rangle + \langle B_q^2 \rangle, \tag{4.28}$$

$$\langle B_q^2 \rangle = \frac{1}{n_f} \sum_{i=1}^{n_f} B_i = \begin{cases} \frac{1}{3}\left(B_u + 2B_d\right), & \text{if} \quad n_f = 3, \\ \frac{1}{2}\left(B_u + B_d\right), & \text{if} \quad n_f = 4, \\ \frac{1}{5}\left(2B_u + 3B_d\right), & \text{if} \quad n_f = 5, \\ \frac{1}{2}\left(B_u + B_d\right), & \text{if} \quad n_f = 6. \end{cases} \tag{4.29}$$

In terms of hard kernels, the cross section is given by

$$\begin{aligned} \widetilde{\sigma}^{\mathrm{NC},\mathrm{e}^\pm} &= x\{K_{\mathrm{NC},\Sigma} \otimes \Sigma_0 + K_{\mathrm{NC},g} \otimes g_0 + K_{\mathrm{NC},+} \otimes (T_{3,0} + \tfrac{1}{3}(T_{8,0} - T_{15,0})) \\ &\quad \mp K_{\mathrm{NC},V} \otimes V_0 \mp K_{\mathrm{NC},-} \otimes (V_{3,0} + \tfrac{1}{3}(V_{8,0} - V_{15,0}))\}, \end{aligned} \tag{4.30}$$

where in Mellin space

$$\begin{aligned} K_{\mathrm{NC},\Sigma}(N) &= (C_{2,q}^s(N) - \tfrac{y^2}{Y_+} C_{L,q}^s(N)) E_{\mathrm{S}}^+ \Gamma_{\mathrm{S}}^{qq}(N) + E_g(C_{2,g}(N) \\ &\quad - \tfrac{y^2}{Y_+} C_{L,g}(N)) \Gamma_{\mathrm{S}}^{gq}(N) + \tfrac{1}{5}(C_{2,q}(N) \\ &\quad - \tfrac{y^2}{Y_+} C_{L,q}(N)) E_{\mathrm{NS}}^+ (\Gamma_{\mathrm{S}}^{24,q}(N) - \Gamma_{\mathrm{S}}^{35,q}(N)), \\ K_{\mathrm{NC},g}(N) &= (C_{2,q}^s(N) - \tfrac{y^2}{Y_+} C_{L,q}^s(N)) E_{\mathrm{S}}^+ \Gamma_{\mathrm{S}}^{qg}(N) + E_g(C_{2,g}(N) \\ &\quad - \tfrac{y^2}{Y_+} C_{L,g}(N)) \Gamma_{\mathrm{S}}^{gg}(N) + \tfrac{1}{5}(C_{2,q}(N) \\ &\quad - \tfrac{y^2}{Y_+} C_{L,q}(N)) E_{\mathrm{NS}}^+ (\Gamma_{\mathrm{S}}^{24,g}(N) - \Gamma_{\mathrm{S}}^{35,g}(N)), \\ K_{\mathrm{NC},+}(N) &= E_{\mathrm{NS}}^+ (C_{2,q}(N) - \tfrac{y^2}{Y_+} C_{L,q}(N)) \Gamma_{\mathrm{NS}}^+(N), \\ K_{\mathrm{NC},V}(N) &= E_{\mathrm{S}}^- \tfrac{Y_-}{Y_+} C_{3,q}(N) \Gamma_{\mathrm{NS}}^v(N) \\ K_{\mathrm{NC},-}(N) &= E_{\mathrm{NS}}^- \tfrac{Y_-}{Y_+} C_{3,q}(N) \Gamma_{\mathrm{NS}}^-(N). \end{aligned} \tag{4.31}$$

The same procedure has been applied to all other DIS observables included in the fit.

4.1.5 Target Mass Corrections

Among the theoretical errors discussed in Chap. 3, there are higher twists and target mass corrections. The former are related to the power suppressed terms $\mathcal{O}(1/Q^2)$ which are ignored in the leading–twist picture of the factorisation theorem. Physical observables are computed using the leading twist perturbation theory, and higher twist corrections are kept under control by the choice of a relatively high kinematic cut, as discussed in Chap. 4. Target mass corrections (TMCs) instead keep into account the finite mass of the target in a fixed–target experiment, which is explicitly written in Eq. (1.15) in the terms proportional to M_N^2. Often the mass of the nucleon is ignored in the computation of the structure functions, but this approximation is accurate only

for $Q^2 \gg M_N^2$, which is not always true, especially when the mass of the target is large.

Since the corrections due to the finite mass of the target are of purely kinematic origin and can be determined exactly [160], we can easily include them in the computation of DIS observables. One way to implement them, as in Ref. [68], consists in rearranging the target mass correction so that it is explicitly factorised into the hard kernel, and can thus be pre-computed along with the perturbative evolution and coefficient functions.

As an example, let us consider the structure function $F_2(x, Q^2)$. From Eq. (4.19) of Ref. [160], $\widetilde{F}_2$ at twist four is given in terms of the leading twist F_2 by

$$\widetilde{F}_2(\xi, Q^2) = \frac{x^2}{\tau^{3/2}} \frac{F_2(\xi, Q^2)}{\xi^2} + 6\frac{M_N^2}{Q^2}\frac{x^3}{\tau^2} I_2(\xi, Q^2), \tag{4.32}$$

where

$$\tau = 1 + \frac{4M_N^2 x^2}{Q^2}, \qquad \xi = \frac{2x}{1+\sqrt{\tau}}, \tag{4.33}$$

M_N is the mass of the target, and

$$I_2(\xi, Q^2) = \int_\xi^1 \frac{dz}{z^2} F_2(z, Q^2). \tag{4.34}$$

Taking Mellin transforms with respect to ξ we get

$$F_2(\xi, Q^2) = \sum_j \int_C \frac{dN}{2\pi i} \xi^{-N} C_{2,j}(N, \alpha_s) f_j(N, Q^2), \tag{4.35}$$

while

$$\begin{aligned} I_2(N, Q^2) &= \int_0^1 d\xi\, \xi^{N-1} \int_\xi^1 \frac{dz}{z^2} F_2(z, Q^2), \\ &= \left[\frac{\xi^N}{N} \int_\xi^1 dz\, \frac{F_2(z, Q^2)}{z^2}\right]_0^1 + \frac{1}{N}\int_0^1 d\xi\, \xi^{N-2} F_2(\xi, Q^2), \\ &= \frac{1}{N} F_2(N-1, Q^2), \end{aligned}$$

so

$$I_2(\xi, Q^2) = \int_{C+1} \frac{dN}{2\pi i} \frac{\xi^{-N}}{N} F_2(N-1, Q^2) = \frac{1}{\xi}\int_C \frac{dN}{2\pi i} \frac{\xi^{-N}}{N+1} F_2(N, Q^2). \tag{4.36}$$

Now, by substituting Eqs. (4.35, 4.36) into Eq. (4.32) we obtain

$$\widetilde{F}_2(\xi,Q^2) = \int_C \frac{dN}{2\pi i}\,\xi^{-N}\left(\frac{x^2}{\tau^{3/2}\xi^2} + \frac{6M_N^2}{Q^2}\frac{x^3}{\xi\tau^2}\frac{1}{(N+1)}\right)\sum_j C_{2,j}(N,\alpha_s) f_j(N,Q^2). \tag{4.37}$$

We can reinterpret the factor in front of $C_{2,j}(N,\alpha_s)$ as the new target mass corrected coefficient function:

$$\widetilde{C}_{2,j}(N,\alpha_s,\tau) = \frac{(1+\tau^{1/2})^2}{4\tau^{3/2}}\left(1+\frac{3\left(1-\tau^{-1/2}\right)}{N+1}\right)C_{2,j}(N,\alpha_s). \tag{4.38}$$

The target mass corrected hard kernel is then simply

$$\widetilde{K}_{F_2,j}(\xi,\alpha_s,\alpha_s^0) = \sum_k \int_C \frac{dN}{2\pi i}\,\xi^{-N}\widetilde{C}_{2,k}(N,\alpha_s,\tau)\Gamma_{kj}(N,\alpha_s,\alpha_s^0). \tag{4.39}$$

The same procedure may be applied to find the target mass corrections to the F_3 and F_L structure functions. For F_3 we have

$$\widetilde{F}_3(\xi,Q^2) = \frac{x}{\tau}\frac{F_3(\xi,Q^2)}{\xi} + \frac{4M_N^2}{Q^2}\frac{x^2}{\tau^{3/2}}\int_\xi^1 \frac{dz}{z}\,F_3(z,Q^2), \tag{4.40}$$

whence we deduce the target mass corrected coefficient function

$$\widetilde{C}_{3,j}(N,\alpha_s,\tau) = \frac{1+\tau^{1/2}}{2\tau}\left(1\;+\;2\frac{1-\tau^{-1/2}}{N}\right)C_{3,j}(N,\alpha_s), \tag{4.41}$$

and thus $\widetilde{K}_{3,j}(\xi,\alpha_s,\alpha_s^0)$ using an equation analogous to Eq. (4.39). Finally,

$$\widetilde{F}_L(x,Q^2) = F_L(x,Q^2) + \frac{x^2(1-\tau)}{\tau^{3/2}}\frac{F_2(\xi,Q^2)}{\xi^2} + 2\frac{M_N^2}{Q^2}\frac{x^3(3-\tau)}{\tau^2}I_2(\xi,Q^2), \tag{4.42}$$

whence

$$\begin{aligned}\widetilde{C}_L(N,\alpha_s) \;=\;& C_L(N,\alpha_s) + \\ & \frac{(1+\tau^{1/2})^2(1-\tau)}{4\tau^{3/2}}\left(1-\frac{(3-\tau)(1+\tau^{1/2})}{4\tau^2}\frac{1}{N+1}\right) \\ & \times C_2(N,\alpha_s).\end{aligned} \tag{4.43}$$

Note that in the limit $M_N^2/Q^2 \to 0$, $\tau \to 1$, $\xi \to x$, and therefore $\widetilde{C}_{I,j}(N,\alpha_s,\tau) \to C_{I,j}(N,\alpha_s)$, and $\widetilde{K}_{I,j}(\xi,\alpha_s,\alpha_s^0) \to K_{I,j}(x,\alpha_s,\alpha_s^0)$ for each of $I = 2,3,L$.

The implementation of target mass corrections sketched above easily matches the approach based on the Mellin space computation of the hard kernel and the pre–computation of their x–space Mellin transform. The effect of the inclusion of target mass corrections is most enhanced at small Q^2 and large x, as expected, yielding a contribution of 3-4% at most.

4.2 The FastKernel method

The method described earlier requires a specific grid for each different experimental measurement performed at a different value of x. For hadronic observables, which depend on two PDFs, a double convolution must be performed. The main bottleneck of the method is the computation of these convolutions, which might end up being too slow for a parton fit. In the FastKernel method, introduced in Ref. [71], the convolution is sped up by means of the use of interpolating polynomials, thereby leading to both fast evolution and fast computation of all observables for which the kernels have been determined.

The introduction of the FastKernel method in the NNPDF2.0 analysis enabled us to use in the fit an exact computation of the Drell–Yan (DY) process, which in other current global PDF fits [42, 43] is instead treated using a K–factor approximation to the NLO (and even NNLO) result, due to lack of a fast–enough implementation.

Several tools for fast evaluation of hadronic observables have been developed recently, based on an idea of Ref. [161]. These have been implemented for the case of jet production and related observables in the FastNLO framework [162]. More recently, the general–purpose interface APPLGRID based on the same idea has been constructed [163]. Also, the method has been used in the fast x–space DGLAP evolution code HOPPET [159]. The FastKernel method developed in Ref. [71] is based on similar ideas, and it allows for the first time the fast and accurate computation of fixed target Drell–Yan and collider weak boson production cross sections.

In the following section I present a description of the new strategy used to solve the PDF evolution equations in the NNPDF20 analysis [71]. Then I turn to the associated technique to compute DIS structure functions. Finally I discuss how analogous techniques can be used for the fast and accurate computation of hadronic observables. Although the method is completely general, for simplicity I restrict the discussion to the Drell–Yan process, since for inclusive jets FastNLO will be used instead [162]. As far as notation is concerned, in the following I use the index I to denote both the kinematical variables which define an experimental point (x, Q^2) and the type of observable, while in the previous section I was only labelling observables.

4.2.1 Fast PDFs evolution

We have seen that, if Γ_{jk} is the matrix of DGLAP evolution kernels and (x_I, Q_I^2) defines the kinematics of a given experimental point, one may write the PDF evolved from a fixed initial scale Q_0^2 to the scale of the experimental point as a convolution between the evolution kernels and the initial scale PDFs, Eq. (4.3). In the method drawn in Sect. 5.1, the integral in Eq. (4.3) was performed numerically by means of a Gaussian sum on a grid of points distributed between x_I and 1, chosen according to the value of x_I. The point here is to use instead a single grid in x, independent of the x_I value. I label the set of points in the grid as x_α by $\alpha = 1, ..., N_x$, with

$$x_{\min} \equiv x_1 < x_2 < ... < x_{N_x-1} < x_{N_x} \equiv 1.$$

Having chosen a grid of points, we define a set of interpolating functions $\mathcal{I}^{(\alpha)}$ such that:

$$\begin{aligned} \mathcal{I}^{(\alpha)}(x_\alpha) &= 1 \\ \mathcal{I}^{(\alpha)}(x_\beta) &= 0\,, \beta \neq \alpha \\ \sum_{\alpha=1}^{N_x} \mathcal{I}^{(\alpha)}(y) &= 1\,, \forall y. \end{aligned} \tag{4.44}$$

An illustrative example is given by the basis of functions drawn in Fig. 4.2 and defined as

$$E^{(\alpha)}(y) = \frac{y - y_{\alpha-1}}{y_\alpha - y_{\alpha-1}} \theta[(y_\alpha - y)(y - y_{\alpha-1})] + \frac{y_{\alpha+1} - y}{y_{\alpha+1} - y_\alpha} \theta[(y_\alpha - y)(y - y_{\alpha+1})]. \tag{4.45}$$

Each function $E^{(\alpha)}$ has a triangular shape centred in x_α and it vanishes outside the interval $(x_{\alpha-1}, x_{\alpha+1})$. For any y, only two triangular functions are non zero and their sum is always equal to one.

With a general interpolation basis, PDFs at the initial scale can be approximated as

$$f_k(y, Q_0^2) \equiv f_k^0(y) = \sum_{\alpha=1}^{N_x} f_k^0(x_\alpha)\, \mathcal{I}^{(\alpha)}(y) + \mathcal{O}[(x_{\alpha+1} - x_\alpha)^p]\,, \tag{4.46}$$

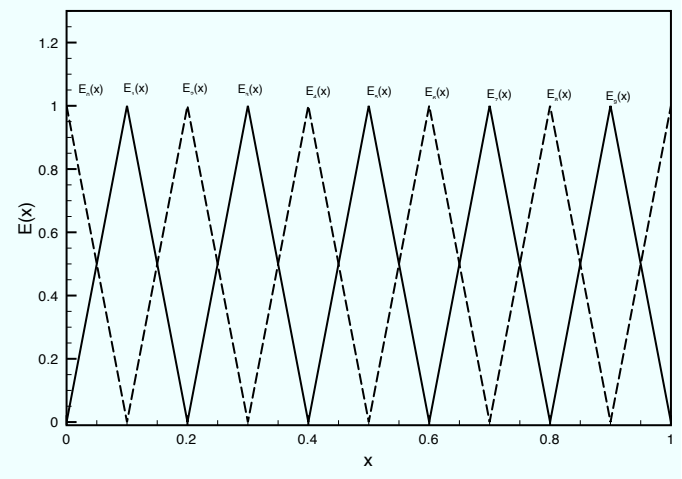


Figure 4.2: Set of interpolating triangular basis functions. The corresponding analytical form is given in Eq. (4.45).

where p is the lowest order neglected in the interpolation. Plugging Eq. (4.46) into Eq. (4.3) and dropping for simplicity the dependence on Q_0^2 and Q_I^2, the latter becomes

$$\begin{aligned} f_j(x_I, Q_I^2) \equiv f_j(x_I) &= \sum_{k=1}^{N_{\rm pdf}} \sum_{\alpha=1}^{N_x} f_k^0(x_\alpha) \int_{x_I}^1 \frac{dy}{y} \Gamma_{jk}\left(\frac{x_I}{y}\right) \mathcal{I}^{(\alpha)}(y) \\ &\quad + \mathcal{O}[(x_{\alpha+1} - x_\alpha)^p] \\ f_j(x_I) &= \sum_{k=1}^{N_{\rm pdf}} \sum_{\alpha=1}^{N_x} \hat{\sigma}_{\alpha k}^{Ij} f_k^0(x_\alpha) + \mathcal{O}[(x_{\alpha+1} - x_\alpha)^p], \end{aligned} \tag{4.47}$$

where

$$\hat{\sigma}_{\alpha k}^{j}(x_I, Q_0^2, Q_I^2) \equiv \hat{\sigma}_{\alpha k}^{Ij} = \int_{x_I}^1 \frac{dy}{y} \Gamma_{jk}\left(\frac{x_I}{y}\right) \mathcal{I}^{(\alpha)}(y). \tag{4.48}$$

In Eq. (4.48) I specifies the data point, α runs over the points in the x–grid and (j,k) run over the PDFs that evolve coupled to each others. Having precomputed the $\hat{\sigma}_{\alpha k}^{Ij}$ coefficients for each point I, the evaluation of the PDFs only requires N_x evaluations of the PDFs at the initial scale, independent of the point at which the evolved PDFs are needed, thereby reducing the computational cost of evolution.

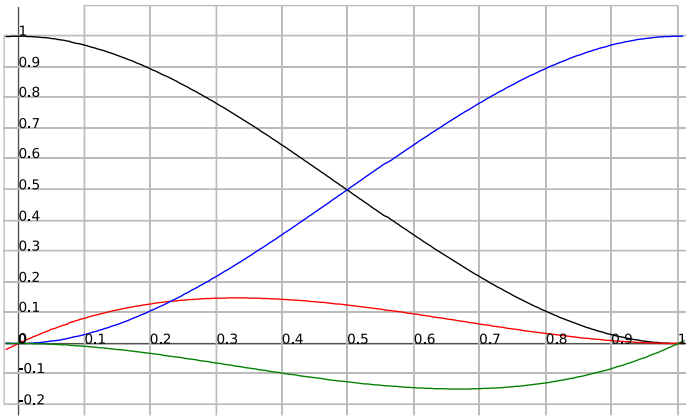


Figure 4.3: Set of interpolating Hermite cubic functions in the [0,1] interval.

If the interpolation is performed on a more complicated set of functions than the triangular basis Fig. 4.2, better accuracy can be obtained with a smaller number of points and thus the computational cost may be further reduced. For PDF evolution, in Ref. [71] the cubic Hermite interpolation drawn in Fig. 4.3 are used. With this choice, for each interval $y \in [x_\alpha, x_{\alpha+1})$ the function to be approximated can be written as

$$\begin{aligned} f_k^0(y) &= h_{00}(t) f_k^0(x_\alpha) + h_{10}(t) h_\alpha m_\alpha + h_{01}(t) f_k^0(x_{\alpha+1}) + h_{11}(t) h_\alpha m_{\alpha+1} \\ &\quad + \mathcal{O}[(x_{\alpha+1} - x_\alpha)^4], \end{aligned}$$

where

$$h_\alpha = g(x_{\alpha+1}) - g(x_\alpha), \qquad t = \frac{g(y) - g(x_\alpha)}{h_\alpha}, \tag{4.49}$$

and $g(y)$ is a monotonic function in [0,1] which determines the distribution of points in the interval (linear, logarithmic, etc.); m_α and $m_{\alpha+1}$ are derivatives of the interpolated function at the right and left–hand side of the interval, which can be defined as finite differences:

$$m_\alpha = \begin{cases} \frac{f_k^0(x_\alpha) - f_k^0(x_{\alpha-1})}{2h_{\alpha-1}} + \frac{f_k^0(x_{\alpha+1}) - f_k^0(x_\alpha)}{2h_\alpha}, & \text{for } 2 \leq \alpha \leq N_x - 1 \\ \frac{f_k^0(x_{\alpha+1}) - f_k^0(x_\alpha)}{h_\alpha}, & \text{for } \alpha = 1 \\ \frac{f_k^0(x_\alpha) - f_k^0(x_{\alpha-1})}{h_{\alpha-1}}, & \text{for } \alpha = N_x. \end{cases} \tag{4.50}$$

Finally the functions h are the $3^{\rm rd}$–order polynomials drawn in Fig. 4.3 and defined as

$$\begin{aligned} h_{00}(t) &= 2t^3 - 3t^2 + 1 = (1+2t)(1-t)^2 \\ h_{10}(t) &= t^3 - 2t^2 + t = t(t-1)^2 \\ h_{01}(t) &= -2t^3 + 3t^2 = t^2(3-2t) \\ h_{11}(t) &= t^3 - t^2 = t^2(t-1) \end{aligned} \tag{4.51}$$

Collecting all terms, Eq. (4.49) becomes

$$\begin{aligned} f_k^0(y) &= f_k^0(x_{\alpha-1})\, A^{(\alpha)}(y) + f_k^0(x_\alpha)\, B^{(\alpha)}(y) + f_k^0(x_{\alpha+1})\, C^{(\alpha)}(y) \\ &\quad + f_k^0(x_{\alpha+2})\, D^{(\alpha)}(y) + \mathcal{O}[(x_{\alpha+1} - x_\alpha)^4]. \end{aligned} \tag{4.52}$$

Hence the function, at any given point y is obtained as a linear combination of f^0 at the four nearest points in the grid. The coefficients of such combination are given by:

$$A^{(\alpha)}(y) = \begin{cases} 0, & \text{for } \alpha = 1 \\ -h_{10}(t)\frac{h_\alpha}{h_{\alpha-1}}, & \text{for } \alpha \neq 1 \end{cases} \tag{4.53}$$

$$B^{(\alpha)}(y) = \begin{cases} h_{00}(t) - h_{10}(t) - \frac{h_{11}(t)}{2}, & \text{for } \alpha = 1 \\ h_{00}(t) - \frac{h_{10}(t)}{2}\left(1 - \frac{h_\alpha}{h_{\alpha+1}}\right) - h_{11}(t), & \text{for } \alpha = N_x - 1 \\ h_{00}(t) - \frac{h_{10}(t)}{2}\left(1 - \frac{h_\alpha}{h_{\alpha+1}}\right) - \frac{h_{11}(t)}{2}, & \text{for } \alpha \neq 1, N_x - 1 \end{cases}$$

$$C^{(\alpha)}(y) = \begin{cases} h_{01}(t) + \frac{h_{11}(t)}{2}\left(1 - \frac{h_\alpha}{h_{\alpha+1}}\right) + h_{10}(t), & \text{for } \alpha = 1 \\ h_{01}(t) + h_{11}(t) + \frac{h_{10}(t)}{2}, & \text{for } \alpha = N_x - 1 \\ h_{01}(t) + \frac{h_{11}(t)}{2}\left(1 - \frac{h_\alpha}{h_{\alpha+1}}\right) + \frac{h_{10}(t)}{2}, & \text{for } \alpha \neq 1, N_x - 1 \end{cases}$$

$$D^{(\alpha)}(y) = \begin{cases} 0, & \text{for } \alpha = N_x - 1 \\ h_{11}(t)\frac{h_\alpha}{2h_{\alpha+1}}, & \text{for } \alpha \neq N_x - 1 \end{cases}$$

Substituting Eq. (4.53) into the integral for the evolution of the PDFs and labelling by ξ the index such that

$$x_\xi \leq x_I < x_{\xi+1},$$

the $\hat{\sigma}$ coefficients are rewritten as:

$$\hat{\sigma}_{\alpha k}^{Ij} = \begin{cases} \int_{x_I}^{x_{\xi+1}} \frac{dy}{y} \Gamma_{jk}\left(\frac{x_I}{y}\right) A^{(\xi)}(y), & \text{for } \alpha = \xi, \\ \int_{x_I}^{x_{\xi+1}} \frac{dy}{y} \Gamma_{jk}\left(\frac{x_I}{y}\right) B^{(\xi)}(y) & \\ \quad +\theta(N_x - (\xi+2)) \int_{x_{\xi+1}}^{x_{\xi+2}} \frac{dy}{y} \Gamma_{jk}\left(\frac{x_I}{y}\right) A^{(\xi+1)}(y), & \text{for } \alpha = \xi+1, \\ \int_{x_I}^{x_{\xi+1}} \frac{dy}{y} \Gamma_{jk}\left(\frac{x_I}{y}\right) C^{(\xi)}(y) & \\ \quad +\theta(N_x - (\xi+2)) \int_{x_{\xi+1}}^{x_{\xi+2}} \frac{dy}{y} \Gamma_{jk}\left(\frac{x_I}{y}\right) B^{(\xi+1)}(y) & \\ \quad +\theta(N_x - (\xi+3)) \int_{x_{\xi+2}}^{x_{\xi+3}} \frac{dy}{y} \Gamma_{jk}\left(\frac{x_I}{y}\right) A^{(\xi+2)}(y), & \text{for } \alpha = \xi+2, \\ \theta(N_x - (I-1)) \int_{x_{\alpha-2}}^{x_{\alpha-1}} \frac{dy}{y} \Gamma_{jk}\left(\frac{x_I}{y}\right) D^{(\alpha-1)}(y) & \\ \quad +\theta(N_x - \alpha) \int_{x_{\alpha-1}}^{x_\alpha} \frac{dy}{y} \Gamma_{jk}\left(\frac{x_I}{y}\right) C^{(\alpha-1)}(y) & \\ \quad +\theta(N_x - (\alpha+1)) \int_{x_\alpha}^{x_{\alpha+1}} \frac{dy}{y} \Gamma_{jk}\left(\frac{x_I}{y}\right) B^{(\alpha)}(y) & \\ \quad +\theta(N_x - (\alpha+2)) \int_{x_{\alpha+1}}^{x_{\alpha+2}} \frac{dy}{y} \Gamma_{jk}\left(\frac{x_I}{y}\right) A^{(\alpha+1)}(y), & \text{for } \xi+3 \leq \alpha \leq N_x+1, \\ 0 & \text{for } \alpha < \xi. \end{cases} \tag{4.54}$$

Despite the complicated bookkeeping, these expressions can be easily pre–computed and input into the fit.

A final remark: because of the divergent behaviour of the x–space evolution kernel in $x = 1$, as we have already seen in Eq. (4.11), the integrals including x_I in the integration interval need to be regularised in $y \sim x_I$. If one considers for instance the first integral of $A^{(\alpha)}$ in Eq.(4.54), one may perform the same subtraction as in Eqs. (4.11) in order to have a finite expression of all precomputed coefficients:

$$\begin{aligned} &\int_{x_I}^{x_{\xi+1}} \frac{dy}{y} \Gamma_{jk}\left(\frac{x_I}{y}\right) A^{(\xi)}(y) \\ &\quad = \int_{x_I}^{x_{\xi+1}} \frac{dy}{y} \Gamma_{jk}\left(\frac{x_I}{y}\right) \left(A^{(\xi)}(y) - \frac{x_I}{y} A^{(\xi)}(x_I)\right) + A^{(\xi)}(x_I) \int_{x_I}^{x_{\xi+1}} \frac{dy}{y^2} \Gamma_{jk}\left(\frac{x_I}{y}\right) \\ &\quad = \int_{x_I}^{x_{\xi+1}} \frac{dy}{y} \Gamma_{jk}\left(\frac{x_I}{y}\right) \left(A^{(\xi)}(y) - \frac{x_I}{y} A^{(\xi)}(x_I)\right) + A^{(\xi)}(x_I) \int_{x_I/x_{\xi+1}}^{1} dz\, \Gamma_{jk}(z) \\ &\quad = \int_{x_I}^{x_{\xi+1}} \frac{dy}{y} \Gamma_{jk}\left(\frac{x_I}{y}\right) \left(A^{(\xi)}(y) - \frac{x_I}{y} A^{(\xi)}(x_I)\right) \\ &\qquad\qquad + A^{(\xi)}(x_I) \left[\Gamma_{jk}(N)\big|_{N=2} - \int_0^{x_I/x_{\xi+1}} dz\, \Gamma_{jk}(z)\right]. \end{aligned} \tag{4.55}$$

As a result all $\hat{\sigma}$ are regularised; they can be stored once and for all for each experimental point, as they do not depend on the PDF at the initial scale.

4.2.2 Accuracy of the fast PDFs evolution

The accuracy of the PDF evolution code described above, has been determined by benchmarking against the Les Houches PDF benchmark tables. In Table 4.2 the relative difference for various combinations of PDFs between our PDF evolution and the

benchmark tables of Ref. [101] at NLO in the ZM–VFNS, for three different grids is shown. In each grid, the interval $[x_{\min}, 1]$ is divided into a logarithmic region at small–x and a linear region at medium–, high–x. The choice of a relatively small grid of 50 points leads to reproducing the Les Houches tables with an accuracy of $\mathcal{O}(10^{-5})$, more than enough for the precision phenomenology one aims to and better than the accuracy displayed in Table 4.1.

4.2.3 Fast computation of DIS observables

Using the strategy described in the previous section, one can easily write down the expressions for the DIS observables included in the fit and show explicitly how their computation works on the interpolation basis. The basic idea is that, starting with the standard factorised expression, the coefficient function C_{Ik} may be absorbed into a modified evolution kernel K_{Ij}, defined in Eq. (4.5). The kernel acts on the j–th PDFs at the initial scale, and it is an observable–dependent linear combination of products of coefficient functions and evolution kernels, as it has been shown in the explicit example of the neutral current reduced cross section, Eq. (4.31). Substituting Eq. (4.47) into the expression for the observable, the latter can be written as:

$$\begin{aligned}\sigma_I^{\rm DIS}(x_I,Q_I^2) &= \sum_{j=1}^{N_{\rm pdf}}\sum_{\alpha=1}^{N_x} f_j^0(x_\alpha)\int_{x_I}^1 \frac{dy}{y}\,K_{Ij}\left(\frac{x_I}{y}\right)\mathcal{I}^{(\alpha)}(y) \qquad (4.56)\\ &= \sum_{j=1}^{N_{\rm pdf}}\sum_{\alpha=1}^{N_x} f_j^0(x_\alpha)\,\hat{\sigma}^I_{\alpha j} + \mathcal{O}[(x_{\alpha+1}-x_\alpha)^p],\end{aligned}$$

where

$$\hat{\sigma}^I_{\alpha j}(x_I,Q_0^2,Q_I^2) \equiv \hat{\sigma}^I_{\alpha j} = \int_{x_I}^1 \frac{dy}{y}\,K_{Ij}\left(\frac{x_I}{y},\alpha_s(Q_I^2),\alpha_s(Q_0^2)\right)\mathcal{I}^{(\alpha)}(y). \quad (4.57)$$

Now the only index running over the PDF basis is j because the other index k is contracted in the definition of K.

Consider for example the expression for the deuteron structure function. We can write down explicitly the terms of Eq. (4.57) as:

$$F_2^d(x_I,Q_I^2) = \sum_{\alpha=1}^{N_x}\sigma^I_{\alpha 10}\,f_{10}(x_\alpha)+\sigma^I_{\alpha 1}\,f_1(x_\alpha)+\sigma^I_{\alpha 2}\,f_2(x_\alpha)+\mathcal{O}[(x_{\alpha+1}-x_\alpha)^p], \quad (4.58)$$

x (30 pts)	$e_{\rm rel}(u_v)$	$e_{\rm rel}(d_v)$	$e_{\rm rel}(\Sigma)$	$e_{\rm rel}(g)$
$1 \cdot 10^{-7}$	$2.5 \cdot 10^{-4}$	$3.5 \cdot 10^{-4}$	$2.1 \cdot 10^{-4}$	$2.1 \cdot 10^{-4}$
$1 \cdot 10^{-6}$	$1.6 \cdot 10^{-3}$	$1.5 \cdot 10^{-3}$	$2.3 \cdot 10^{-4}$	$2.5 \cdot 10^{-4}$
$1 \cdot 10^{-5}$	$1.5 \cdot 10^{-3}$	$1.4 \cdot 10^{-3}$	$2.5 \cdot 10^{-4}$	$2.8 \cdot 10^{-4}$
$1 \cdot 10^{-4}$	$6.5 \cdot 10^{-4}$	$5.1 \cdot 10^{-4}$	$3.0 \cdot 10^{-4}$	$3.4 \cdot 10^{-4}$
$1 \cdot 10^{-3}$	$6.5 \cdot 10^{-4}$	$4.7 \cdot 10^{-4}$	$3.4 \cdot 10^{-4}$	$3.9 \cdot 10^{-4}$
$1 \cdot 10^{-2}$	$1.4 \cdot 10^{-3}$	$1.9 \cdot 10^{-3}$	$3.4 \cdot 10^{-4}$	$5.3 \cdot 10^{-4}$
$1 \cdot 10^{-1}$	$7.0 \cdot 10^{-4}$	$1.0 \cdot 10^{-3}$	$1.1 \cdot 10^{-4}$	$4.1 \cdot 10^{-4}$
$3 \cdot 10^{-1}$	$1.9 \cdot 10^{-5}$	$8.6 \cdot 10^{-5}$	$1.3 \cdot 10^{-5}$	$5.8 \cdot 10^{-5}$
$5 \cdot 10^{-1}$	$1.5 \cdot 10^{-4}$	$1.8 \cdot 10^{-4}$	$1.0 \cdot 10^{-4}$	$1.1 \cdot 10^{-4}$
$7 \cdot 10^{-1}$	$3.8 \cdot 10^{-4}$	$3.9 \cdot 10^{-5}$	$3.1 \cdot 10^{-4}$	$2.8 \cdot 10^{-4}$
$9 \cdot 10^{-1}$	$8.5 \cdot 10^{-3}$	$9.5 \cdot 10^{-2}$	$3.4 \cdot 10^{-3}$	$2.0 \cdot 10^{-2}$

x (50 pts)	$e_{\rm rel}(u_v)$	$e_{\rm rel}(d_v)$	$e_{\rm rel}(\Sigma)$	$e_{\rm rel}(g)$
$1 \cdot 10^{-7}$	$2.1 \cdot 10^{-4}$	$2.3 \cdot 10^{-4}$	$2.7 \cdot 10^{-5}$	$4.7 \cdot 10^{-6}$
$1 \cdot 10^{-6}$	$8.9 \cdot 10^{-5}$	$8.4 \cdot 10^{-5}$	$3.0 \cdot 10^{-5}$	$2.1 \cdot 10^{-5}$
$1 \cdot 10^{-5}$	$9.3 \cdot 10^{-5}$	$6.0 \cdot 10^{-5}$	$2.3 \cdot 10^{-5}$	$2.0 \cdot 10^{-5}$
$1 \cdot 10^{-4}$	$4.5 \cdot 10^{-5}$	$2.8 \cdot 10^{-5}$	$4.4 \cdot 10^{-5}$	$4.2 \cdot 10^{-5}$
$1 \cdot 10^{-3}$	$3.0 \cdot 10^{-5}$	$1.7 \cdot 10^{-5}$	$4.0 \cdot 10^{-5}$	$3.5 \cdot 10^{-5}$
$1 \cdot 10^{-2}$	$7.9 \cdot 10^{-5}$	$6.8 \cdot 10^{-5}$	$4.5 \cdot 10^{-5}$	$5.8 \cdot 10^{-5}$
$1 \cdot 10^{-1}$	$1.7 \cdot 10^{-4}$	$2.1 \cdot 10^{-4}$	$1.6 \cdot 10^{-5}$	$3.9 \cdot 10^{-5}$
$3 \cdot 10^{-1}$	$9.1 \cdot 10^{-6}$	$3.9 \cdot 10^{-5}$	$1.1 \cdot 10^{-5}$	$1.9 \cdot 10^{-7}$
$5 \cdot 10^{-1}$	$2.4 \cdot 10^{-5}$	$2.2 \cdot 10^{-5}$	$2.2 \cdot 10^{-5}$	$2.2 \cdot 10^{-5}$
$7 \cdot 10^{-1}$	$9.1 \cdot 10^{-5}$	$1.5 \cdot 10^{-5}$	$7.8 \cdot 10^{-5}$	$1.2 \cdot 10^{-4}$
$9 \cdot 10^{-1}$	$1.0 \cdot 10^{-3}$	$3.3 \cdot 10^{-3}$	$8.0 \cdot 10^{-4}$	$2.8 \cdot 10^{-3}$

x (100 pts)	$e_{\rm rel}(u_v)$	$e_{\rm rel}(d_v)$	$e_{\rm rel}(\Sigma)$	$e_{\rm rel}(g)$
$1 \cdot 10^{-7}$	$3.2 \cdot 10^{-5}$	$5.0 \cdot 10^{-5}$	$5.4 \cdot 10^{-6}$	$2.0 \cdot 10^{-5}$
$1 \cdot 10^{-6}$	$2.6 \cdot 10^{-6}$	$1.3 \cdot 10^{-6}$	$5.7 \cdot 10^{-6}$	$5.9 \cdot 10^{-6}$
$1 \cdot 10^{-5}$	$1.1 \cdot 10^{-5}$	$2.2 \cdot 10^{-5}$	$3.7 \cdot 10^{-6}$	$1.0 \cdot 10^{-5}$
$1 \cdot 10^{-4}$	$1.8 \cdot 10^{-5}$	$3.3 \cdot 10^{-6}$	$1.3 \cdot 10^{-5}$	$6.9 \cdot 10^{-6}$
$1 \cdot 10^{-3}$	$1.3 \cdot 10^{-6}$	$4.9 \cdot 10^{-6}$	$4.7 \cdot 10^{-6}$	$7.7 \cdot 10^{-6}$
$1 \cdot 10^{-2}$	$1.6 \cdot 10^{-5}$	$1.7 \cdot 10^{-5}$	$4.8 \cdot 10^{-6}$	$1.1 \cdot 10^{-6}$
$1 \cdot 10^{-1}$	$3.4 \cdot 10^{-5}$	$2.9 \cdot 10^{-5}$	$8.7 \cdot 10^{-6}$	$2.1 \cdot 10^{-6}$
$3 \cdot 10^{-1}$	$2.0 \cdot 10^{-6}$	$2.5 \cdot 10^{-5}$	$7.9 \cdot 10^{-6}$	$3.9 \cdot 10^{-6}$
$5 \cdot 10^{-1}$	$1.7 \cdot 10^{-5}$	$1.3 \cdot 10^{-5}$	$1.7 \cdot 10^{-5}$	$3.1 \cdot 10^{-5}$
$7 \cdot 10^{-1}$	$7.1 \cdot 10^{-5}$	$8.3 \cdot 10^{-6}$	$6.3 \cdot 10^{-5}$	$1.3 \cdot 10^{-4}$
$9 \cdot 10^{-1}$	$3.9 \cdot 10^{-5}$	$3.8 \cdot 10^{-4}$	$2.5 \cdot 10^{-5}$	$1.7 \cdot 10^{-3}$

Table 4.2: Relative accuracy of FastKernel evolution compared to the Les Houches benchmark tables for PDFs evolved to the scale $Q^2 = 10^4$ GeV^2. The interpolation is performed on cubic Hermite polynomials and the grid is composed of 30 points (top), 50 points (middle), or 100 points (bottom), distributed logarithmically in the small–x region and linearly in the medium– and large–x region.

with

$$\begin{aligned}
\sigma_{\alpha 10} &= \int_{x_I}^1 \frac{dy}{y}\frac{1}{18}\left(C_{2,q}\otimes\Gamma^{-}\right)\left(\frac{x_I}{y}\right)\mathcal{I}^{(\alpha)}(y)\\
\sigma_{\alpha 1} &= \int_{x_I}^1 \frac{dy}{y}\Big[-\frac{1}{18}\left(C_{2,q}\otimes\Gamma^{15,q}\right)\left(\frac{x_I}{y}\right)+\frac{1}{30}\left(C_{2,q}\otimes\Gamma^{24,q}\right)\left(\frac{x_I}{y}\right)\\
&\quad -\frac{1}{30}\left(C_{2,q}\otimes\Gamma^{35,q}\right)\left(\frac{x_I}{y}\right)+\frac{5}{18}\left(C_{2,q}\otimes\Gamma^{S,qq}\right)\left(\frac{x_I}{y}\right)\\
&\quad -c_g(n_f)\left(C_{2,g}\otimes\Gamma^{S,gq}\right)\left(\frac{x_I}{y}\right)\Big]\mathcal{I}^{(\alpha)}(y)\\
\sigma_{\alpha 2} &= \int_{x_I}^1 \frac{dy}{y}\Big[-\frac{1}{18}\left(C_{2,q}\otimes\Gamma^{15,q}\right)\left(\frac{x_I}{y}\right)+\frac{1}{30}\left(C_{2,q}\otimes\Gamma^{24,g}\right)\left(\frac{x_I}{y}\right)\\
&\quad -\frac{1}{30}\left(C_{2,q}\otimes\Gamma^{35,g}\right)\left(\frac{x_I}{y}\right)+\frac{5}{18}\left(C_{2,q}\otimes\Gamma^{S,qg}\right)\left(\frac{x_I}{y}\right)\\
&\quad -c_g(n_f)\left(C_{2,g}\otimes\Gamma^{S,gg}\right)\left(\frac{x_I}{y}\right)\Big]\mathcal{I}^{(\alpha)}(y)
\end{aligned}\tag{4.59}$$

where all kernels and coefficient functions are defined in Ref. [68] and $f_{10}^0 = T_{8,0}$, $f_1^0 = \Sigma_0$ and $f_2^0 = g_0$, in the evolution basis of Eq. (1.58). The same procedure may be applied to any other DIS observable and similar expressions are obtained.

4.2.4 Fast computation of hadronic observables

The FastKernel implementation of hadronic observables requires a double convolution of the coefficient function with two parton distributions. We could follow the same strategy used for DIS: construct a kernel for each observable and each pair of initial PDFs, and then compute the double convolution with a suitable generalisation of the method introduced in Sect. 4.2.3. However, for hadronic observables, we adopt a somewhat different strategy, which allows us to treat in a more symmetric way processes for which a fast interface already exists (such as jets) and those (such as DY) for which we have to develop our own interface. Namely, instead of including the coefficient function into the kernel according to Eq. (4.57), we compute the convolution Eq. (4.1) using the fast interpolation method.

To see how this works, let us consider first the case of a process with only one parton in the initial state. Starting from Eq. (4.1), we can project the evolved PDF f_k onto an interpolation basis as follows:

$$\begin{aligned}
\sigma_I^{\rm DIS}(x_I,Q_I^2) &= \sum_{k=1}^{N_{\rm pdf}}\sum_{\alpha=1}^{N_y} f_k(y_\alpha,Q_I^2)\int_{x_I}^1\frac{dy}{y}\,C_{Ik}\left(\frac{x_I}{y},\alpha_s(Q_I^2)\right)\mathcal{I}^{\alpha}(y)\\
&+ \qquad \mathcal{O}[(y_{\alpha+1}-y_\alpha)^q],
\end{aligned}\tag{4.60}$$

where q indicates the first order neglected in the interpolation of the evolved PDFs. This defines another grid of points, $\{y_\alpha\}$, upon which the coefficients can be pre–computed before starting the fit:

$$\int_{x_I}^1 \frac{dy}{y} C_{Ik}\left(\frac{x_I}{y}, \alpha_s(Q_I^2)\right) \mathcal{I}^\alpha(y) \equiv C_{Ik}^\alpha. \tag{4.61}$$

If, on top of this interpolation, we interpolate the parton distributions at the initial scale on the $\{x_\alpha\}$ grid as we did in the previous subsection, we get

$$\begin{aligned} \sigma_I^{\rm DIS}(x_I, Q_I^2) &= \sum_{k=1}^{N_{\rm pdf}} \sum_{\alpha=1}^{N_y} f_k(y_\alpha, Q_I^2)\, C_{Ik}^\alpha + \mathcal{O}[(y_{\alpha+1} - y_\alpha)^q] \\ &= \sum_{k,n=1}^{N_{\rm pdf}} \sum_{\alpha=1}^{N_y} \sum_{\beta=1}^{N_x} C_{Ik}^\alpha\, \hat{\sigma}_{\beta kn}^{\alpha,I}\, f_n^0(x_\beta) \\ &+ \qquad \mathcal{O}[(y_{\alpha+1} - y_\alpha)^q (x_{\beta+1} - x_\beta)^p]. \end{aligned} \tag{4.62}$$

Notice that the two interpolations are independent of each other. The number of points N_x and N_y in each grid, the interpolating functions, and the interpolation orders p and q are not necessarily the same.

We now may apply this to the rapidity–differential Drell–Yan cross section to exemplify the procedure. The NLO cross section is given by

$$\begin{aligned} \frac{d\sigma^{\rm DY}}{dQ_I^2 dY_I} &= \frac{4\pi\alpha^2}{9Q_I^2 s} \sum_{j=1}^{N_q} e_j^2 \int_{x_1^0}^1 dx_1 \int_{x_2^0}^1 dx_2 \\ &\Big\{ \left[q_j(x_1, Q_I^2)\bar{q}_j(x_2, Q_I^2) + q_j(x_2, Q_I^2)\bar{q}_j(x_1, Q_I^2)\right] \, (D^{\rm qq}(x_1, x_2, Y_I)) \\ &\quad + g(x_1, Q_I^2)\left[q_j(x_2, Q_I^2) + \bar{q}_j(x_2, Q_I^2)\right] \, (D^{\rm gq}(x_1, x_2, Y_I)) \\ &\quad + g(x_2, Q_I^2)\left[q_j(x_1, Q_I^2) + \bar{q}_j(x_1, Q_I^2)\right] \, (D^{\rm qg}(x_1, x_2, Y_I)) \Big\}, \end{aligned} \tag{4.63}$$

where the coefficient functions can be found in Refs. [164, 165].

For each point of the interpolation grid, we define a set of two–dimensional interpolating functions as the product of one–dimensional functions defined in Eq. (4.44):

$$\mathcal{I}^{(\alpha,\beta)}(x_1, x_2) \equiv \mathcal{I}^{(\alpha)}(x_1)\mathcal{I}^{(\beta)}(x_2). \tag{4.64}$$

The product of two functions can be approximated by means of these interpolating functions as

$$\begin{aligned} f(y_1)h(y_2) &= \sum_{\alpha,\beta=1}^{N_y} f(y_{1,\alpha})h(y_{2,\beta})\,\mathcal{I}^{(\alpha,\beta)}(y_1,y_2) \\ &+ \mathcal{O}[(y_{1,\alpha+1}-y_{1,\alpha})^q(y_{2,\beta+1}-y_{2,\beta})^q]. \end{aligned} \tag{4.65}$$

Applying Eq. (4.65) to the PDFs in Eq. (4.64), we get

$$\begin{aligned} \frac{d\sigma^{\rm DY}}{dQ_I^2 dY_I} &= n(Q_I^2)\sum_{j=1}^{N_q} e_j^2 \sum_{\alpha,\beta=1}^{N_x} \left[q_j(y_{1,\alpha})\bar{q}_j(y_{2,\beta}) + \bar{q}_j(y_{1,\alpha})q_j(y_{2,\beta})\right] \\ &\int_{x_1^0}^1 dx_1 \int_{x_2^0}^1 dx_2 \mathcal{I}^{(\alpha,\beta)}(x_1,x_2)\, D^{\rm qq}(x_1,x_2,Y_I) \\ &+\left[g(y_{1,\alpha})(q_j(y_{2,\beta})+\bar{q}_j(y_{2,\beta}))\right] \int_{x_1^0}^1 dx_1 \int_{x_2^0}^1 dx_2 \mathcal{I}^{(\alpha,\beta)}(x_1,x_2)\, D^{\rm gq}(x_2,x_1,Y_I) \\ &+\left[g(y_{1,\alpha})(q_j(y_{2,\beta})+\bar{q}_j(y_{2,\beta}))\right] \int_{x_1^0}^1 dx_1 \int_{x_2^0}^1 dx_2 \mathcal{I}^{(\alpha,\beta)}(x_1,x_2)\, D^{\rm gq}(x_1,x_2,Y_I) \\ &+\mathcal{O}[(y_{1,\alpha+1}-y_{1,\alpha})^q(y_{2,\beta+1}-y_{2,\beta})^q], \end{aligned} \tag{4.66}$$

where at next–to–leading order $D^{\rm qg}(x_1,x_2,Y_I) = D^{\rm gq}(x_2,x_1,Y_I)$. Therefore, we can define

$$C_{I,ij}^{(\alpha,\beta)} \equiv \int_{x_1^0}^1 dx_1 \int_{x_2^0}^1 dx_2 \mathcal{I}^{(\alpha,\beta)}(x_1,x_2)\, D^{\rm ij}(x_1,x_2,Y_I), \tag{4.67}$$

where i,j run over the non–zero combinations of $q,\bar{q}$ and g. By substituting them into Eq. (4.67), we end up with the expression

$$\begin{aligned} \frac{d\sigma^{\rm DY}}{dQ_I^2 dY_I} &= n(Q_I^2)\sum_{j=1}^{N_q} e_j^2 \sum_{\alpha,\beta=1}^{N_y} C_{I,qq}^{(\alpha,\beta)} \left[q_j(y_{1,\alpha})\bar{q}_j(y_{2,\beta}) + \bar{q}_j(y_{1,\alpha})q_j(y_{2,\beta})\right] \\ &+C_{I,gq}^{(\alpha,\beta)} \left[g(y_{1,\alpha})(q_j(y_{2,\beta})+\bar{q}_j(y_{2,\beta}))\right] \\ &+C_{I,qg}^{(\alpha,\beta)} \left[(q_j(y_{1,\alpha})+\bar{q}_j(y_{1,\alpha}))g(y_{2,\beta})\right] \\ &+\mathcal{O}[(y_{1,\alpha+1}-y_{1,\alpha})^q(y_{2,\beta+1}-y_{2,\beta})^q]\,, \end{aligned} \tag{4.68}$$

which is the analogue of Eq. (4.60) for a hadronic observable. The physical basis $\{q\}_j$ and the evolution basis $\{f\}_j$ are related by a matrix A:

$$q_j = A_{jr} f_r \qquad \bar{q}_j = \bar{A}_{js} f_s \,.$$

Each PDF f is evolved at the physical scale of the process, and the evolution matrix Γ which relates the initial scale PDFs to the evolved ones is

$$f_r(x, Q^2) = \Gamma_{rn}(x, Q_0^2, Q^2) \otimes f_n(x, Q_0^2).$$

Therefore Eq. (4.68) becomes

$$\begin{aligned} \frac{d\sigma^{\rm DY}}{dQ_I^2 dY_I} &= n(Q_I^2) \sum_{j=1}^{N_q} e_j^2 \sum_{\alpha,\beta=1}^{N_y} \sum_{r,s=1}^{N_{\rm pdf}} C_{I,qq}^{(\alpha,\beta)} (A_{jr}\bar{A}_{js} + \bar{A}_{jr}A_{js})\, f_r(y_{1,\alpha}) f_s(y_{2,\beta}) \\ &\quad + \left[C_{I,qg}^{(\alpha,\beta)} \delta_{r2} (A_{js} + \bar{A}_{js}) + C_{I,qg}^{(\alpha,\beta)} (A_{jr} + \bar{A}_{js}) \delta_{q2} \right] \\ &\quad \times f_r(y_{1,\alpha}) f_s(y_{2,\beta}) + \mathcal{O}[(y_{1,\alpha+1} - y_{1,\alpha})^q (y_{2,\beta+1} - y_{2,\beta})^q]. \end{aligned} \tag{4.69}$$

Defining

$$\begin{aligned} c_{rs} &\equiv \sum_{j=1}^{N_q} e_j^2 (A_{jr}\bar{A}_{js} + \bar{A}_{jr}A_{js}) \\ d_{rs} &\equiv \sum_{j=1}^{N_q} e_j^2 \left[\delta_{r2}(A_{js} + \bar{A}_{js}) + (A_{jr} + \bar{A}_{jr})\delta_{s2} \right] \end{aligned} \tag{4.70}$$

and applying Eq. (4.47) to the evolved PDFs, we end up with a result which is similar to Eq. (4.63):

$$\begin{aligned} \frac{d\sigma^{\rm DY}}{dQ_I^2 dY_I} &= n(Q_I^2) \sum_{\gamma,\delta=1}^{N_x} \sum_{l,m=1}^{N_{\rm pdf}} \Bigg[\sum_{\alpha,\beta=1}^{N_y} \sum_{r,s=1}^{N_{\rm pdf}} c_{rs} C_{I,qq}^{(\alpha,\beta)} \hat{\sigma}_{\gamma rl}^{\alpha,I} \hat{\sigma}_{\delta sm}^{\beta,I} \\ &\quad + [d_{rs} C_{I,qg}^{(\alpha,\beta)} + d_{sr} C_{I,qg}^{(\alpha,\beta)}] \hat{\sigma}_{\gamma rl}^{\alpha,I} \hat{\sigma}_{\delta sm}^{\beta,I} \Bigg] f_l^{(0)}(x_{1,\gamma}) f_m^{(0)}(x_{2,\delta}) \\ &\quad + \mathcal{O}[(y_{1,\alpha+1} - y_{1,\alpha})^q (y_{2,\beta+1} - y_{2,\beta})^q (x_{1,\gamma+1} - x_{1,\gamma})^p (x_{2,\delta+1} - x_{2,\delta})^p]. \end{aligned} \tag{4.71}$$

In order to define the coefficients in Eq. (4.71), we have to make an explicit choice of an interpolating basis. For the interpolation of the evolved PDFs we use the triangular interpolating basis drawn in Fig. 4.2 and defined in Eq. (4.45). Projecting the PDFs

on the triangular basis,we get

$$q(y) = \sum_{\alpha=1}^{N_x} q(y_\alpha)\, E^{(\alpha)}(y) + \mathcal{O}[(y_{\alpha+1} - y_\alpha)^2]$$

and define

$$C_{K,ij}^{(\alpha,\beta)} = \int_{x_1^0}^1 dx_1 \int_{x_2^0}^1 dx_2\, E^{(\alpha)}(x_1)\, E^{(\beta)}(x_2) D_{ij}^{(K)}(x_1, x_2), \tag{4.72}$$

where K indicates the perturbative order and i, j run over the non–zero combinations of $q, \bar{q}$ and g. To be more explicit, defining the index ξ and ζ in such a way that

$$x_\xi < x_1^0 < x_{\xi+1} \qquad x_\zeta < x_2^0 < x_{\zeta+1}, \tag{4.73}$$

we can give the precise definition of the NLO coefficients:

$$C_{K,ij}^{(\alpha,\beta)} = \begin{cases} \int_{x_1^0}^{x_{\alpha+1}} dx_1 \int_{x_2^0}^{x_{\beta+1}} dx_2\, E^{(\alpha)}(x_1)\, E^{(\beta)}(x_2) D_{ij}^{(K)}(x_1, x_2), & \alpha = \xi, \xi+1, \beta = \zeta, \zeta+1 \\ \int_{x_1^0}^{x_{\alpha+1}} dx_1 \int_{x_{\beta-1}^{x_{\beta+1}}} dx_2\, E^{(\alpha)}(x_1)\, E^{(\beta)}(x_2) D_{ij}^{(K)}(x_1, x_2), & \alpha \leq \xi+1,\ \beta \geq \zeta+2, \\ \int_{x_{\alpha-1}}^{x_{\alpha+1}} dx_1 \int_{x_2^0}^{x_{\beta+1}} dx_2\, E^{(\alpha)}(x_1)\, E^{(\beta)}(x_2) D_{ij}^{(K)}(x_1, x_2), & \alpha \geq \xi+2,\ \beta \leq \zeta+1, \\ \int_{x_{\alpha-1}}^{x_{\alpha+1}} dx_1 \int_{x_{\beta-1}}^{x_{\beta+1}} dx_2\, E^{(\alpha)}(x_1)\, E^{(\beta)}(x_2) D_{ij}^{(K)}(x_1, x_2), & \alpha \geq \xi+2,\ \beta \geq \zeta+2, \\ 0 & \alpha \leq \xi-1,\ \beta \leq \zeta-1. \end{cases}$$

The expression for the LO is trivial, given that $D_{q\bar{q}}^{(0)}(x_1, x_2) = \delta(x_1 - x_1^0)\delta(x_2 - x_2^0)$.

The FastKernel method for hadronic observables is easily interfaced to other existing fast codes, such as FastNLO for inclusive jets [162], by simply using FastKernel for the interpolation at the initial scale and parton evolution, and exploiting the existing interface for the convolution of the evolved PDF with the appropriate coefficient functions. In the particular case of the inclusive jet measurements used in the present analysis, the analogues of the coefficients $C_{I,ij}^{(\alpha,\beta)}$ in Eq. (4.71) can be directly extracted from the FastNLO precomputed tables through its interface, although in such case the relevant PDFs combinations are different than those of the DY process Eq. (4.68).

4.2.5 FastKernel benchmarking

It is straightforward to extend the FastKernel method described in the previous section to all fixed–target DY and collider vector boson production datasets using the appropriate couplings and PDF combinations.

In order to assess the accuracy of the method, I have benchmarked the results obtained with our code to those produced by an independent code [166] which computes the exact NLO cross sections for all relevant Drell–Yan distributions. The comparison is performed by using a given set of input PDFs and evaluating the various cross–sections for all observables included in the NNPDF20 [71] fit in the kinematical points which correspond to the included data.

The benchmarking of the FastKernel code for the Drell–Yan process has been performed for the following observables:

- The rapidity and x_F distributions and asymmetries for fixed target Drell–Yan in pp and pCu collisions (E605 and E866 kinematics)
- The W rapidity distribution and asymmetries at hadron colliders (Tevatron kinematics)
- The Z rapidity distribution at hadron colliders (Tevatron kinematics)

The results of this benchmark comparison are displayed in Fig. 4.4, where the relative accuracy between the FastKernel implementation and the exact code is shown for all data points included in the NNPDF2.0. This accuracy has been obtained with a grid of 100 points distributed as the squared root of the logarithm from $x_{\min}$ to 1.

It is clear from Fig. 4.4 that with a linear interpolation performed on a 100–points grid, we get a reasonable accuracy for all points, 1% in the worse case, which is suitable because the experimental uncertainties of the available datasets are rather larger. This accuracy can be improved arbitrarily by increasing the number of data points in the grid, with a very small cost in terms of speed: this is demonstrated in Fig. 4.5, where we show the improvement in accuracy obtained by using a grid of 500 points.

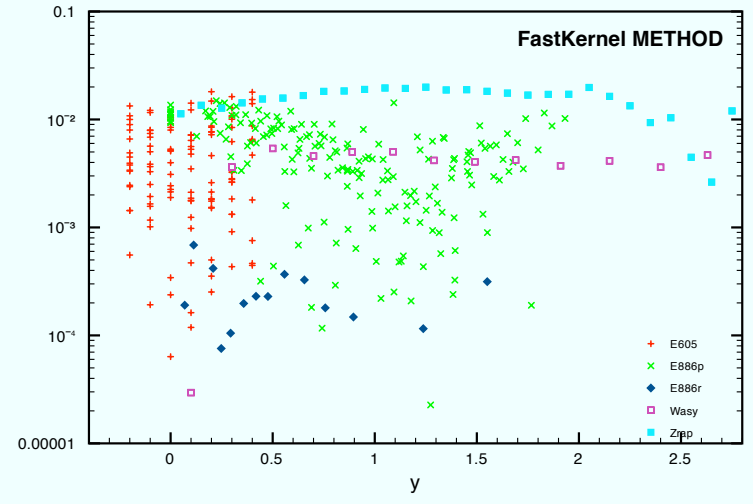


Figure 4.4: Relative accuracy for NLO Drell–Yan rapidity distributions using the FastKernel method, compared to the code of [166], as a function of rapidity y. Each point corresponds to the kinematics of a data point included in the NNPDF2.0 fit. The accuracy refers to a grid of 100 points distributed as the squared root of the logarithm from $x_{\min}$ to 1.

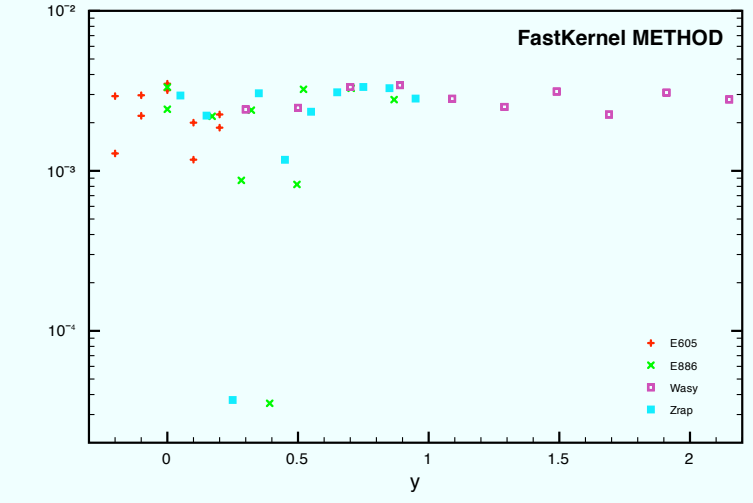


Figure 4.5: Same as Fig. 4.4, for 40 points in the kinematical range covered by the data points included in the NNPDF2.0 fit, using a grid of 500 points distributed as the squared root of the logarithm from $x_{\min}$ to 1.

Chapter 5

Application of the NNPDF method to phenomenology

In this chapter I illustrate three phenomenological studies carried out by exploiting the NNPDF method. First I present the determination of the strange content of the proton formulated in the NNPDF1.2 analysis. With PDF uncertainties under control, detailed precision physics studies become possible: the NNPDF1.2 parton analysis leads at the same time to the solution of the NuTeV anomaly and to the precise determination of the $|V_{cs}|$ and $|V_{cd}|$ elements of the CKM matrix. Secondly, I discuss the effect of the combined PDF and strong coupling uncertainties and their impact on the standard candle processes at the LHC. Finally I introduce a reweighting technique which allows us to include new experiments into the NNPDF analyses without need of refitting; in particular this method is applied to the inclusion of W lepton asymmetry data.

5.1 The strange content of the proton

The determination of the strange and anti–strange quark distributions of the nucleon is of considerable phenomenological interest, because many final states in the Standard Model and beyond couple directly to strangeness. For instance the determination of the electroweak mixing angle by the NuTeV collaboration might provide evidence for physics beyond the Standard Model and is very sensitive to the strange content of the nucleon [167, 168].

However, neutral-current deep-inelastic scattering, which constitutes the bulk of the data which is used for parton determination, have minimal sensitivity to flavour sep-

aration, and no sensitivity at all to the separation of quark and antiquark contributions. As a consequence, until very recently in parton fits, such as CTEQ6.5 [41], MRST2006 [66] and NNPDF1.0 [68], the strange and anti–strange quark distributions were assumed to be equal and proportional to the total light antiquark sea.

This situation has recently changed, due to the availability of a wider set of inclusive neutrino DIS scattering data [169, 88] and, more importantly, of data for deep-inelastic neutrino and anti-neutrino production of charm [89, 170, 171] ("dimuon" data), which is directly sensitive to the strange and anti–strange parton distributions. Indeed, from Eq. (3.38) we see that the measured dimuon cross–section is proportional to the charm production cross–section. The latter is given by

$$\begin{aligned}\tilde{\sigma}^{\nu(\bar{\nu}),c}(x,y,Q^2) &\equiv \frac{1}{E_\nu}\frac{d^2\sigma^{\nu(\bar{\nu}),c}}{dx\,dy}(x,y,Q^2) \qquad (5.1)\\ &= \frac{G_F^2 M_N}{2\pi(1+Q^2/M_W^2)^2}\left[\left(\left(Y_+ - \frac{2M_N^2x^2y^2}{Q^2} - y^2\right)\left(1+\frac{m_c^2}{Q^2}\right)+y^2\right)\right.\\ &\left.\times F_2^{\nu(\bar{\nu}),c}(x,Q^2) - y^2 F_L^{\nu(\bar{\nu}),c}(x,Q^2) \pm Y_-\, xF_3^{\nu(\bar{\nu}),c}(x,Q^2)\right],\end{aligned}$$

where

$$Q^2 = 2M_N E_\nu xy, \qquad Y_\pm = 1 \pm (1-y)^2. \qquad (5.2)$$

We refer to Chap. 2 and 4 for the definition of all kinematical quantities involved in the above equation. The charm production cross–section is determined by the charm structure functions $F_2^{\nu(\bar{\nu}),c}$,$F_L^{\nu(\bar{\nu}),c}$ and $xF_3^{\nu(\bar{\nu}),c}$, which in the quark model are given by

$$\begin{aligned}F_2^{\nu,c}(x,Q^2) &= xF_3^{\nu,c}(x,Q^2) \qquad (5.3)\\ &= 2x\left(|V_{cd}|^2\, d(x) + |V_{cs}|^2\, s(x) + |V_{cb}|^2\, b(x)\right), \qquad (5.4)\\ F_2^{\bar{\nu},c}(x,Q^2) &= -xF_3^{\bar{\nu},c}(x,Q^2)\\ &= 2x\left(|V_{cd}|^2\, \bar{d}(x) + |V_{cs}|^2\, \bar{s}(x) + |V_{cb}|^2\, \bar{b}(x)\right),\end{aligned}$$

with $F_L^{\nu(\bar{\nu}),c} = 0$. It is clear from the above equation that the dimuon cross–section provides a direct constraint on the strange and anti–strange distributions. It may be explicitly shown that the same constraint, even if weaker, is given by the charged–current observables [68].

As a consequence, dedicated analyses of the strange quark distribution have been performed [172, 173, 62, 174], and independent parametrisations of the strange and anti–

strange distributions have been included in most recent parton fits [43]. However, the method of parton determination used in these analyses, based on fitting the parameters of a fixed functional form, is known to be hard to handle when experiments are relatively unconstraining. Indeed, it is not uncommon that the addition of new experimental information to a parton fit of this kind, actually leads to an increase rather than to a decrease of uncertainty bands , because the new data require the use of a more general parametrisation. For this reason in most parton fits [43, 42] the strange is parametrised by a very restrictive functional form, which might bias the result and artificially reduce its uncertainty. In conclusion, the scarceness of the experimental information on these quantities, makes it difficult to separate the genuine information from theoretical bias.

The methodology developed by the NNPDF collaboration is particularly able to deal with this kind of issues. This methodology is largely free of bias related to parton parametrisation, and handles in a satisfactory way incomplete information, contradictory data, and the addition of new data within a single framework. Indeed, in the NNPDF1.1 analysis [69], 74 free parameters were introduced to parametrise the strange and anti–strange parton distributions and, even if they were basically unconstrained by data, results are statistically consistent.

In the NNPDF1.2 parton determination [70], we added the dimuon data to the global deep-inelastic scattering dataset on which the NNPDF1.0 and NNPDF1.1 fits were based, and constructed a new parton set including a determination of the strange and anti–strange distributions. In Fig. 5.1 and 5.2 the $s^{\pm}(x,Q_0^2)$, $s(x,Q_0^2)$ and $\bar{s}(x,Q_0^2)$ strange PDFs are shown at the input scale and compared to the most recent CTEQ6.6 [42] and MSTW08 [43] sets, as well as to the NNPDF1.0 set [68]. Whereas the CTEQ collaboration has not performed a full determination of the s^- uncertainty band, a study of the dependence of the best-fit s^- on assumptions on its functional form was performed in Ref. [62]. We see that in the data region $x \gtrsim 0.03$ all determinations of $s^{\pm}$ agree, however the NNPDF1.2 has a much larger uncertainty than other existing determinations.

The origin of this can be understood by looking at Fig. 5.3, where 25 randomly chosen replicas out of the full NNPDF1.2 set are displayed, and the mean and standard deviation computed from them: the large uncertainty is a consequence of the great flexibility afforded by the neural network parametrisation. This is particularly noticeable in the case of s^-, which must have at least one node because of the sum rule Eq. (3.31): individual replicas cross the x–axis in different places, with different sign (from positive to negative or conversely), and some replicas have more than one crossing. It is interesting to observe that the "neck" in the uncertainty on s^- around

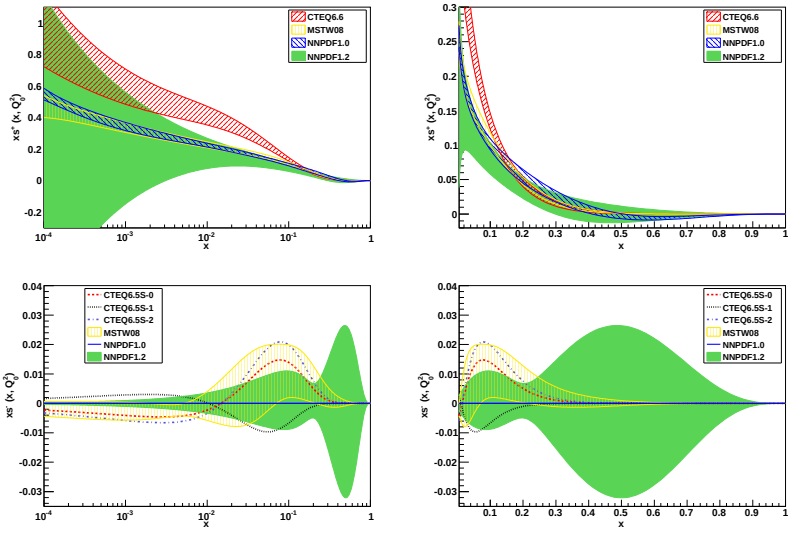


Figure 5.1: From top to bottom, the strange C-even and C-odd combinations $s^+(x,Q_0^2)$, $s^-(x,Q_0^2)$ PDFs, plotted at the initial scale $Q_0^2 = 2$ GeV^2 versus x on a on a log (left) or linear(right) scale, computed from the NNPDF1.2 set of $N_{\rm rep} = 1000$ replicas. The NNPDF1.2 result is compared to the MSTW08 [43] and CTEQ6.6 [42] global fits. For s^- some of the results from obtained the CTEQ6.5s strangeness series [62] are also shown.

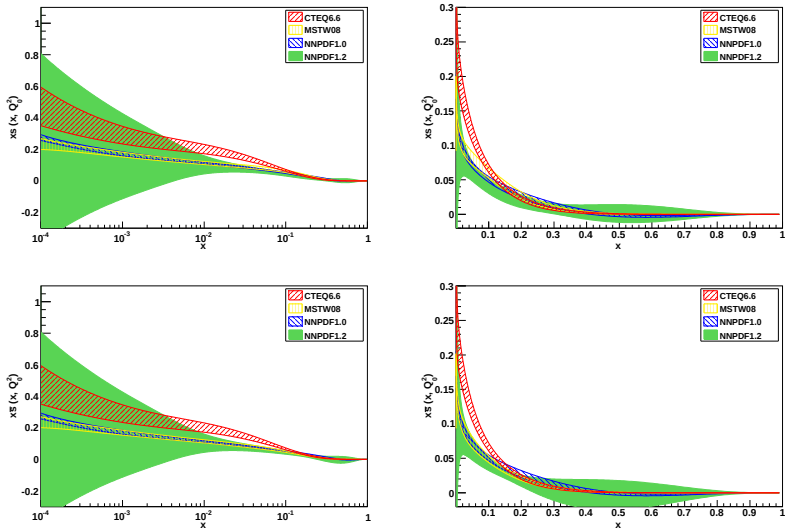


Figure 5.2: Same as Fig. 5.1 with strange $s(x,Q_0^2)$ (top) and anti–strange $\bar{s}(x,Q_0^2)$ (bottom) PDFs plotted at the initial scale $Q_0^2 = 2$ GeV^2.

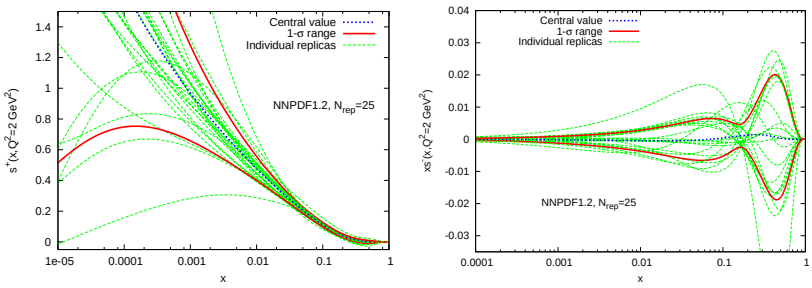


Figure 5.3: A set of randomly chosen $N_{\rm rep} = 25$ replicas of the strange PDFs $s^+(x,Q_0^2)$ (left), $s^-(x,Q_0^2)$ (right) out of the full set of Fig. 5.1, and the PDFs computed from them.

$x \approx 0.1$ corresponds to the value of x at which the crossing is most likely to occur. The role played by the valence sum rule Eq. (3.31) in determined these features of the strangeness asymmetry s^- can be elucidated by repeating the fit without imposing it. The results, displayed in Fig. 5.4, show that, even without the sum rule constraint, many replicas still cross the x–axis.

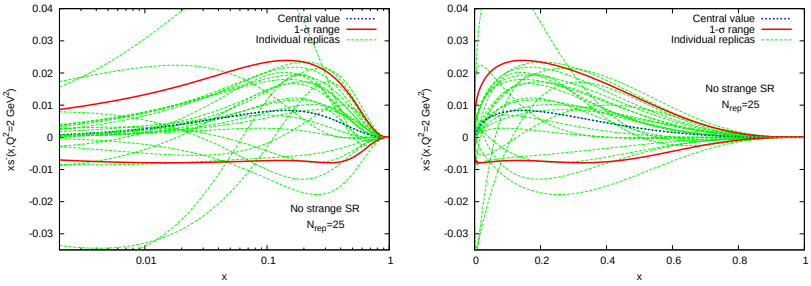


Figure 5.4: A set of randomly chosen $N_{\rm rep} = 25$ replicas of the strange valence $s^-(x,Q_0^2)$ out of the full set of Fig. 5.1, and the PDFs computed from them. Here sum rule Eq. (3.31) is not imposed during the fitting procedure. Both small–x (left) and large–x (right) regions are shown.

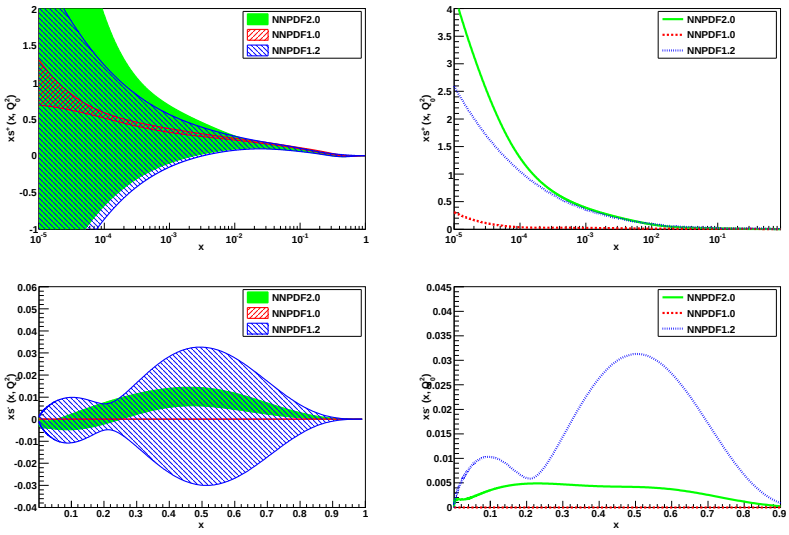


Figure 5.5: $xs^+(x,Q_0^2)$ (top), $xs^-(x,Q_0^2)$ (bottom) PDFs combinations (left) and their absolute uncertainties (right), plotted at the input scale $Q_0^2 = 2$ GeV^2 versus x, computed from the NNPDF1.2 set of $N_{\rm rep} = 1000$ replicas. The NNPDF2.0 [71], NNPDF1.2 [70] and NNPDF1.0 [68] distributions and their absolute errors are compared.

In the NNPDF2.0 analysis, the Drell-Yan data have been added in the fit. They provide a strong constraint to the valence–type PDFs, as it was shown in Chap. 4. On the other hand, the new parton distributions remain consistent with those determined in the previous analysis. The same holds for the $s^{\pm}$, as it is shown in Fig. 5.5.

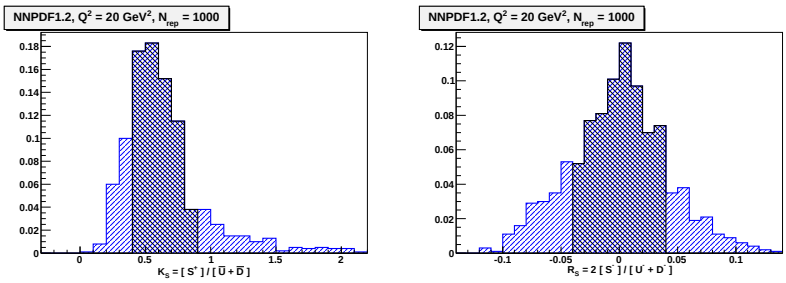


Figure 5.6: Probability distribution of K_S (left) and R_S (right) at $Q^2 = 20$ GeV^2 computed from the reference set of $N_{\rm rep} = 1000$ NNPDF1.2 PDF replicas. The region filled in black corresponds to the central 68% confidence interval, $K_S\left(Q^2 = 20\,{\rm GeV}^2\right) = 0.71^{+0.19\,{\rm stat}}_{-0.31}$ and $R_S\left(Q^2 = 20\,{\rm GeV}^2\right) = 0.006 \pm 0.045^{\rm stat}$.

This might have a strong impact in all analyses presented in Ref. [70]. For this reason, in what follows, I refer to the NNPDF1.2 results and, when available, I show how they are modified by using the NNPDF2.0 parton set as an input distribution.

5.1.1 Implications for LHC observables

The main interesting features of the strange distributions for phenomenological applications are the momentum fractions, defined as

$$\left[S^{\pm}\right](Q^2) \equiv \int_0^1 dx\, x\, s^{\pm}(x,Q^2)\,, \tag{5.5}$$

with similar definitions for moments of other PDF combinations, and in particular their ratio to the light sea and the light valence momentum fractions:

$$K_S(Q^2) \equiv \frac{\int_0^1 dx\; x\; s^+\left(x,Q^2\right)}{\int_0^1 dx\; x\left(\bar{u}\left(x,Q^2\right)+\bar{d}\left(x,Q^2\right)\right)} = \frac{[S^+]}{\left[\bar{U}+\bar{D}\right]}\,, \tag{5.6}$$

$$R_S(Q^2) \equiv 2\frac{\int_0^1 dx x s^-(x,Q^2)}{\int_0^1 dx x\left(u^-(x,Q^2)+d^-(x,Q^2)\right)} = 2\frac{[S^-]}{[U^-+D^-]}\,. \tag{5.7}$$

In parton fits where the strange content of the proton is determined by flavor assumptions, these quantities are assumed to be fixed at the starting scale($K_S(Q_0^2) \approx 0.5$, $R_S(Q_0^2) = 0$).

The value and uncertainty on these quantities can be determined from the NNPDF1.2 set by performing averages over PDFs replicas. The probability distribution of K_S at $Q^2 = 20$ GeV^2 is shown in Fig. 5.6, and turns out to be quite far from Gaussian. This is not unexpected, given that the denominator in Eq. (5.6) can assume rather small values. Therefore, we compute the 1σ uncertainty as a central 68% confidence integral, namely requiring the two outer tails of the probability distribution (lighter blue region in Fig. 5.6) to correspond to 16% probability each, with the central value still given by the average. The median of the probability distribution is equal to $K_S^{\rm med} = 0.59$, significantly different from the average because of the asymmetry.

In the case of the strange momentum asymmetry R_S the denominator is fixed by knowledge of the valence content of the nucleon, which is known quite accurately: hence we expect the uncertainty to be symmetric and dominated by the uncertainty of the numerator. Indeed, as we see in Fig. 5.6 the probability distribution for R_S turns out to be approximately Gaussian so that the uncertainty computed from its central 68% confidence essentially coincides with the standard deviation of the distribution, while its central value and uncertainty are essentially proportional to those of the strangeness asymmetry $[S^-]$.

In Ref. [71] the addition of fixed–target Drell–Yan data lead to significantly stricter constraints on the shape of the strange distributions $s^{\pm}(x)$, as it is shown in Fig. 5.5: the new determination of s^+ and especially s^- at large–x have a much reduced uncertainty in the NNPDF2.0 analysis. As a consequence, the strange momentum fraction and strangeness asymmetry at $Q^2 = 20$ GeV^2 are

$$K_S = \begin{cases} 0.71^{+0.19}_{-0.31}{}^{\rm stat} \pm 0.26^{\rm syst} & ({\rm NNPDF1.2}) \\ 0.503 \pm 0.075^{\rm stat}; & ({\rm NNPDF2.0}) \end{cases} \tag{5.8}$$

$$R_S = \begin{cases} 0.006 \pm 0.045^{\rm stat} \pm 0.010^{\rm syst} & ({\rm NNPDF1.2}) \\ 0.019 \pm 0.008^{\rm stat} & ({\rm NNPDF2.0}), \end{cases} \tag{5.9}$$

i.e. the PDF uncertainty on K_S is reduced by more than a factor two, while that on R_S is reduced by a factor 5, with all results consistent within uncertainties. The systematic uncertainty is determined by keeping into account several factors like the effect of the heavy quark masses, which in the NNPDF1.2 analysis is included only in the computation of the dimuon cross–section by using the Improved ZM-VFN scheme. On top of that we considered the effect of nuclear corrections, estimated by repeating the fit with CHORUS and NuTeV data corrected using the de Florian-Sassot [175] and HKN07 [176] models. Finally we considered the effect of the momentum sum rules by repeating a fit without including them. As it is show in Tab. 5.1, the effect of any of these systematics is rather moderate, even if very conservatively estimated.

	K_S (mean)	R_S
Reference	$0.71^{+0.19}_{-0.31}$	$(6 \pm 45) \cdot 10^{-3}$
ZM-VFN	$0.47^{+0.10}_{-0.20}$	$(8 \pm 39) \cdot 10^{-3}$
Nuclear - dFS03	$0.74^{+0.21}_{-0.40}$	$(12 \pm 48) \cdot 10^{-3}$
Nuclear - HKN07	$0.68^{+0.24}_{-0.29}$	$(0 \pm 40) \cdot 10^{-3}$
LO	$0.61^{+0.33}_{-0.22}$	$(1 \pm 38) \cdot 10^{-3}$
No strange SR	$0.62^{+0.20}_{-0.21}$	$(17 \pm 32) \cdot 10^{-3}$

Table 5.1: The strange relative total and valence momentum fractions K_S and R_S, Eqs. (5.6, 5.7), at the scale $Q^2 = 20$ GeV^2. The first row gives the value computed from the reference NNPDF1.2 set of $N_{\rm rep} = 1000$ replicas, while the other rows give results from sets of $N_{\rm rep} = 100$ replicas each obtained from alternative fits discussed in text. All uncertainties are 1σ or 68% central confidence intervals.

Moreover even in a fit in which the sum rule Eq. (3.31) is not imposed the result changes very little. In the NNPDF2.0 analysis we have made no attempt to provide a new determination of systematic theoretical uncertainties on K_S and R_S by assuming that they should be similar to those determined in Ref. [70].

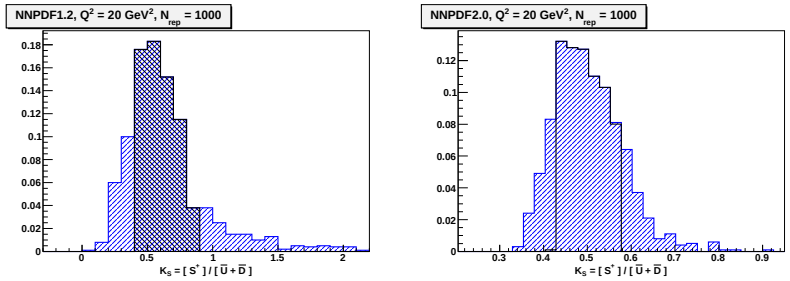


Figure 5.7: Distribution of K_S at $Q^2 = 20$ GeV^2 computed from the reference set of $N_{\rm rep} = 1000$ NNPDF1.2 (left) and NNPDF2.0 (right) PDF replicas. The central region corresponds to the 68% confidence interval.

The distribution of K_S values for 1000 NNPDF2.0 replicas is shown on the right–hand side of Fig. 5.7. The narrower distribution which one gets in the NNPDF2.0 analysis is closer to Gaussian and no difference is found between the 68% confidence level and (symmetric) 1σ intervals. This is most likely related to the results of Chap. 4, where it was shown that in the data region PDFs uncertainties are basically Gaussian, and that they start deviating from the Gaussian behaviour only in the extrapolation region. The strange distribution being much more constrained by data in the NNPDF2.0 fit, this might provide an explanation for what is observed in Fig. 5.7.

Analysis	Reference	K_S	$[S^-] \cdot 10^3$
NNPDF1.2	[70]	$0.71^{+0.19}_{-0.31}$	0.5 ± 8.6
NNPDF2.0	[71]	0.503 ± 0.075	3.3 ± 1.5
MSTW08(*)	[43]	0.56 ± 0.03	1.4 ± 1.2
Tung:2006tb(*)	[42]	0.72 ± 0.05	1.2 ± 1.2
AKP08	[174]	0.59 ± 0.08	1.0 ± 1.3

Table 5.2: The relative strange momentum fraction $K_S\left(Q^2 = 20\,\mathrm{GeV}^2\right)$ Eq. (5.6) and the strangeness momentum asymmetry $\left[S^-\right]\left(Q^2 = 20\,\mathrm{GeV}^2\right)$ Eq. (5.5), as determined from various parton sets. All uncertainties correspond to 68% confidence levels. For the sets marked by (*) the value for $\left[S^-\right](Q^2)$ is obtained evolving the published value to $Q^2_{\mathrm{ref}} = 20\ \mathrm{GeV}^2$ through NLO perturbative evolution.

Results are summarised in Table 5.2, along with the results found using other parton sets. The NNPDF1.2 uncertainty is much larger than that found in other fits, for the reasons discussed above. Note that, however, all values are essentially consistent with the simple assumption $K_S = 0.5$ used in older parton fits. In the same table the second moment for the strange valence, proportional to the R_S coefficient is shown. Results when necessary are evolved to a common scale, exploiting the fact that at NLO $[S^-]$ evolves multiplicatively. In this case, too, the NNPDF1.2 uncertainty is much larger than that obtained in other fits: while for all other sets, including NNPDF2.0, there is an indication that a positive value of $[S^-]$ is favoured (all results being nevertheless compatible with zero), this indication loses its significance in our analysis due to the very large uncertainty.

Now we look at the phenomenological implications of the NNPDF1.2 analysis, as studied in Ref. [177]. We first studied the sensitivity of the Z/W ratio

$$r_{ZW} \equiv \sigma_Z/(\sigma_{W^+} + \sigma_{W^-})$$

to the uncertainty in the strange distribution. This observable is particularly interesting since PDF uncertainties are greatly reduced when considering the ratio of two cross-sections, and thus provides an excellent candidate for a measurement of the LHC luminosity. However, as it was shown in Ref. [42], although the impact of the uncertainties on the W and Z cross sections due to the strange PDF are rather small on their own, the ratio r_{ZW} has a greater sensitivity to the strange uncertainty in the region $0.01 < x < 0.05$ because PDF uncertainties do not completely cancel in the ratio. Therefore, r_{ZW} is potentially affected by the larger uncertainty of strange PDF found in the NNPDF1.2 analysis with respect to other parton densities determinations. The NNPDF 1.2 prediction for the ratio r_{ZW} and the associated correlation[1] $\rho\left[\sigma(Z), \sigma(W^{\pm})\right]$ are given in Tab. 5.3, together with the results obtained using the

[1] Ref. [42] instead uses the notation $\cos\varphi$ to denote the correlation between two PDFs/observables.

	NNPDF1.2	NNPDF1.0
r_{ZW}	0.0961 ± 0.0005	0.0965 ± 0.0003
$\rho\left[\sigma(Z),\sigma(W^{\pm})\right]$	0.976	0.994
	CTEQ6.6	CTEQ6.5
r_{ZW}	0.0964 ± 0.0004	0.0957 ± 0.0002
$\rho\left[\sigma(Z),\sigma(W^{\pm})\right]$	0.983	0.994

Table 5.3: Comparison of the values for the ratio of the Z and W cross sections at the LHC as well as their associated correlation $\rho\left[\sigma(Z),\sigma(W^{\pm})\right]$, Eq. 5.10, computed with different PDF sets. Again, all numbers shown correspond to $\sqrt{s}$=14 TeV.

NNPDF1.0, CTEQ6.5 and CTEQ6.6 sets. In the Hessian approach, the correlation between the two observables considered, $\sigma(W^{\pm})$ and $\sigma(Z)$, is computed using the method described in Ref. [42]. In the Monte Carlo approach, the corresponding expression is given by

$$\rho\left[\sigma(Z),\sigma(W^{\pm})\right] = \frac{\left\langle\sigma(Z)\sigma(W^{\pm})\right\rangle_{\rm rep} - \left\langle\sigma(Z)\right\rangle_{\rm rep}\left\langle\sigma(W^{\pm})\right\rangle_{\rm rep}}{\sqrt{\left\langle\sigma(Z)^2\right\rangle_{\rm rep} - \left\langle\sigma(Z)\right\rangle^2_{\rm rep}}\sqrt{\left\langle\sigma(W^{\pm})^2\right\rangle_{\rm rep} - \left\langle\sigma(W^{\pm})\right\rangle^2_{\rm rep}}}\,, \quad (5.10)$$

where the averages are performed over the $N_{\rm rep}$ replicas of the NNPDF sets. In Fig. 5.8 we compare the σ_Z-σ_W 1σ correlation ellipses for the NNPDF1.2, NNPDF1.0, CTEQ6.6 and CTEQ6.5 sets. We note that, despite the fact that the error band on the strange parton densities is in general much larger for the NNPDF1.2 set than for CTEQ6.6, the uncertainty on the ratio r_{ZW} is of the same size. This is a consequence of the fact that, as previously mentioned, this ratio is mostly correlated to the strange PDFs in a limited region of relatively small x, where the NNPDF1.2 and CTEQ6.6 uncertainties on s^+ are roughly of the same size. This might be due to the fact that the NuTeV dimuon data, which constrains the strangeness in the two analyses, covers precisely the kinematical range relevant for r_{ZW}. As the error on the WZ ratio is small despite the much larger strangeness uncertainties, this is a hint for its validity as a standard candle.

The situation could be different in the differential rapidity distribution. Indeed, while the total cross-sections probe mostly the central rapidity region, at forward rapidities one might expect differences. In order to check this, in Fig. 5.9 the rapidity distribution of the ratio r_{ZW}

$$\frac{dr_{ZW}}{dy}(y) \equiv \frac{d\sigma^Z(y)/dy}{d\sigma^W(y)/dy} \quad (5.11)$$

is shown together with the associated PDF uncertainties. We observe a sizeable increase in the PDF uncertainty from NNPDF1.0 to NNPDF1.2 at forward rapidities

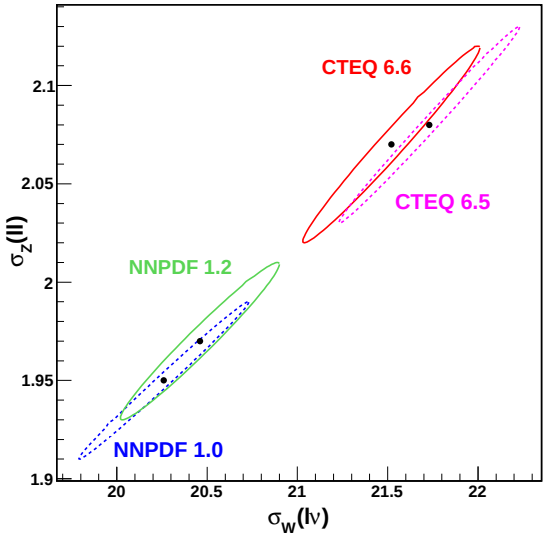


Figure 5.8: Comparison of the W and Z one sigma correlation ellipses at the LHC obtained from different fits: NNPDF 1.2 (green), NNPDF 1.0 (blue), CTEQ 6.6 (red) and CTEQ 6.5 (purple).

due to the increase of the strange PDF uncertainties at small-x as compared to other sets. However in the central region, which provides the dominant contribution to r_{ZW}, the uncertainties of CTEQ6.6 and NNPDF1.2 turn out to be comparable, confirming the agreement of PDF uncertainties shown in Table 5.3.

To conclude, I present another analysis carried out in Ref. [177], focused on a LHC process which in principle could be used to measure the strange PDF, which at the moment is poorly constrained in the region $x < 10^{-2}$ [70]. This process is the Wc associated production. The associated production of a vector boson and a charm quark at hadronic colliders is directly sensitive to the strange sea PDF. The dominant production channel for Wc production is $gq \to Wc$, with q a down-type quark. As in the case of neutrino dimuon production, the down and bottom initiated contributions are suppressed with respect to the strange one by the smallness of the corresponding CKM matrix elements, and therefore at LO, neglecting CKM mixing, the cross section is proportional to the strange PDF. For this reason, the associated W and charm production looks like a promising channel for providing a direct constraint on the strange PDF at the energy scale of M_W, one order of magnitude above the typical energy of the NuTeV dimuon data. For this reason, in the past it has been proposed as a can-

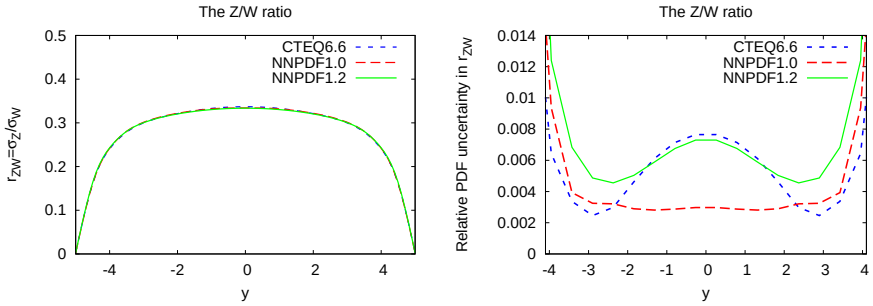


Figure 5.9: Left: the differential rapidity distribution of the ZW ratio, defined in Eq. 5.11. The very small PDF uncertainties are not shown. Right: the relative PDF uncertainties in r_{ZW}, as a function of the rapidity y.

didate for constraining the strange PDF at the Tevatron and the LHC colliders [62]. In Ref. [177] we revisited this proposal by comparing our the NNPDF predictions for the total Wc cross section with recent measurements at the Tevatron and giving predictions for the LHC.

The CDF experiment has published a measurement of the Wc production cross section, obtained using ~ 1.8 fb^{-1} of $p\overline{p}$ collisions at $\sqrt{s} = 1.96$ TeV [178]. A NLO prediction for this observable can be obtained easily using the MCFM [179, 180] code. The result of the CDF collaboration measurement is

$$\sigma_{Wc}(p_{Tc} > 20\text{GeV}, |\eta_c| < 1.5) \times \text{BR}(W \to l\nu) = 9.8 \pm 3.2\text{pb}, \tag{5.12}$$

to be compared with the NLO prediction obtained using MCFM and the NNPDF 1.2 set:

$$\sigma_{Wc}(p_{Tc} > 20\text{GeV}, |\eta_c| < 1.5) \times \text{BR}(W \to l\nu) = 10.11 \pm 1.24(\text{PDF})^{+0.74}_{-0.92}(\text{scale})\ \text{pb}, \tag{5.13}$$

where the first error is the one coming from PDF uncertainties and the second one is due to the variation of the renormalisation and factorisation scales in the perturbative computation. The expected precision of the experimental result, when extrapolated to the full Run II dataset ($\sim 6-7$ fb^{-1}), is $\sim 15\%$, comparable to the present uncertainty on the theoretical prediction. The theoretical uncertainty at the Tevatron is dominated by PDF uncertainties, with a sizeable contribution from scale dependence.

In order to investigate the possibility of using this very same channel as a strangeness constraint at the LHC, we also computed the Wc cross section for the LHC assuming a centre of mass energy of 14 TeV and standard (ν, p_T) cuts both for the charm quark

and the leptons coming from the W decay. The result that we obtain is

$$\sigma_{Wc}(p_{Tc} > 20\mathrm{GeV}/c, |\eta_c| < 4.) \times \mathrm{BR}(W \to l\nu) = 631 \pm 46(\mathrm{PDF})^{+38}_{-63}(\mathrm{scale})\ \mathrm{pb}\,. \quad (5.14)$$

It shows that the uncertainty on the theoretical prediction due to scale variations, and thus to higher order corrections, is comparable to the uncertainty due to the strange PDFs. The result seems to suggest that, contrarily to what is found in Ref. [62], where theoretical uncertainties were not considered, it might be difficult to use this process to constrain PDFs. Unless scale uncertainties can be reduced by higher order computations or by a better understanding of the charm mass treatment[2], the theoretical error limits the usefulness of Wc production as a constraint for the strangeness distributions.

In the comparison with other parton analyses we need to keep into account that there are significant sources of systematics. The most relevant of those, especially in the Wc analysis, is the effect of the heavy quark mass. The treatment of heavy quark mass effects entails various ambiguities related to the prescription used to deal with subleading terms [181]. In our case, a source of systematics is due to the approximate treatment of the charm mass.

5.1.2 Solving the NuTeV anomaly

The coupling which controls neutral current neutrino DIS depends on the electroweak mixing angle, which can thus be extracted from its experimental measurement. Specifically, in the parton model the θ_{W} electroweak mixing angle can be related to the experimentally measurable Paschos-Wolfenstein ratio

$$\begin{aligned} R_{\mathrm{PW}} &\equiv \frac{\sigma(\nu\mathcal{N} \to \nu X) - \sigma(\bar{\nu}\mathcal{N} \to \bar{\nu} X)}{\sigma(\nu\mathcal{N} \to \ell X) - \sigma(\bar{\nu}\mathcal{N} \to \bar{\ell} X)} \quad (5.15) \\ &= \frac{1}{2} - \sin^2\theta_{\mathrm{W}} + \left[\frac{([U^-] - [D^-]) + ([C^-] - [S^-])}{[\mathcal{Q}^-]} \frac{1}{6} \left(3 - 7\sin^2\theta_{\mathrm{W}}\right) \right], \end{aligned}$$

where $[S^-]$ is the strange valence momentum fraction Eq. (5.5), $[U^-]$, $[D^-]$ and $[C^-]$ the valence momentum fractions of other quark flavors, and $[\mathcal{Q}^-] \equiv ([U^-]+[D^-])/2$.

The experimental determination in Ref. [167]

$$\sin^2\theta_{\mathrm{W}}\Big|_{\mathrm{NuTeV}} = 0.2277 \pm 0.0014^{\mathrm{stat}} \pm 0.0009^{\mathrm{sys}} = 0.2277 \pm 0.0017^{\mathrm{tot}}\,, \quad (5.16)$$

is obtained using Eq. (5.15) under the assumption that for an isoscalar nucleon target $[U^-] - [D^-] = 0$ and under the assumption of vanishing charm and strange valence

[2] This is discussed in Chap. 7

$[C^-]$=$[S^-]$=0. With these assumptions the term in square brackets in Eq. (5.15) vanishes. The result Eq. (5.16) disagrees at the 3σ level with the value determined in global precision electroweak fits, such as [182, 183]

$$\sin^2\theta_{\rm W}\Big|_{\rm EWfit} = 0.2223 \pm 0.0003\,. \tag{5.17}$$

Possible explanations for this include nuclear effects, electroweak corrections, QCD corrections, and physics beyond the standard model [168]. However, one may also question the validity of the assumption of the vanishing of the contribution in square brackets in Eq. (5.15). The possibility that $[U^-]-[D^-] \neq 0$ for NuTeV iron target due to isospin violation induced by QED evolution effects was discussed in Ref. [184]: it could easily explain about a third of the observed discrepancy.

In the NNPDF1.2 analysis, isospin symmetry is assumed, and furthermore $[C^-] = 0$. We are then left with the correction

$$\delta_s \sin^2\theta_{\rm W} = -R_S \frac{1}{6}\left(3 - 7\sin^2\theta_{\rm W}\right), \tag{5.18}$$

with R_S defined in Eq. (5.7). Using the value of R_S Eq. (5.9), obtained at the typical scale $Q^2 = 20$ GeV^2 of the NuTeV data we obtain

$$\delta_s \sin^2\theta_{\rm W} = -0.001 \pm 0.011^{\rm PDFs} \pm 0.002^{\rm th}, \tag{5.19}$$

where the theoretical uncertainty comes from the effects discussed in the previous section and summarised in Table 5.1, and it is not to be confused with the experimental systematics in the NuTeV measurement Eq. (5.16).

Even neglecting these theoretical uncertainties, the additional PDF uncertainty due to strangeness alone is thus about twice the observed discrepancy in $\sin^2\theta_{\rm W}$. Applying the correction Eq. (5.19) the NuTeV result becomes

$$\sin^2\theta_{\rm W}\Big|_{\rm NuTeV} = 0.2263 \pm 0.0014^{\rm stat} \pm 0.0009^{\rm sys} \pm 0.0107^{\rm PDFs}. \tag{5.20}$$

We concluded therefore that the apparent inconsistency between the NuTeV measurement and the global electroweak fit disappears once the uncertainty on the strange distribution is properly taken into account.

In the previous section, we showed how, due to the addition of hadronic data, the determination of $[S^-]$ in the NNPDF2.0 analysis [71] is much more precise, see Eq. (5.9). It is natural to ask what is the implication of this reduced uncertainty on the NuTeV analysis. The results shown in Fig. 5.10 are striking: the NuTeV determination of the Weinberg angle [171], using the values of R_S Eqs. (5.9), is compared to the result

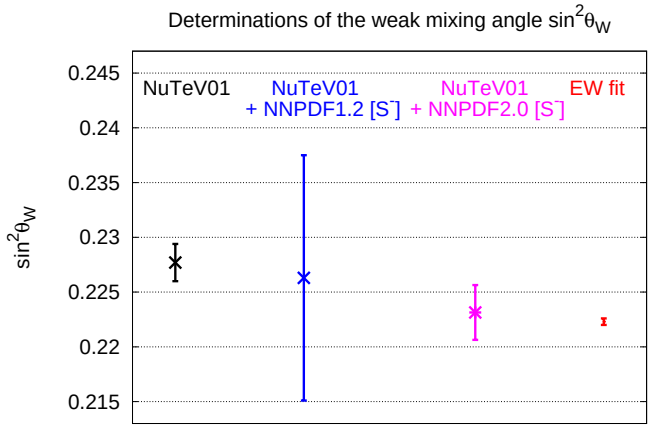


Figure 5.10: Determination of the Weinberg angle from the uncorrected NuTeV data [171], with $\left[S^{-}\right]$ correction using NNPDF1.2 (Eq. (5.9)) and NNPDF2.0 (Eq. (5.9)) results, and from a global electroweak fit [182]. Note that only statistical uncertainties are included in the NNPDF2.0 correction.

of a global electroweak fit [182]. The two corrected values, unlike the uncorrected NuTeV value, are in perfect agreement with the electroweak fit and with each others. However, the uncertainty on the Weinberg angle with NNPDF1.2 correction is considerably larger than the uncertainty after NNPDF2.0 correction.

5.1.3 Precise determination of $|V_{cs}|$ and $|V_{cd}|$

Neutrino DIS data, and especially dimuon data, can be used to perform direct measurements of electroweak parameters [185, 186]. However the potential precision of these measurements can be spoiled by PDF uncertainties. Indeed the uncertainties on the strange distributions are quite large, typically larger by almost one order of magnitude than those found in previous global fits. In this section I show that, in spite of the large uncertainties found in the NNPDF1.2 analysis, in Ref. [70] we provided the most precise direct determination of the CKM matrix element $|V_{cs}|$ within a single experiment. We also provided a determination of $|V_{cd}|$ with an accuracy consistent with previous results from neutrino data. These results are possible because PDF uncertainties are free from parametrisation bias, thus they may be disentangled from the uncertainty on the physical parameters.

The CKM matrix elements [18]-[19] control the strength of the coupling of various partons to neutrinos, as it is shown in Table 1.1. Since the CDHS studies [187], neutrino DIS has been used as a means to directly determine them: the parton–model expressions for the neutrino and anti-neutrino dimuon production Eqs. (5.3, 5.5) provide two equations which relate two experimentally measurable cross sections to the two unknowns $|V_{cd}|$ and $|V_{cs}|$.

However, these equations also contain as unknowns the second moments of the light quark PDFs (the total cross section is proportional to the second moment of the PDF). With the assumption $S^- \approx 0$ [187, 188, 189], the linear combination $F_2^{\nu,c} - F_2^{\bar{\nu},c}$ only depends on the $|V_{cd}|$ and the u and d valence components, which are well measured by other experiments, so it can be used to determine $|V_{cd}|$. On the other hand, the orthogonal combination $F_2^{\nu,c} + F_2^{\bar{\nu},c}$ depends on the $|V_{cs}|/|V_{cd}|$ ratio, but also on K_S, as can be seen from Eq. (5.6), and thus it can only be used to determine the combination $|V_{cs}|K_S$. Indeed, the PDG quotes a value of $|V_{cd}| = 0.23 \pm 0.11$ obtained from the average neutrino dimuon experiments as the best current direct determination [189]. Only the bound $|V_{cs}| \geq 0.74$ at 90% confidence level [188] was quoted in previous PDG [190] editions, but this is now superseded by a direct determination $|V_{cs}| = 1.04 \pm 0.06$ from D decays. Of course, the values obtained from the current global CKM fits [191, 192, 189] are much more precise than these direct determinations (see Table 5.5 below).

In the NNPDF1.2 reference fit, $|V_{cd}|$, $|V_{cs}|$, and $|V_{cb}|$ are each fixed to the current PDG value [189], obtained from the global CKM unitarity fit. To extract both $|V_{cs}|$ and $|V_{cd}|$ from the fit, we performed a scan over the values of $|V_{cs}|$ and $|V_{cd}|$ used in the fit, holding $|V_{cb}|$ fixed, but relaxing the unitarity constraint (in practice, because of its smallness, the precise value chosen for $|V_{cb}|$ is inconsequential). The best–fit value and uncertainty for the CKM parameters are then determined in the standard way by maximum likelihood from the χ^2 profile.

The χ^2 determined from a set of $N_{\rm dat}$ data points fluctuates, with a standard deviation equal to $\sigma_{\chi^2} = \sqrt{2N_{\rm dat}}$. In order to determine the χ^2 profile as the underlying parameters are varied, these fluctuations must be kept under control. Within our Monte Carlo approach, this could be done by using the same set of data replicas each time the χ^2 is recomputed with different values of the underlying parameters. This might however bias the result in a random way depending on the particular set of replicas which has been chosen in the first place. We prefer thus to vary randomly the set of replicas which is used for different parameter values: fluctuations are then kept under control by using a sufficiently large set of replicas, given the fluctuation of the χ^2 computed from a replica average has a standard deviation equal to $\sigma_{\chi^2}/\sqrt{N_{\rm rep}}$. Since only dimuon data are sensitive to the CKM matrix elements, we can determine their values from the dependence of the χ^2 of the fit to these data only, rather than for that

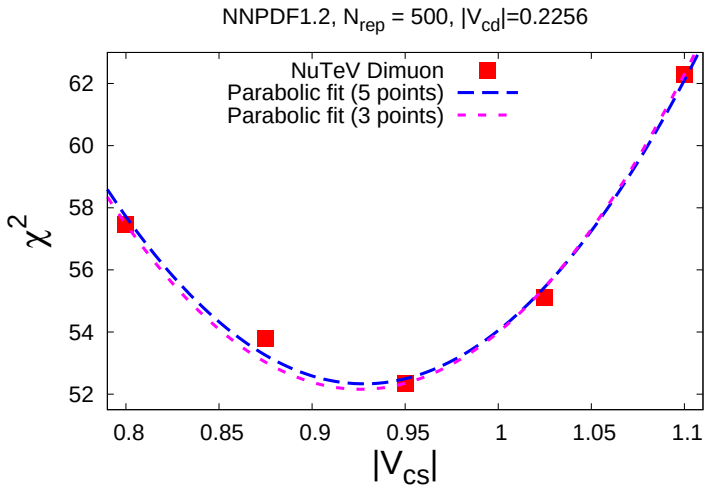


Figure 5.11: The χ^2 of the NuTeV dimuon data as a function of $|V_{cs}|$ when $|V_{cd}|$ is kept fixed at its best unitarity fit value. The solid line is the parabolic fit from which the central value and 1σ uncertainty Eq. (5.21); the dashed curve is a parabolic fit to the central and two outer points only.

of the fit to the global dataset. Given that in the NNPDF1.2 analysis [70], 84 dimuon data points are included, a set of $N_{\rm rep} = 500$ replicas is sufficient to guarantee that point-by-point fluctuations are smaller than $\Delta\chi^2 = 1$.

First, we varied independently each of the two CKM matrix elements, keeping the other fixed at its central value in the CKM unitarity fit. The χ^2 profile is computed for five equally spaced values of the parameter which is being varied. The values have been chosen on the basis of a preliminary exploration of the space of parameters based on fits with a small number of replicas; they are displayed in Fig. 5.13. The ensuing χ^2 profile is displayed in Fig. 5.11 for $|V_{cs}|$ and in Fig. 5.12 for $|V_{cd}|$. We observe well-defined minima in both cases. A parabolic fit leads to

$$|V_{cs}| = 0.93 \pm 0.06, \tag{5.21}$$

$$|V_{cd}| = 0.248 \pm 0.012, \tag{5.22}$$

where the 1σ uncertainty is obtained from the condition $\Delta\chi^2 = 1$. The fit is quite stable upon the choice of different subsets of the five available points: if it is repeated by only retaining the central and two outer points neither the central values nor the

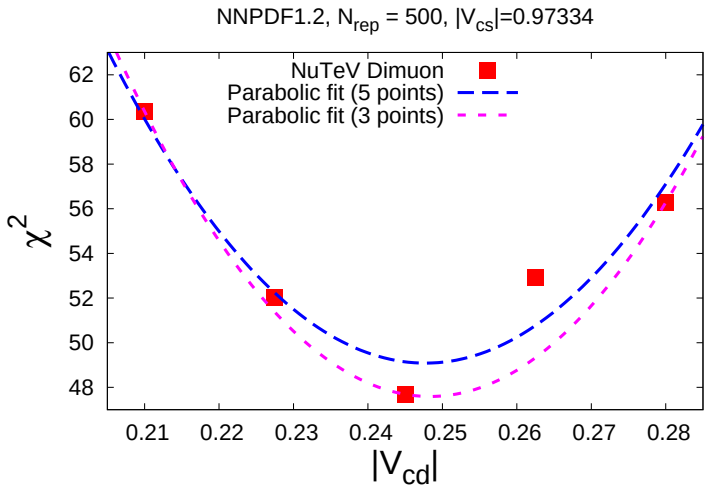


Figure 5.12: The χ^2 of the NuTeV dimuon data as a function of $|V_{cd}|$ when $|V_{cs}|$ is kept fixed at its best unitarity fit value. The solid line is the parabolic fit from which the central value and 1σ uncertainty Eq. (5.22); the dashed curve is a parabolic fit to the central and two outer points only.

uncertainties Eqs. (5.21, 5.22) vary significantly. This confirms that the number of replicas used to compute the χ^2 is sufficiently large for the result not to be biased by statistical fluctuations. Both fits are shown in Figs. 5.11-5.12.

This shows that either CKM matrix element can be determined from our data, with comparable uncertainty, by keeping the other fixed. We can thus perform a simultaneous determination of both these CKM matrix elements. In order to improve the accuracy of this determination, we compute the χ^2 at four more points in the ($|V_{cd}|$, $|V_{cs}|$) plane, denoted by squares in Fig. 5.13. The χ^2 in these additional points is computed from a smaller set of $N_{\rm rep} = 100$ replicas. The result of the combined fit is then

$$|V_{cs}| = 0.96 \pm 0.05, \tag{5.23}$$

$$|V_{cd}| = 0.244 \pm 0.012. \tag{5.24}$$

The uncertainties turn out to be almost identical to the diagonal uncertainties, and the correlation coefficient is relatively small $\rho = 0.21$, reflected in a moderate shift in central values in comparison to the separate fits Eqs. (5.21-5.22). The location of the

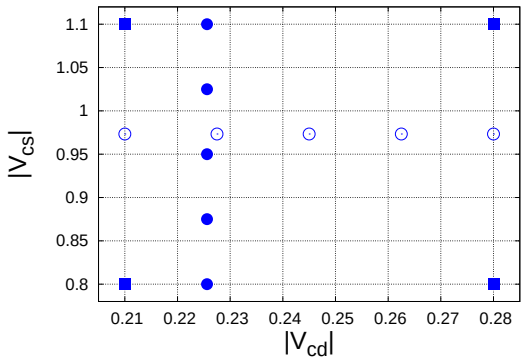


Figure 5.13: The grid of points used in the determination of the CKM matrix elements $|V_{cs}|$ and $|V_{cd}|$. Open circles denotes points used for the determination of $|V_{cd}|$ Eq. (5.22), and full circles points used for the determination of $|V_{cs}|$ Eq. (5.21). All points are used in the joint determination Eq. (5.24).

	$\|V_{cd}\|$	$\|V_{cs}\|$
Statistical	± 0.012	± 0.05
Mass effects	± 0.007	± 0.02
Higher order QCD	± 0.010	± 0.03
Nuclear corrections	± 0.008	± 0.03
Total systematic uncertainty	± 0.014	± 0.05
Total uncertainty	± 0.019	± 0.07

Table 5.4: Summary of statistical and systematic uncertainties in the present determination of $|V_{cs}|$ and $|V_{cd}|$.

best-fit point and 1σ ($\Delta\chi^2 = 1$) ellipse in the ($|V_{cd}|$, $|V_{cs}|$) plane for the best-fit χ^2 paraboloid is shown in Fig. 5.14.

This determination Eq. (5.24) is affected by the same systematics that we examined in the previous section, namely, higher order QCD corrections, treatment of heavy quark effects and modeling of nuclear corrections. In order to assess their impact in the CKM element determination, we have repeated the determination of each of the two parameters as the other is kept fixed, Eqs. (5.21, 5.22), by recomputing the χ^2 for a smaller set of $N_{\rm rep} = 100$ replicas along the points denoted as circles in Fig. 5.13, with each of these three effects varied in turn. We then take the shift in central value as an estimate of the corresponding uncertainty. The results are summarised in Table 5.4.

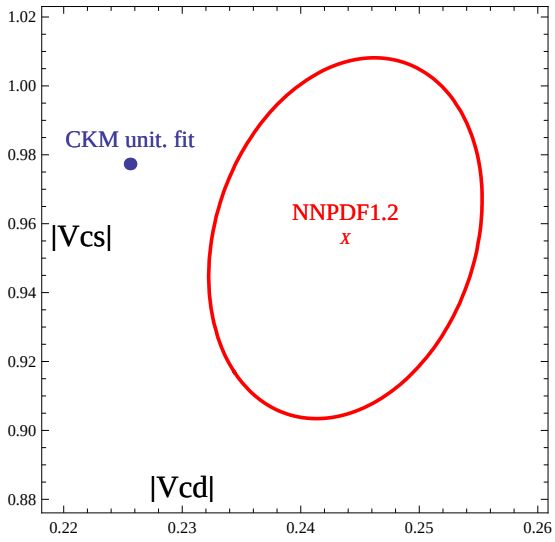


Figure 5.14: Location of the best-fit point and 1σ (statistical $\Delta\chi^2 = 1$ uncertainty) ellipse in the $(|V_{cd}|, |V_{cs}|)$ plane for the best-fit χ^2 paraboloid obtained from the χ^2 computed at the points displayed in Fig. 5.13. The best unitarity fit result [189] is also shown for comparison.

Putting everything together, we found

$$|V_{cs}| = 0.96 \pm 0.07^{\rm tot} , \tag{5.25}$$

$$|V_{cd}| = 0.244 \pm 0.019^{\rm tot} . \tag{5.26}$$

In Table 5.5 we compare the final results Eqs. (5.25, 5.26) with the best CKM unitarity fit results and with other direct determinations. The determination of $|V_{cd}|$ is consistent with other direct determinations, and of comparable accuracy, though one should bear in mind that previous determinations from dimuon data were based on fits with a fixed functional form, and thus subject to potentially large systematics bias. Despite these increased uncertainties, we find that, perhaps surprisingly, the dimuon data are sufficient to determine $|V_{cs}| = 0.96 \pm 0.07^{\rm tot}$. This is one order of magnitude more precise than any other direct determination from neutrino deep-inelastic scattering, and is comparable to the current PDG best average of direct determinations from D meson decays. It would be interesting to repeat the analysis after the inclusion of the Drell-Yan and jet data, i.e. by using the NNPDF2.0 fit. This has not been done yet but is going to be performed in the near future.

Analysis	Description	Reference	$\|V_{cs}\|$
NNPDF1.2	Direct determination from global PDF analysis	[70]	$0.96 \pm 0.07^{\rm tot}$
CDHS	LO determination from $\nu N \to \mu^+\mu^- X$	[187]	≥ 0.59 (90% C.L.)
CCFR	NLO determination from $\nu N \to \mu^+\mu^- X$	[87, 188]	≥ 0.74 (90% C.L.)
PDG08	Avg of determinations from D decays	[189]	1.04 ± 0.06
Hocker	Avg of determinations from $\nu N \to \mu^+\mu^- X$	[193]	1.04 ± 0.16
DELPHI	Direct measurement from $W^+ \to c\bar{s}$ decays	[194]	$0.94^{+0.32}_{-0.26} \pm 0.13$
LEP (avg)	Direct measurement from $W^+ \to c\bar{s}$ decays	[195]	0.95 ± 0.08
PDG08	CKM unitarity fit	[189]	0.97334 ± 0.00023

Analysis	Description	Reference	$\|V_{cd}\|$
NNPDF1.2	Direct determination from global PDF analysis	[70]	$0.244 \pm 0.019^{\rm tot}$
CDHS	LO determination from $\nu N \to \mu^+\mu^- X$	[187]	0.24 ± 0.03
CCFR	NLO determination from $\nu N \to \mu^+\mu^- X$	[188]	$0.232^{+0.017}_{-0.019}$
PDG08	Avg of direct determinations from $\nu N \to \mu^+\mu^- X$	[189]	0.230 ± 0.011
PDG08	Avg of determinations from $D \to K/\pi l\nu$ decays	[189]	0.218 ± 0.023
PDG08	CKM unitarity fit	[189]	0.2256 ± 0.0010

Table 5.5: Comparison of the present determination of the CKM matrix elements $|V_{cs}|$ (upper table) and $|V_{cd}|$ (lower table) with other available direct measurements, averages and CKM constrained fits.

5.2 Prediction for LHC standard Candle

As for several relevant processes at the LHC the uncertainties coming from Parton Distribution Functions (PDFs) and from the strong coupling constant $\alpha_s\left(M_Z\right)$ are the dominant ones, it is important to have a reliable estimate of their uncertainties and to use a correct prescription to combine them, keeping into account the fact that PDFs, especially the gluon, and α_s are strongly correlated. In particular, when performing a comparison between predictions evaluated with different PDFs sets, it is desirable to know how much of the discrepancy is due to the difference in α_s and how much instead is due to the uncertainty on PDFs.

In this section the NNPDF predictions for some of the so–called standard candle processes at the LHC are presented. They are compared to the results of the other two global analyses, MSTW2008 and CTEQ6.6. For the comparison I consistently use a common value of α_s. Then the dependence of the NNPDF2.0 global set of partons upon the central value of α_s is investigated and the correlation between the gluon PDF and the strong coupling constant is evaluated. Finally I discuss the combination of PDF and α_s uncertainty.

5.2.1 NNPDF predictions for LHC standard candles

Schematically, the measured cross section for a process X is given by

$$\sigma_X = \frac{N_X}{\epsilon_X \mathcal{L}}, \tag{5.27}$$

where N_X is the number of observed events generated by the process X, ϵ_X is the experimental efficiency and $\mathcal{L}$ is the integrated luminosity. The latter measures the beam intensity and depends upon several factors, like the bunch sizes, the crossing angles between beams, the density of the beams and so on. The measurement of the integrated luminosity is needed and its accuracy influences all predictions. There are several ways for measuring it directly from beam parameters. A good accuracy of the direct measurements requires sophisticated techniques and depends on the details of the detector. Another way for measuring the luminosity is to perform an indirect measurement. Inverting the above equation one has

$$\mathcal{L} = \frac{N_Y}{\epsilon_Y \sigma_Y}. \tag{5.28}$$

If Y is some well–known process, both from the experimental and theoretical points of view, one can infer the luminosity from the measured number of events, the experimental efficiency and the theoretical prediction σ_Y. In this measurement several factors are relevant: the understanding of the detector and of the background contamination, and the accuracy of the theoretical prediction.

These processes are the so–called standard candles. Their measurements might also be used to check the theoretical framework, like factorisation and DGLAP evolution equations. Two typical standard candles are the production of the Z and W bosons, for which there are realistic prospects of experimental and theoretical accuracy at a few % level. Most of the theoretical uncertainty due to the error on the PDFs. Indeed, the theoretical uncertainty due to the scale variation in $\sigma(Z)$ is about 1%, while the PDF uncertainty is about 2%. In Refs. [43, 196], a study of the correlated Z and $W^{\pm}$ production cross section has been performed with the MSTW input PDFs, and the error ellipses show that the total 1σ uncertainty for the NNLO prediction is about 4%. Another process which has been proposed as a candidate to measure the luminosity is the total $t\bar{t}$ cross section, $\sigma\left(t\bar{t}\right)$. Several recent studies [42, 197, 198] motivated such proposal. For this process the PDF uncertainty is only slightly larger than the scale variation uncertainty associated with higher orders in perturbation theory, which is about 3%. Finally it is important to provide an estimate of the PDF uncertainty on the Higgs production, even if in that case the uncertainty due to the scale variation is much more significant than the uncertainty due to the PDFs. On the other hand,

given that the main contribution to the Higgs production at the LHC comes from the gluon–gluon fusion, the uncertainty on α_s sizeably affects the total uncertainty.

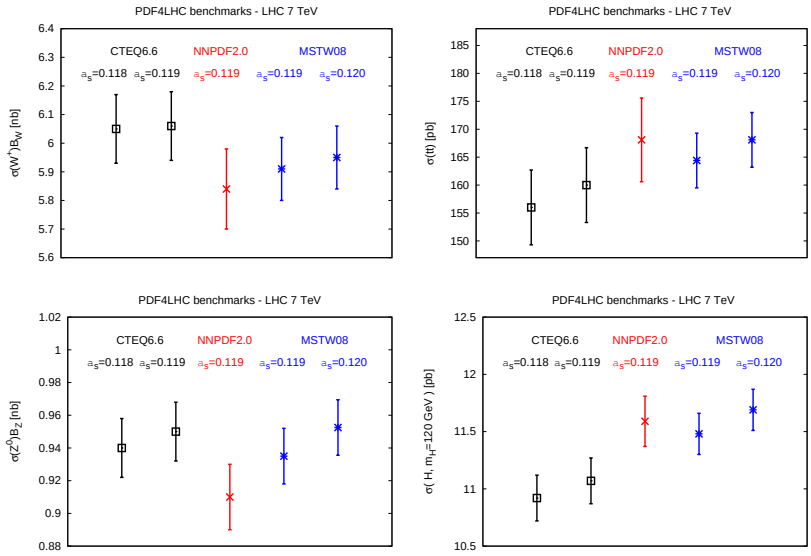


Figure 5.15: Comparison of predictions for LHC observables for NNPDF2.0, MSTW08 and CTEQ6.6 sets for the LHC at centre of mass energy of 7 TeV.

In Fig. 5.15 and in Table 5.6 predictions for the observables mentioned above are compared from the three global PDF fits: NNPDF2.0 [71], MSTW08 [43] and CTEQ6.6 [42]. The observables are computed at a centre–of–mass energy of 7 TeV, like in the early LHC runs by using the public next–to–leading order MCFM code [179]. For CTEQ and MSTW we show results both at the default value of α_s and for a common value $\alpha_s\left(M_Z\right) = 0.119$ in order to disentangle the discrepancy due to the use of a different α_s reference value and the discrepancy due to the PDFs. In order to produce the CTEQ6.6 and MSTW08 predictions with $\alpha_s = 0.119$, the sets from Refs. [199, 200] have been used. We also assumed that the PDF uncertainty for these two PDF sets does not depend in a statistically significant way on the value of α_s when switching from the default to the common value of α_s (which in both cases differ by $\delta\alpha_s = 0.001$).

It is clear that using a common value of the strong coupling improves the agreement between global PDF sets. If predictions with $\alpha_s = 0.119$ are compared, the three global PDF sets are in reasonable agreement both in central values and in uncertainties.

	$\sigma(W^+)\text{Br}\left(W^+ \to l^+\nu_l\right)$ [nb]	$\sigma(Z^0)\text{Br}\left(Z^0 \to l^+l^-\right)$ [nb]
NNPDF2.0	5.84 ± 0.14	0.91 ± 0.02
CTEQ6.6 - $\alpha_s = 0.118$	4.10 ± 0.09	0.94 ± 0.02
CTEQ6.6 - $\alpha_s = 0.119$	4.11 ± 0.09	0.95 ± 0.02
MSTW08 - $\alpha_s = 0.119$	4.16 ± 0.08	0.94 ± 0.02
MSTW08 - $\alpha_s = 0.120$	5.95 ± 0.11	4.19 ± 0.08

	$\sigma(t\bar{t})$ [pb]	$\sigma(H, m_H = 120\,\text{GeV})$ [pb]
NNPDF2.0	168.1 ± 7.5	11.59 ± 0.22
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 5.6: 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 NNPDF2.0, CTEQ6.6 and MSTW08 PDF sets. All uncertainties shown are 1σ level. See Fig. 5.15 for the graphical representation of the results of this table.

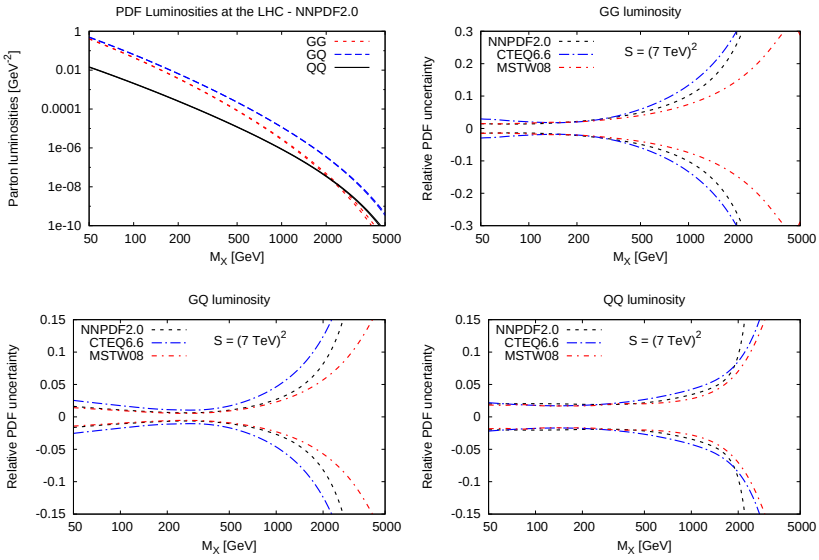


Figure 5.16: Parton–parton luminosities Eq. (5.29) in the various partonic channels, computed from the NNPDF2.0 set at the LHC for $\sqrt{s} = 7$ (top left). Relative PDF uncertainties on parton–parton luminosities Eq. (5.29) for the NNPDF2.0, CTEQ6.6 and MSTW2008 PDF sets, as function of the mass of the produced heavy object M_X at the LHC for 7 TeV. In a clockwise sense, the gluon-gluon luminosity, the gluon-quark luminosity and the quark-quark luminosity are shown.

To understand the comparable uncertainty between different predictions, it is useful to look at the qq, qg and gg parton fluxes, defined as

$$\Phi_{gg}\left(M_X^2\right) = \frac{1}{s}\int_\tau^1 \frac{dx_1}{x_1} g\left(x_1, M_X^2\right) g\left(\tau/x_1, M_X^2\right) , \tag{5.29}$$

$$\Phi_{gq}\left(M_X^2\right) = \frac{1}{s}\int_\tau^1 \frac{dx_1}{x_1} \left[g\left(x_1, M_X^2\right) \Sigma\left(\tau/x_1, M_X^2\right) + (1 \to 2)\right] ,$$

$$\Phi_{qq}\left(M_X^2\right) = \frac{1}{s}\int_\tau^1 \frac{dx_1}{x_1} \sum_{i=1}^{N_f} \left[q_i\left(x_1, M_X^2\right) \bar{q}_i\left(\tau/x_1, M_X^2\right) + (1 \to 2)\right] ,$$

Each of them is relevant for the considered processes, for instance Φ_{qq} is mostly relevant for the $W^{\pm}$ and Z production, while Φ_{gg} is more relevant for the $t\bar{t}$ and light Higgs production. In Fig. 5.16 these quantities are evaluated with the NNPDF2.0 parton set. In the same figure, the relative uncertainty is compared to the uncertainty of the parton fluxes evaluated with the CTEQ6.6 and the MSTW08 parton sets. The comparable uncertainty is the origin of the comparable uncertainty of the predictions for the considered standard candles.

To understand the remaining differences in central values between PDF sets, one has to keep into account that a different treatment of the heavy quark masses is used, as it is explained in Chap. 3. For deeper investigations a detailed benchmarking on the lines of the HERA-LHC benchmarks [101] would be required.

5.2.2 Determination of α_s

Even though the strong coupling can be determined by a parton fit [43], its most accurate determination is obtained by combining results from many high–energy processes, most of which do not depend on PDFs at all. A recent combined determination [12] is

$$\alpha_s(M_Z^2) = 0.1184 \pm 0.0007, \tag{5.30}$$

while the current average in the PDG [109] has a rather more conservative assessment of the 1σ uncertainty:

$$\alpha_s(M_Z^2) = 0.1176 \pm 0.002. \tag{5.31}$$

The PDG world average is dominated by NNLO results, with the NLO results included carrying little weight.

On the other hand, various parton fitting groups use different values of α_s. Specifically, for the PDF sets that I am going to include in the following analysis the reference values are

$$\begin{aligned} \alpha_s(M_Z^2) &= 0.118, \qquad \text{for CTEQ6.6,} \\ \alpha_s(M_Z^2) &= 0.119, \qquad \text{for NNPDF2.0,} \\ \alpha_s(M_Z^2) &= 0.12018, \qquad \text{for MSTW08.} \end{aligned} \tag{5.32}$$

However, in recent publications, sets of partons with varying α_s have been presented by the fitting collaborations. Specifically CTEQ [42] has released, on top of the set evaluated with the central value of α_s reported above, four more sets with $\alpha_s =$

$0.116, 0.117, 0.119, 0.120$. However for these sets only the best–fit set is provided without the error sets. In the CTEQ analyses α_s is taken as an external parameter which is fixed before starting the fit.

The MSTW collaboration has performed a simultaneous determination of PDFs and α_s [199], which is treated as a parameter of the fit and whose fitted value is the one reported in Eq. (5.33). This set may be used for the determination of the correlation between α_s and both the central value and the uncertainties of PDFs. The value of α_s found in the fit must be used. On top of that, MSTW has released sets of PDFs with α_s taken as a external parameters for twenty values of α_s varied in steps of 0.001 for from $\alpha_s = 0.110$ up to $\alpha_s = 0.130$.

In the NNPDF2.0 analysis, α_s is taken as an external parameter. In order to study the dependence of the results on the choice of $\alpha_s(M_Z^2)$, the fit with α_s varied by one and two standard deviations about this value is performed, by repeating the fit and producing PDFs sets of 100 replicas for each central value of α_s from 0.114 to 0.124 in steps of 0.001. The same was done in the previous NNPDF1.0 and NNPDF1.2 analysis: the dependence of PDFs on α_s was found [71, 156] to be noticeable but weak, so much so that when α_s was varied by $\Delta\alpha_s = \pm 0.002$ most PDFs were found to be statistically indistinguishable from those obtained with α_s fixed to its central value (i.e. to be at a distance $d \approx 1$ from them). The gluon (and to a lesser extent the singlet PDF) was found to change in a statistically significant way, but still within its uncertainty band when α_s was varied in this range.

The dependence of NNPDF2.0 PDFs on α_s is shown in Figs. 5.17-5.18, where the ratio of the four α_s PDF sets to the central set are shown for all PDFs except the total strangeness s^+ which is found not to vary significantly. All PDFs are still within the central uncertainty band when $\Delta\alpha_s = \pm 0.002$. However, there appears to be now a greater sensitivity to α_s. Firstly, now not only the gluon but also the triplet, singlet and valence, when α_s is varied in the range $\Delta\alpha_s = \pm 0.002$, move close to the edge of the 1σ range for the central PDF. This corresponds to a distance $d \approx 7$, well above the threshold of statistical significance, and even for the gluon it is a somewhat larger variation than observed in NNPDF1.2. Furthermore, the triplet appears to be as sensitive as the gluon to the value of α_s. The increased sensitivity of quark distributions to the value of α_s is likely a consequence of the inclusion of Drell-Yan data, which undergo large NLO corrections and are thus sensitive to α_s. This increased sensitivity with respect to α_s suggests that the strong coupling could be determined from the global PDF analysis with competitive accuracy, following a procedure similar to that used to obtain the accurate determination of the CKM matrix element $|V_{\rm cs}|$.

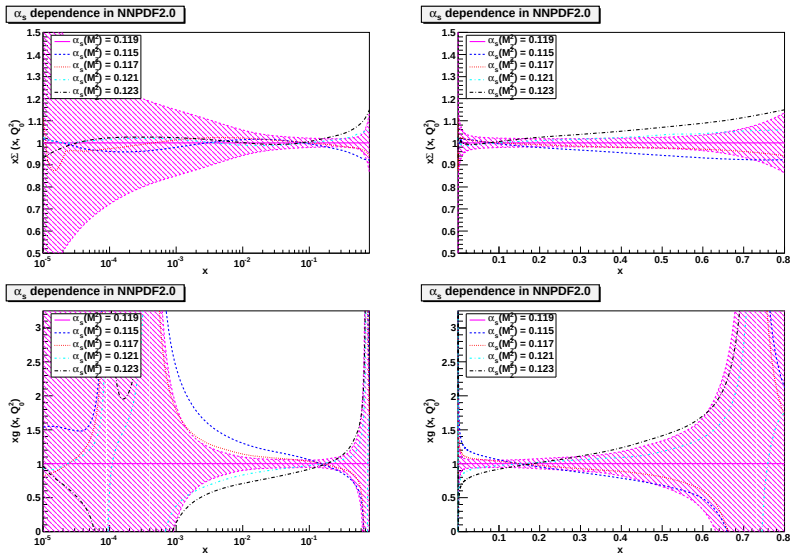


Figure 5.17: Ratios of PDFs with α_s varied in the range $0.115 \leq \alpha_s \leq 0.123$ to the central NNPDF2.0 determination, compared to the PDF uncertainty band: the singlet at small– and large–x, the gluon at small– and large–x (from top to bottom and from left to right).

The qualitative behaviour of the PDFs with varying α_s shown might be understood by taking into account the correlation between the parton distributions and the central value of α_s. For instance, in a purely DIS fit, the gluon is determined by scaling violation of deep–inelastic structure functions mostly in the medium–small x regions. In the high–x region it is purely determined by momentum sum rules. Hence at large–x a higher value of α_s requires a smaller value of the gluon and vice versa. This qualitative picture is supported by a quantitative study of the correlation between α_s and the gluon, or any other PDFs. As it was shown in Chap. 4, the correlation coefficient may be evaluated as

$$\rho[\alpha_s(M_Z^2), f(x,Q^2)] = \frac{\left\langle \alpha_s(M_Z^2) f(x,Q^2) \right\rangle_{\rm rep} - \left\langle \alpha_s(M_Z^2) \right\rangle_{\rm rep} \left\langle f(x,Q^2) \right\rangle_{\rm rep}}{\sigma_{\alpha_s} \sigma_f}. \quad (5.33)$$

In the above equation the distribution of α_s can be evaluated as follows. Given sets of replicas determined with different values of α_s, it is possible to assume a distribution of values for α_s. For instance, the average over Monte Carlo replicas of a general quantity which depends on both α_s and the PDFs, $\mathcal{F}\left(\mathrm{PDF}, \alpha_s\right)$ can be computed as

$$\left\langle \mathcal{F} \right\rangle_{\rm rep} = \frac{1}{N_{\rm rep}} \sum_{j=1}^{N_\alpha} \sum_{k_j=1}^{N_{\rm rep}^{\alpha_s^{(j)}}} \mathcal{F}\left(\mathrm{PDF}^{(k_j,j)}, \alpha_s^{(j)} \right) , \quad (5.34)$$

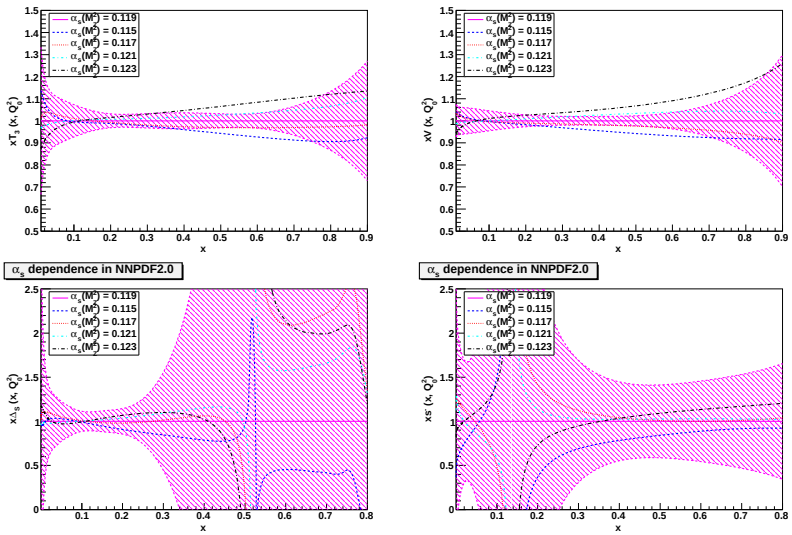


Figure 5.18: Ratios of PDFs with α_s varied in the range $0.115 \leq \alpha_s \leq 0.123$ to the central NNPDF2.0 determination, compared to the PDF uncertainty band: triplet, valence, sea asymmetry and strange valence (from top to bottom and from left to right).

where $\mathrm{PDF}^{(k_j,j)}$ stands for the k_j–th replica of the PDF fit with $\alpha_s = \alpha_s^{(j)}$, and the numbers $N_{\mathrm{rep}}^{\alpha_s^{(j)}}$ of replicas for each value of α_s in the total sample are determined by the probability distribution of values of α_s, with the constraint

$$N_{\mathrm{rep}} = \sum_{j=1}^{N_{\alpha_s}} N_{\mathrm{rep}}^{\alpha_s^{(j)}} \,. \tag{5.35}$$

Specifically, assuming that global fit values of α_s are Gaussianly distributed, the number of replicas is given by

$$N_{\mathrm{rep}}^{\alpha_s^{(j)}} \propto \exp\left(-\frac{\left(\alpha_s^{(j)} - \alpha_s^{(0)}\right)^2}{2\left(\delta_{\alpha_s}^{(68)}\right)^2} \right) \,. \tag{5.36}$$

with the normalisation condition Eq. (5.35). Plugging the above equations into Eq. (5.33) one may compute the correlation coefficient and plot it as a function of x. In Fig. 5.19 the correlation between α_s and the gluon determined both in the NNPDF1.2 and NNPDF.20 analyses are plotted. The same pattern is observed in both analyses: a

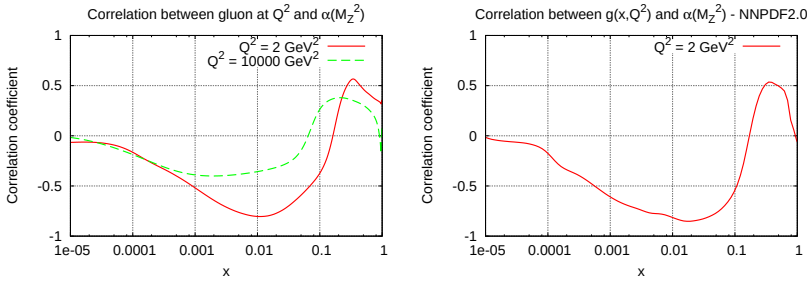


Figure 5.19: Correlation coefficient between the gluon PDF in the NNPDF1.2 (left) and NNPDF2.0 (right) analyses. The correlations is computed according to Eqs. (5.33, 5.34)

sizeable correlation at high–x and a strong anti–correlation at medium–x, mostly due to momentum sum rules.

5.2.3 PDFs+α_s uncertainty for LHC standard candles

To conclude, in this section I present a method for combining the PDF+α_s uncertainty keeping into account their correlation. This method is then used to evaluate the combined uncertainties on the standard–candle processes considered in the previous section, and to compare them with the results obtained by estimating the error as a sum in quadrature.

A way of computing the total PDF+α_s uncertainty when the correlation is kept into account has been proposed in Refs. [199, 156]. In the MSTW methodology of Ref. [199] this is done by relying on a simultaneous determination of PDFs and α_s. As a consequence, the value and range of variation of α_s must be those obtained in this determination. The procedure is the following: the total upper and lower (generally asymmetric) uncertainties on an observable F are determined as

$$\begin{aligned}(\Delta F)_+^{\rm PDF+\alpha_s} &= \max_{\alpha_s}\left(\{F^{\alpha_s}(S_0)+(\Delta F^{\alpha_s}_{\rm PDF})_+\}\right)-F^{\alpha_s^0}(S_0) \\ (\Delta F)_-^{\rm PDF+\alpha_s} &= F^{\alpha_s^0}(S_0)-\min_{\alpha_s}\left(\{F^{\alpha_s}(S_0)-(\Delta F^{\alpha_s}_{\rm PDF})_-\}\,,\right)\end{aligned} \tag{5.37}$$

where $F^{\alpha_s}(S_0)$ is the observable computed using the central PDF set S_0 and the value α_s of the strong coupling, $(\Delta F^{\alpha_s}_{\rm PDF})_\pm$ is the PDF uncertainty on the observable for given fixed value of α_s, as determined from the Hessian PDF eigenvectors [43, 199], and the maximum and minimum are determined from a set of five results, each

computed with one distinct value of α_s (central, $\pm$ half confidence level, $\pm$ confidence level).

With the NNPDF methodology one is free to choose any value and range for α_s, inasmuch as the corresponding Monte Carlo PDF replicas are available. In the approach proposed in Ref. [156] the uncertainty is simply given by the standard deviation of the joint distribution of PDF replicas and α_s values

$$\Delta F^{\rm PDF+\alpha_s} = \sigma_F \equiv \left[\frac{1}{N_{\rm rep}-1} \sum_{j=1}^{N_\alpha} \sum_{k_j=1}^{N_{\rm rep}^{\alpha_s^{(j)}}} (F[\left\{q^{(k_j,j)}\right\}] - F[\{q^0\}])^2 \right]^{1/2} \tag{5.38}$$

where the number of replicas $N_{\rm rep}^{\alpha_s^{(j)}}$ for each value $\alpha_s^{(j)}$ of the strong coupling is determined in the Gaussian case by Eq. (5.36). In this case, we have taken as central value and uncertainty on α_s the one given in Eq. (5.33).

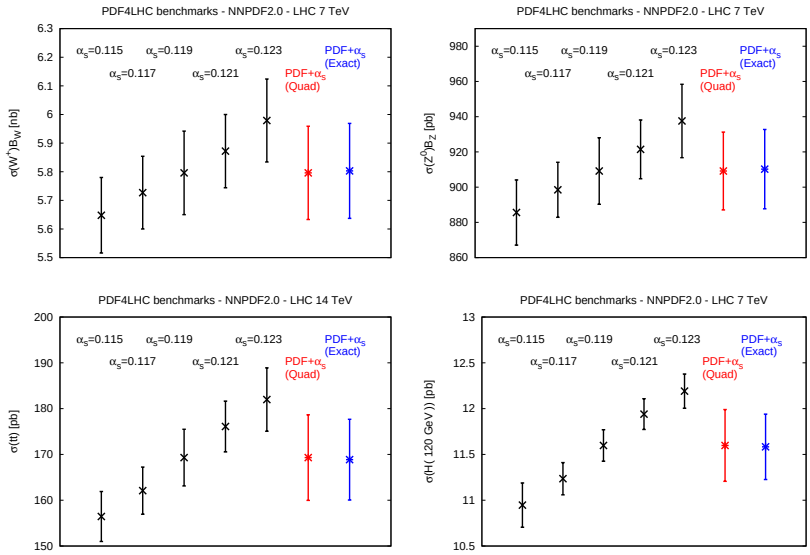


Figure 5.20: Predictions for some important LHC observables computed at 7 TeV. From top to bottom and from left to right: W^+ and Z production, $t\bar{t}$ production, and Higgs production in gluon-gluon fusion for $m_H = 120$ GeV. Results are shown for different values of $\alpha_s\,(M_Z)$ as well as for the combined PDF+α_s uncertainties.

The results for the uncertainty obtained in this way for W^+ and Z^0 production, $t\bar{t}$ production and Higgs production in gluon–fusion for $m_H = 120$ GeV are shown in

Fig. 5.20. We computed predictions for various values of α_s in order to determine the combined PDF+α_s uncertainties for these observables. The exact combined PDF+α_s uncertainty is compared to the one obtained by adding in quadrature the PDF uncertainties and the α_s uncertainties, in turn obtained either with fixed PDFs, or by taking the PDF set that corresponds to each value of α_s.

It is clear that the two methods, quadrature and exact propagation, yield essentially identical results, confirming the results found in Ref. [156]. The effect of the correlation between α_s and PDF uncertainties is indeed quite small: as one might expect, it is in fact smaller than the effect of the correlation between α_s and PDF central values shown in Fig. 5.19. Moreover PDF uncertainties are independent of α_s for any reasonable range of α_s. For processes which depend on α_s at leading order like Higgs or $t\bar{t}$ production, the combined PDF+α_s uncertainty is as expected sizeably larger than the PDF uncertainty alone: for such processes, comparing predictions from different PDF sets using a common value of α_s is mandatory to obtain a meaningful comparison.

5.3 Reweighting

The use of statistical inference to determine the probability distribution function of derived quantities starting from some initial probability density of parton distribution functions (PDFs) was first advocated in Ref. [110]. In the original work, such initial probability density for the PDFs, $\mathcal{P}_{\rm init}[\{f\}]$, is projected from the functional space of all possible functional forms assumed by the parton densities into the $N_{\rm par}-$dimensional space of parameters characterising the chosen functional form. The space of parameters is sampled by generating a Monte Carlo (MC) ensemble consisting of $N_{\rm rep}$ random sets of parameters. The expectation values with respect to the initial density are then computed as averages over the MC ensemble.

The method has several interesting applications depending on the *derived* quantities that are considered. First it allows one to propagate easily the PDF uncertainty to a new observable without having to perform the usual error propagation that involves the derivative of the observable with respect to the parameters. Moreover it can be used to include new observables into an existing fit without having to perform further minimisations.

The limitations of the method proposed in Ref. [110] are related to the generation of the Monte Carlo ensemble. Indeed, the sampling of the parameter space is not particularly efficient due to the presence of many flat directions, which require the generation of very large ensembles which cannot be used for practical purposes.

In the NNPDF approach [67, 68, 70, 71] this problem is solved by generating $N_{\rm rep}$ MC replicas in the space of the data included into the fit, distributed according to the probability distribution of the experimental data; the MC ensemble of PDFs are obtained by fitting each data replica, and therefore the sampling in PDF space is determined by the data themselves. This ensemble of $N_{\rm rep}$ best-fit PDFs is the final output of the NNPDF fits. Hence the method developed within the NNPDF collaboration is ideally suited for the application of the techniques based on statistical inference.

In the next sections, I first briefly describe the statistical principles upon which the reweigthing is based and then I show several examples of application of the reweighting technique.

5.3.1 Bayesian reweighting

Given a NNPDF ensemble of PDFs $\mathcal{E} = \{f_k; k = 1, \ldots, N\}$, the mean, the error, the correlation of any quantity depending on the PDFs may be evaluated by computing these statistical estimators over the set of replicas, according to the formulae reported in Appendix C. For instance the integral in the space of functions is approximated by the average over the ensemble $\mathcal{E}$, so that the mean value of an observable $\mathcal{O}$ depending on PDFs is given by

$$\begin{aligned} \langle \mathcal{O} \rangle &= \int \mathcal{O}[f]\, \mathcal{P}_{\rm old}(f)\, Df \\ &= \frac{1}{N} \sum_{k=1}^{N} \mathcal{O}[f_k]\,. \end{aligned} \tag{5.39}$$

Consider a set of n new data that have not been included in the determination of the initial probability density distribution:

$$\{y\} = y_1, y_2, \cdots, y_n.$$

The experimental uncertainties are summarised by the $d \times d$ experimental covariance matrix cov_{ij} defined in Eq. (2.3). Assuming that the new experiments are not correlated with any of the experiments used in the determination of the initial probability

density, the probability density distribution of $\{y\}$ is given by:

$$\begin{aligned}\mathcal{P}(y) &= \int \mathcal{P}(y|f)\mathcal{P}_{\rm old}(f)\,Df \\ &= \tfrac{1}{N}\sum_{k=1}^{N}\mathcal{P}(y|f_k),\end{aligned} \tag{5.40}$$

where $\mathcal{P}(y|f_k)$ is the conditional probability density distribution, often called likelihood function. Assuming that uncertainties are Gaussian, the probability $d^nP(y|f_k)$ that the new data lie in an infinitesimal volume d^ny of the space of possible data given the k-th element f_k of the ensemble of PDFs is

$$d^nP(y|f_k) = \mathcal{P}(y|f_k)d^ny = (2\pi)^{-n/2}(\det {\rm cov}_{ij})^{-1/2}e^{-\frac{1}{2}\chi^2(y,f_k)}d^ny. \tag{5.41}$$

where $\chi^2(y,f_k)$ is calculated using the new data:

$$\chi^2(y,f_k) = \sum_{i,j=1}^{n}(y_i - f_k(x_i)){\rm cov}_{ij}^{-1}(y_j - f_k(x_j)) \tag{5.42}$$

The probability density for the new dataset is then obtained by averaging over replicas according to Eq. (5.40):

$$d^nP(y) = \mathcal{P}(y)d^ny = \tfrac{1}{N}\sum_{k=1}^{N}\mathcal{P}(y|f_k)d^ny. \tag{5.43}$$

Similarly the probability density for the χ^2 to the new dataset may be evaluated:

$$dP(\chi^2|f_k) = \mathcal{P}(\chi^2|f_k)d\chi^2 = 2^{-n/2}(\Gamma(n/2))^{-1}(\chi^2(y,f_k))^{n/2-1}e^{-\frac{1}{2}\chi^2(y,f_k)}d\chi^2, \tag{5.44}$$

where $\mathcal{P}(\chi^2|f_k)$ is now the χ^2 distribution. This may be readily derived from Eq. (5.41) by diagonalising the covariance matrix and rescaling the data to a set $\{Y\}$ of independent Gaussian variables each with unit variance. Then $d^ny = (\det\sigma_{ij})^{1/2}d^nY$, and $\chi^2 = \sum_{i=1}^{n}Y_i^2$. Choosing n-dimensional spherical co-ordinates in the space of data (with $\sqrt{\chi^2}$ as the radial co-ordinate), we may write $d^nY = A_n\frac{1}{2}(\chi^2)^{n/2-1}d\chi^2d^n\Omega$, where $d^{n-1}\Omega$ is the measure on the sphere and $A_n = 2\pi^{n/2}(\Gamma(n/2))^{-1}$ is the area of the unit sphere in n-dimensions. The probability Eq. (5.41) may thus be written

$$\begin{aligned}d^nP(y|f_k) &= (2\pi)^{-n/2}e^{-\frac{1}{2}\chi^2(y,f_k)}d^nY \\ &= 2^{-n/2}(\Gamma(n/2))^{-1}(\chi^2(y,f_k))^{n/2-1}e^{-\frac{1}{2}\chi^2(y,f_k)}d\chi^2d^{n-1}\Omega.\end{aligned} \tag{5.45}$$

in agreement with Eq. (5.44) provided

$$d^n P(y|f_k) = dP(\chi^2|f_k)d^{n-1}\Omega. \tag{5.46}$$

Again the probability density $\mathcal{P}(\chi^2)$ for the χ^2 of the new dataset is obtained by averaging over replicas:

$$dP(\chi^2) = \mathcal{P}(\chi^2)d\chi^2 = \tfrac{1}{N}\sum_{k=1}^{N}\mathcal{P}(\chi^2|f_k)d\chi^2. \tag{5.47}$$

so combining Eq. (5.43) and Eq. (5.47)

$$d^n P(y) = dP(\chi^2)d^{n-1}\Omega, \tag{5.48}$$

since $d^{n-1}\Omega$ is independent of the choice of replica, and may thus be taken out of the sum.

If the new data are simply consistent with the old data as summarised in the probability density $\mathcal{P}_{\rm old}(f)$ the probability density $\mathcal{P}(\chi^2|f)$ should be a χ^2 distribution for n degrees of freedom: for n large this is peaked at n, with width $\sqrt{2n}$. To test this one might evaluate $\langle\chi^2\rangle$ and the standard deviation σ_{χ^2}. A value of $\frac{1}{n}\langle\chi^2\rangle$ much greater than one suggests that the new data are inconsistent with the old, while if the value of $\sigma_{\chi^2_{\rm new}}$ is much larger than $\sqrt{n}$, the new data should be useful to constrain PDFs. Smaller values mean that the errors on the new data are probably overestimated.

Using the formalism of statistical inference, we can update the old probability density $\mathcal{P}_{\rm old}(f)$ by taking into account the new data to give an improved probability density $\mathcal{P}_{\rm new}(f)$. The new probability density is the conditional probability for $\{f\}$ taking into account the new data $\{y\}$, i.e. $\mathcal{P}_{\rm new}(f) = \mathcal{P}(f|y) = \mathcal{P}(f|\chi^2)$. It may thus be determined from the probability densities $\mathcal{P}(y|f)$ or $\mathcal{P}(\chi^2|f)$ using Bayes' theorem, which relates the conditional probabilities $P(A|B)$ and $P(B|A)$:

$$P(A|B)P(B) = P(B|A)P(A), \tag{5.49}$$

where $P(A)$ and $P(B)$ are the unconditional probabilities of A and B. Note that Eq. (5.49) relates probabilities, not probability densities: to apply it to probability densities needs some care with regard to measure factors. Thus for example

$$(\mathcal{P}(f|\chi^2)Df)(\mathcal{P}(\chi^2)d\chi^2) = (\mathcal{P}(\chi^2|f)d\chi^2)(\mathcal{P}(f)Df). \tag{5.50}$$

Note that the marginalization Eq. (5.47) follows directly on integration over f, since if $\mathcal{P}(f|\chi^2)$ is correctly normalized, $\int\mathcal{P}(f|\chi^2)Df = 1$. Now, cancelling the $d\chi^2$ from

either side of Eq. (5.50) (since this is just a pre-assigned interval),

$$\mathcal{P}(f|\chi^2)Df = \frac{\mathcal{P}(\chi^2|f)}{\mathcal{P}(\chi^2)}\mathcal{P}(f)Df. \tag{5.51}$$

Multiplying on both sides by some observable $\mathcal{O}[f]$ and integrating over the PDFs,

$$\begin{aligned}\langle\mathcal{O}\rangle_{\rm new} &= \int \mathcal{O}[f]\,\mathcal{P}(f|\chi^2)\,Df,\\ &= \int \mathcal{O}[f]\,\frac{\mathcal{P}(\chi^2|f)}{\mathcal{P}(\chi^2)}\mathcal{P}(f)\,Df,\\ &= \frac{1}{N}\sum_{k=1}^{N}\frac{\mathcal{P}(\chi^2|f_k)}{\mathcal{P}(\chi^2)}\mathcal{O}[f_k]\,,\end{aligned} \tag{5.52}$$

where in the last line we used Eq. (5.39). It follows that we can sample the probability density $\mathcal{P}(f|\chi^2)$ using the replicas f_k, but reweighted:

$$\langle\mathcal{O}\rangle_{\rm new} = \tfrac{1}{N}\sum_{k=1}^{N} w_k\,\mathcal{O}[f_k]\,, \tag{5.53}$$

where

$$w_k = \frac{\mathcal{P}(\chi^2|f_k)}{\mathcal{P}(\chi^2)} = \frac{P(f_k|\chi^2)}{P(f_k)}. \tag{5.54}$$

the second equality coming from another application of Bayes' theorem. Importance sampling of the old probability distribution Eq. (5.39) guarantees that all the $P(f_k)$ are equal (to $1/N$): the weights w_k are thus the relative probabilities of the replicas given the χ^2 to the new data. Combining Eq. (5.54) and Eq. (5.44), and noting that the density $\mathcal{P}(\chi^2)$ can be taken out of the sum,

$$w_k = \mathcal{N}_\chi(\chi^2(y,f_k))^{n/2-1}e^{-\frac{1}{2}\chi^2(y,f_k)}, \tag{5.55}$$

where $\mathcal{N}_\chi$ is a normalization factor. Applying Eq. (5.52) to the unit operator, $\langle 1\rangle_{\rm new} = 1$, so $\sum_{k=1}^{N} w_k = N$, and

$$\mathcal{N}_\chi = N\Big/\sum_{k=1}^{N}(\chi^2(y,f_k))^{n/2-1}e^{-\frac{1}{2}\chi^2(y,f_k)}, \tag{5.56}$$

consistent with Eq. (5.54) and Eq. (5.47).

One can therefore re–evaluate the average, the standard deviation and the correlation for all uncertainties using the new probability density instead of the old one by simply replacing the average in Eq. (5.39) with the weighted average in Eq. (5.53). The same can be done for the evaluation of the uncertainties and correlations. In particular if $\mathcal{O}$ is one of the PDFs, one can re–evaluate the PDFs after the inclusion of the new data and use the re–weighted PDF set into the analysis rather than the PDF set used to evaluate the weights.

A useful measure of the effectiveness of the reweighting is

$$N_{\rm eff} \equiv \frac{(\sum_{k=1}^{N} w_k)^2}{\sum_{k=1}^{N} w_k^2} = \frac{N^2}{\sum_{k=1}^{N} w_k^2} : \tag{5.57}$$

this gives the number of replicas left after the reweighting. Clearly, $0 < N_{\rm eff} < N$. If $N_{\rm eff}$ becomes too low, the reweighting procedure will no longer be reliable, either because the new data contain a lot of information on the PDFs, necessitating a full refitting, or because the new data are inconsistent with the old. These two cases can be distinguished by examining the χ^2 profile of the new data: if there are very few replicas with a χ^2 per degree of freedom of order unity, the errors in the new dataset have probably been underestimated.

5.3.2 Phenomenological applications

First of all, as a cross-check of the procedure, the reweighting procedure has been tested in a well–known situation. The NNPDF2.0 reference fit includes a large number of DIS, Drell–Yan and inclusive jet data, as it is illustrated in Chap. 4. As a test of the reweighting procedure within the NNPDF approach, we compare the NNPDF2.0 reference fit to the one obtained by fitting only the DIS and the Drell–Yan data (which we call it NNPDF2.0_DYP+DIS) and including inclusive jet data through Bayesian reweighting. Given the consistency of the inclusive jet data with the DIS and Drell–Yan data included in the NNPDF2.0 analysis [71], we expect the refitting and the reweigthing to provide statistically equivalent results.

Results are shown in Fig. 5.21. The re–weighted PDFs reproduce exactly the parton shapes obtained in the NNPDF2.0 reference fit. In particular the error of the medium and large–x gluon is reduced by the inclusion of the inclusive jet data, while the other PDFs remain more or less the same. In Fig. 5.22 the distribution of the χ^2 per degree of freedom, $\chi^{2,(k)}_{\rm jets}/N_{\rm dat}$, over the ensemble of 1000 replicas and the corresponding distribution of weights are displayed. We see that the distribution of the χ^2 is peaked between 1 and 2 and it is pretty narrow about that region. Consequently the weights

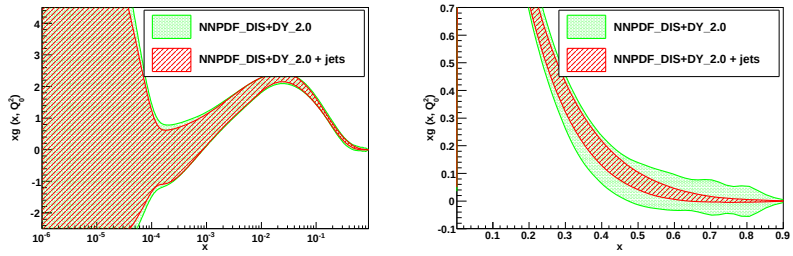


Figure 5.21: Small–x (left) and large–x gluon distribution at the initial scale $Q_0^2 = 2$ GeV^2. The gluon is evaluated before and after the inclusion of the jet data by reweighting.

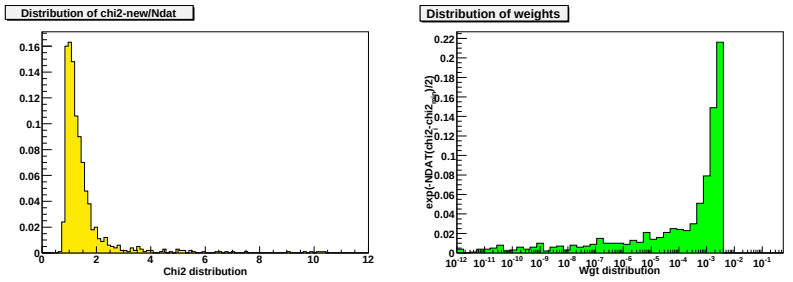


Figure 5.22: Distribution of the $\chi^{2,(k)}_{\rm new}$ and the weights in the reweighting of the jets inclusive cross–sections on the NNPDF20_DIS+DYP_1000 PDF set.

are mostly concentrated about $1/N_{\rm rep}$. The computation of the effective number of replicas ($N_{\rm eff} = 443$ in this case) confirms what we can infer from Fig. 5.22, i.e. that half of the replicas constribute to the redefinition of the new probability density.

Once the method of Bayesian reweighting has been cross–checked, it can be used in several contexts. One possibility is the study of the potential impact of the Tevatron lepton asymmetry data at Run II. This analysis is interesting in several respects. First of all, it allows us to study the compatibility between the W lepton asymmetry data and the other data included into the NNPDF2.0 analysis, namely the CDF W asymmetry data [95] and the DIS structure functions data and the low–mass Drell–Yan data. These data are found to be problematic within other global analyses, see for instance Ref. [201]. Secondly, it enables us to assess the impact of these data in further constraining the parton densities. In proton–antiproton scattering, $W^{\pm}$ bosons are produced mainly by the annihilation of a u(d) quarks in the proton with the $\bar{d}(\bar{u})$ in the anti–proton. Any asymmetry in the W^+ and W^- rapidity distributions is the result of a difference between the u and d distributions in the proton. Indeed the W

charge asymmetry at the Tevatron is defined as

$$A_W(y_W) = \frac{d\sigma(W^+)/dy_W - d\sigma(W^-)/dy_W}{d\sigma(W^+)/dy_W + d\sigma(W^-)/dy_W} \sim \frac{u(x_1)d(x_2) - d(x_1)u(x_2)}{u(x_1)d(x_2) + d(x_1)u(x_2)}, \quad (5.58)$$

where the last equality is valid only at leading–order neglecting the sea contributions and $x_{1,2} = \frac{M_W}{\sqrt{S}} \exp(\pm y_W)$. The $W^\pm$ bosons are detected through their decays into a $l\nu_l$ pair. Due to the unknown longitudinal momentum of the neutrino, the boson rapidity y_W cannot be directly measured, unless some extra information of the transverse energy of the resolved lepton and on the missing energy is used on a event–by–event basis, as in a recent CDF analysis [95]. What is typically measured, see e.g. Refs. [146, 145, 147], is the lepton charge asymmetry which is a convolution between the W boson production asymmetry and the parity violating asymmetry from the W boson decay

$$A(\eta_l) = \frac{d\sigma(l^+)/d\eta_l - d\sigma(l^-)/d\eta_l}{d\sigma(l^+)/d\eta_l + d\sigma(l^-)/d\eta_l}, \quad (5.59)$$

where η_l is the pseudo–rapidity of the charged lepton. If one defines the emission angle of the charged lepton relative to the proton beam in the W rest frame by $\cos\theta_R = 1 - 4E_T^2/M_W^2$, being E_T the transverse energy of the charged lepton, the lepton pseudo–rapidity and the W rapidity are related by

$$\eta_l \; = \; y_W + \frac{1}{2}\ln\left(\frac{1+\cos\theta_R}{1-\cos\theta_R}\right). \quad (5.60)$$

Since at the edge of the phase space for $\cos\theta_R \sim \pm 1$ the leading sea contribution $\bar{u}\bar{d}$ is enhanced relatively to the valence–valence contributions, on top of measuring the shapes of the up and down quarks, the lepton charge asymmetry probes the separation into valence and sea quarks. The enhancement, and therefore the constraint on the separation, is bigger near the edge of the phase space, for small E_T. This is one of the reasons why in some experimental analyses [146, 147], the asymmetry is measured in different bins of E_T.

In the NNPDF analysis only the W boson asymmetry data of Ref. [95] is included. They are implemented at next–to–leading order without relying on any K–factor approximation thanks to the FastKernel method introduced in Ref. [71]. The other datasets are not included in the analysis due to the lack of a fast implementation. The recent development of the APPLGRID [202] interface is likely to facilitate the future inclusion of these data directly in our fits. However we can already study the impact of their inclusion into the analysis through the reweighting technique. In this analysis we consider the sets of measurements performed by the D0 collaboration and published in Refs. [146, 145].

In Ref. [146] a measurements of the electron charge asymmetry in $p\bar{p} \to W + X \to e\nu + X$ events at a centre–of–mass energy of 1.96 TeV using 0.75 fb^{-1} of data collected with the D0 detector at the Run II of the Tevatron collider is presented. The asymmetry is measured as a function of the electron pseudo–rapidity η in the interval $|\eta_e| < 3.2$ with the following cuts

$$\not{E} > 25 \text{ GeV} \qquad |\eta_e| < 3.2 \qquad |M_T| < 50 \text{ GeV}.$$

Three sets of measurements in 12 bins of the electron pseudo–rapidity which differ because of the different cuts imposed to the transverse energy of the electron

$$E_T > 25 \text{ GeV (bin A)} \quad 25 \text{ GeV} > E_T > 35 \text{ GeV (bin B)} \quad E_T > 35 \text{ GeV (bin C)}.$$

Here I only study the inclusion of the bin A data. The analysis will be extended to the other two bins in the near future. In Ref. [145] a measurement of the muon charge asymmetry from W boson decay using 0.3 fb^{-1} of data collected at $\sqrt{s} = 1.96$ GeV between 2002 and 2004 with the D0 detector at the Tevatron collider is presented. The measurements are performed in the $|\eta_\mu| < 2$ pseudo–rapidity range and the following cuts are applied: $p_{T,\mu} > 20$ GeV and $M_T > 40$ GeV. The range in η_μ constrains the PDF in the $0.005 < x < 0.3$ $x-$region. There is a more recent set of data introduced in Ref. [203] but the data are not public yet.

In what follows, we use the DYNNLO code [204] to compute the theoretical predictions for the lepton asymmetries. The code is a parton level Monte Carlo program designed to compute the cross–section distributions for the vector–boson colliders in the proton–proton and proton–antiproton collisions. It calculates exclusive processes and it enables the user to implement the same cuts and the same isolation as those implemented in the experimental analyses.

The predictions obtained with the next–to–leading order computation with DYNNLO and several sets of PDFs are compared in Fig. 5.23 for the considered datasets. The error bands refer only to the PDF errors. The latter are evaluated by running $N_{\rm mem}$ times the DYNNLO code, once for each member of input parton distribution, where $N_{\rm mem} = 44$ for NLO CTEQ6.6 [41], $N_{\rm mem} = 40$ for the NLO MSTW08 [43] and $N_{\rm set}$ =1000 for NNPDF20 [71]. All sets are available in the common LHAPDF interface [137]. It is quite remarkable that, even if these data are not included into the NNPDF2.0 fit, the prediction obtained from the NNPDF20_1000 is much closer to the experimental data than the predictions obtained with the other parton sets.

In this analysis, we first reweight one data set at the time, in order to study the impact of each of the two separately. In Fig. 5.24 the effect of the addition of the $D0$ electron (bin A) data through reweighting into the NNPDf20 fit (prior probability). One can see that the effect consists in a slight reduction of the error of the Singlet and the

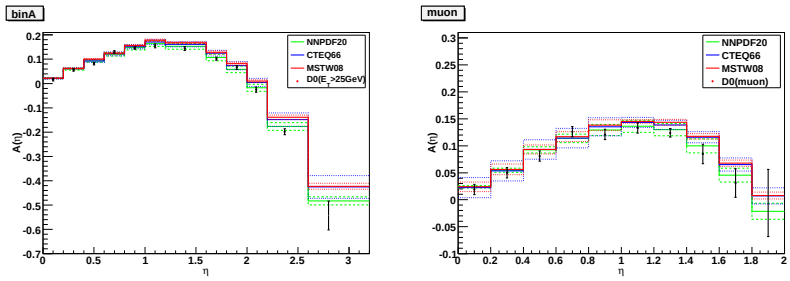


Figure 5.23: Prediction for the D0 W electron and muon asymmetries [146, 145] (left, electron binA: $E_T > 25$ GeV; ; right, muon $E_T > 20$ GeV) obtained with DYNNLO at next-to-leading order, using the NNPDF20_1000 [71], CTEQ6.6 [41] and MSTW08 [43] parton sets. The uncertainty are the PDFs uncertainty, evaluated according to the HEPDATA recipe for MSTW08 and CTEQ6.6 and with the Monte Carlo prescription for the NNPDF2.0 set reported in Appendix C.

Total Strangeness PDFs in the small–x region and a more sensible reduction of the uncertainty of the total Valence in the small and middle$-x$ region. Instead the effect of the addition of the muon data through re-weighting is pretty mild and it does not produce any effect on the PDF shapes or errors, as one can see in Fig. 5.25 in the case of the Singlet PDF.

In order to understand the difference between the two sets, in Fig. 5.26 the histogram with the distribution of $\chi^{2,(k)}_{\rm new}$ over the 1000 replicas and the corresponding distribution of weights are displayed for both sets. For these sets

$$\begin{aligned} N_{{\rm dat},e} &= 12, \qquad N_{{\rm dat},\mu} = 10, \\ N_{{\rm eff},e} &= 266, \qquad N_{{\rm eff},\mu} = 758. \end{aligned}$$

While for the D0 electron data, the number of effective replicas reduces to a quarter of the original set, still remaining significant, and the distribution of the weights is pretty spread, for the D0 muon data, $N_{\rm eff}$ is much bigger. However, the weights for this dataset are distributed about 10^{-3}, according to a very narrow distribution, which means that the reweighting does not have much effect.

In Fig. 5.27 we have evaluated the average and the standard deviation of the observable before and after the re-weighting, i.e. with the old and with the new probability distributions respectively. We notice that in both cases the predictions get closer to

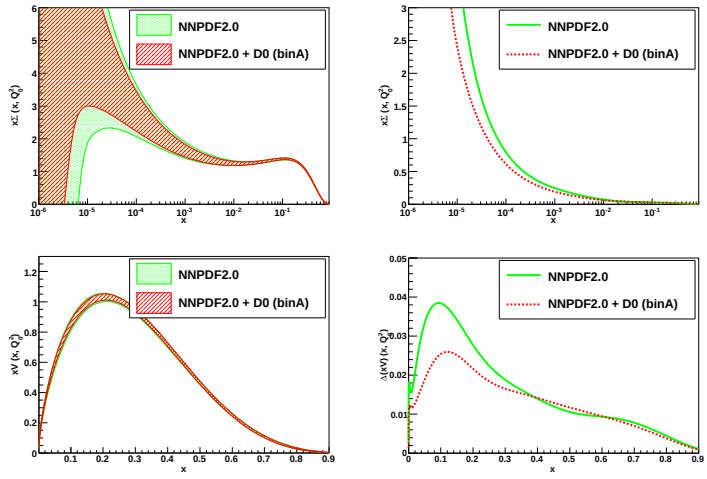


Figure 5.24: Singlet (top) and Valence (bottom) PDFs and their absolute errors at $Q^2 = Q_0^2$ for NNPDF2.0 and NNPDF2.0 + D0 electron data (bin A) added by reweighting.

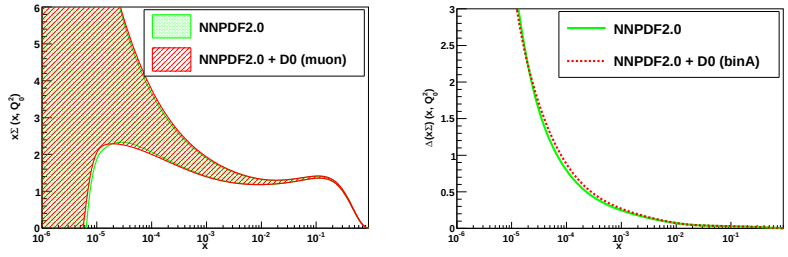


Figure 5.25: Singlet PDF and its absolute error at $Q^2 = Q_0^2$ for NNPDF2.0 and NNPDF2.0 + D0 muon data added by reweighting

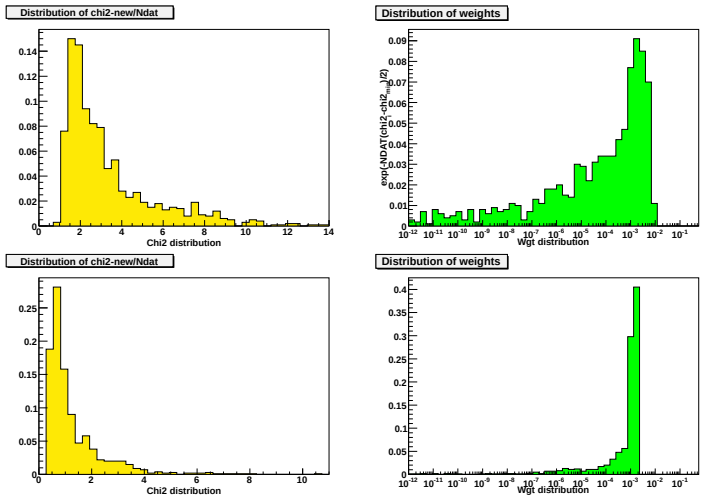


Figure 5.26: Distribution of the $\chi^{2,(k)}_{\rm new}$ and the weights in the reweighting of the D0 W electron asymmetry (top) and of the D0 muon (bottom) using the NNPDF20_1000 set as a initial probability distribution.

the data, while only for the reweighting of the D0 electron data the PDF uncertainty is substantially reduced. This confirms what we said earlier. To conclude, if we evaluate the χ^2 per degree of freedom of the experiments already included in the NNPDF2.0 analysis before and after the addition of the $D0$ data through the reweighting procedure, we see that none of the experiments displays a deterioration of the χ^2 due to the addition of these data. This is a sign that there is no manifest incompatibility between these data and the other Drell–Yan, jets and DIS data included into the NNPDF2.0 analysis. More care must be adopted when we will be considering the data in separate bins (bin B and bin C) because these data might display incompatibility and therefore be more problematic to include.

The technique that I have applied to a set of real data, might be applied to any other dataset or to any pseudo–dataset. The method just discussed enables one to easily assess the potential of future experiment to further constrain the parton content of the nucleon without relying PDFs fitting collaborations. For this reason, it is particularly handy for external users and easily generalisable.

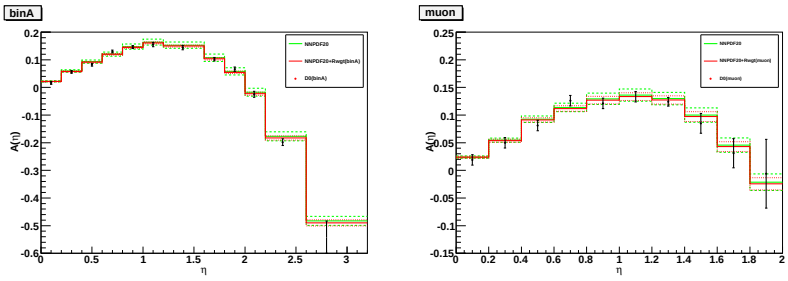


Figure 5.27: W electron asymmetry computed on the NNPDF2.0 set before and after the reweighting of the D0 W electron asymmetry (left) and D0 muon asymmetry (right).

Chapter 6

Processes with heavy quarks in the initial states

As it has been discussed in Chap. 2, there are two possible ways of performing a calculation of a high-energy process involving heavy quarks. One option is the most straightforward from the conceptual point of view. The heavy quark mass is taken to be of the same order of magnitude as the other hard scales involved in the process. The heavy quark does not contribute to the proton wavefunction and can only be generated as a massive final state. In practice the theory which is used in the massive calculation is an effective theory with n_l light quarks, where the heavy quarks are decoupled and do not enter in the computation of the running coupling constant and in the evolution of the PDFs. Alternatively, one may consider the heavy quark masses much smaller than the other scales involved in the process and consequently ignore them. The heavy quarks are treated as massless partons which do constitute the hadron and may appear in the initial state.

Both schemes present several advantages and disadvantages and can be applied in complementary regimes depending on the relative size of the heavy quark masses. In the massless scheme the calculation is highly simplified and potentially large logarithms $\mathcal{O}(\mu^2/M_Q^2)$ due to the collinear splitting of the initial heavy quarks and gluons are consistently resummed in the heavy quark PDF. Analogously logarithms of $\mathcal{O}(p_{T,Q}^2/M_Q^2)$ appear due to the heavy quark splitting in the final state and they may be resummed in the fragmentation functions or absorbed into a suitable definition of inclusive heavy quark jets. In the massive scheme instead, the computation is more complicated due to the addition of massive final states; however the kinematic description of the heavy quark is correctly taken into account at the leading order and can be

coherently studied at the next–to–leading order. In the latter case, the implementation in parton shower codes is also straightforward. The downside is that possibly large logarithms developing both in the initial and the final states are not resummed.

To all orders in perturbation theory the two schemes are identical, but the way of ordering the perturbative expansion is different and at a finite order the results do not match exactly. For some processes the difference between calculations performed in the two schemes may be very significant at the leading order, yielding to predictions which might differ up to an order of magnitude. One of the most famous and glaring example is the discrepancy observed in the inclusive Higgs production initiated by b quarks. For this reason matching schemes based on the combination of the massless and massive computations with the suitable subtraction of the double–counting were proposed [46, 51, 205, 53].

In this work we critically reconsider the motivations for using schemes (massless or improved) that allow the resummation of potentially large logarithms, by performing a thorough analysis of their origin and relevance in various processes. The first section is devoted to a brief overview of the studies which have been performed in literature. In the following sections we analyse in detail some representative processes that can be treated analytically and cover a broad spectrum of possibilities. We start by considering the b quark production at DIS in both schemes. In this case the scales involved in the process are the mass of the bottom quark, the virtuality of the intermediate vector boson and the energy of the partonic process. Then we move to a more general case where the masses of the produced quarks are different, as in the case of the single top production. Finally, we consider the W boson production associated to a charm quark. For each of the above processes we proceed as follows. We first go through the massive calculation and determine analytically the scale associated to the collinear emission of the heavy quarks. Then we study the associate cross–section distributions and assess the size of the logarithms which are resummed in the evolution of the heavy parton distribution. This allows us to carefully estimate the size of the potentially large logarithms and therefore quantitatively assess their impact in predictions for processes at the LHC.

6.1 Phenomenological review

In this section I give a general overview of the impact that the choice of different heavy quark schemes has on the predicted cross sections. This topic has been broadly discussed in literature and this section cannot provide a comprehensive review. Rather, it aims to describe the context where the analysis which we have performed finds its natural location and inspiration. In particular, we mention some studies performed in

literature on the DIS production of heavy quarks, the associated production of heavy quarks and vector or Higgs bosons and we finally mention the single top process.

The production of heavy quarks in Deep–Inelastic lepton–hadron scattering has been extensively analysed from both the theoretical and the experimental points of view. Calculations have been performed in the massive scheme, where the heavy quark and

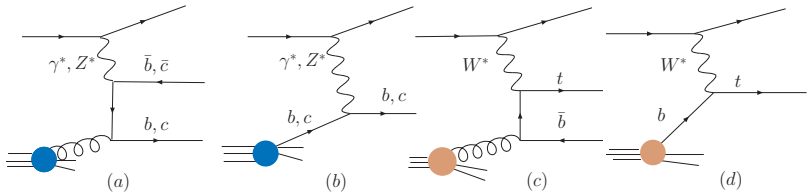


Figure 6.1: Leading–order diagrams for heavy quark production in the massive (a) and massless (b) scheme. Leading order diagrams for single top production in massive (c) and massless (d) schemes.

anti–quark appear in pairs and are produced via photon–gluon fusion, as it is shown in Fig. 6.1 (a). The LO calculation starts at $\mathcal{O}(\alpha_s)$ and has been evaluated at the end of the seventies (see for instance Ref. [206]). The $\mathcal{O}(\alpha_s^2)$ corrections have been evaluated in the nineties [44]. The same collaboration has provided a code [207], which integrates the heavy quark contributions to the fully differential heavy structure functions $F_i^{b,c}(x, Q^2)$ over x and Q^2 to produce rates and differential distributions relevant for DIS charm and bottom production. We'll make use of this code in the following analyses. In addition, other methods to study heavy flavor production were advocated, like the intrinsic quark approach [208] and the variable flavor number scheme [46], whose leading–order contribution is drawn in Fig. 6.1 (b). The latter was one of the first processes evaluated according to the so–called ACOT scheme, discussed in Chap. 2. The main difference between the two production mechanisms which describe the LO contribution in case of massive or massless (as well as VFNS) calculations can be attributed to the fact that for massive heavy quark production the quark and the anti–quark are produced in pairs, while in the other approach only one heavy quark is produced at the leading–order. At NLO, however, these striking differences are milder as, for instance, the gluon splitting process appears also in the massless scheme. The experiments carried out at HERA, which have provided a wealth of information about charm and bottom production [82], indicate that the bulk of the heavy quark structure function in the region explored by HERA is given by gluon–initiated processes.

On the other hand in Ref. [209] the size of the logarithmic terms of (Q^2/m_Q^2) has been assessed by comparing the fixed–order massive calculation at NLO to an asymptotic

calculation which only contains $\log^n(Q^2/m_Q^2)$ and $\log^n(\mu^2/m_Q^2)$ and terms which survive in the limit $Q^2 \to \infty$. In this analysis it was shown that the logarithms mentioned above dominated the structure function, in the sense that the ratio between the asymptotic calculation where these logarithms are resummed to all orders and the logarithms included at fixed order in the massive calculation is large, except in the narrow threshold region where $s = Q^2(1-z)/z \sim 4m_Q^2$. In the same reference, to answer the

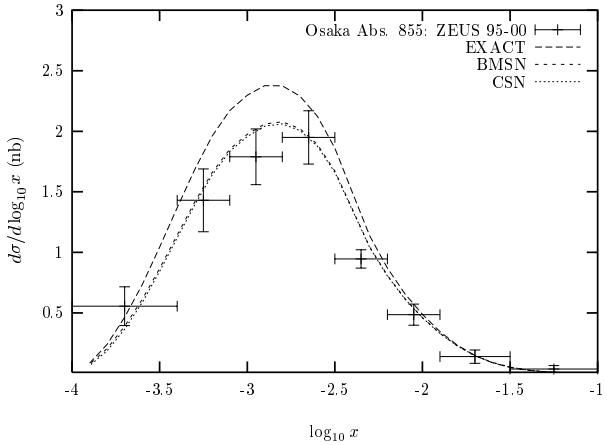


Figure 6.2: The combined Osaka and published ZEUS-data [210] for $d\sigma/d\log_{10}x$ (nb) for DIS production of $D^{*,\pm}$-mesons. The dashed line is the NLO fixed flavor number scheme result $F_{2,c}^{\rm EXACT}$, from the program HVQDIS [207]. The dotted line (BMSN–scheme [205]) and dashed-dotted line (CSN–scheme [211]) are based on a variable flavor number scheme computation $F_{2,c}^{\rm VFNS}$. Taken from Ref. [209].

question whether these logarithms bedevil the convergence of the perturbation theory, a comparison was performed between the massive calculation and the variable–flavor number calculation performed in the BMSN scheme [39]. It was observed that, in spite of their dominance, these logarithms do not vitiate the convergence of the perturbation series so that their resummation in the heavy quark PDF is in principle not necessary. Indeed, as it is shown in Fig. 6.2, it is actually hard to distinguish between the two approaches except in the vicinity of $x_B = 10^{-3}$, where the VFNS seems to describe better the data.

The second class of processes that we consider is the associated production of heavy quarks with electroweak gauge or Higgs bosons. Among the various channels which might be exploited to search for Higgs bosons at hadron colliders, Higgs radiation off bottom quarks is of special interest [212]. This process is the dominant Higgs-

boson production mechanism in supersymmetric theories at large $\tan\beta$, where the bottom–Higgs Yukawa coupling is strongly enhanced[1]. In a four-flavour scheme with

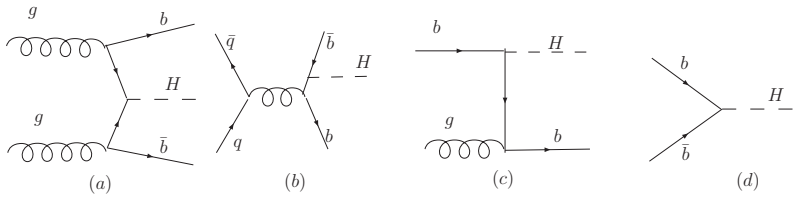


Figure 6.3: Leading–order diagrams for: (a),(b) inclusive Higgs production in the four–flavor scheme, (c) inclusive Higgs plus 1 b–jet production in the five–flavor scheme, (d) inclusive Higgs production in the five–flavor scheme.

no b quarks in the initial state, the lowest-order QCD processes for associated $b\bar{b}H$ production are gluon–gluon fusion, Fig. 6.3 (a), and quark–antiquark annihilation[2], Fig. 6.3 (b)

$$gg \to b\bar{b}H \quad \text{and} \quad q\bar{q} \to b\bar{b}H\,. \tag{6.1}$$

As we mentioned earlier, the massive approach does not resum the logarithmic contribution of the bottom mass, (Q^2/m_b^2), being Q the typical scale of the process. However, these terms can be summed to all orders in perturbation theory by introducing bottom parton densities (the five-flavour scheme) [214]. For the inclusive Higgs production cross–section, where no b-jet is necessarily tagged in the final state, the leading order diagram of the five–flavor calculation is drawn in Fig. 6.3 (d). The leading order diagram of the four–flavor scheme, drawn in Fig. 6.3 (a), enters in the five–flavor calculation only at NNLO. In the latter scheme, the incoming b partons are given zero transverse momentum at leading order, and acquire transverse momentum at higher orders. If on the other hand one demands that at least one b-jet is observed ($p_T^b > p_T^{\rm cut}$; $|\eta^b| < \eta^{\rm cut}$), then the leading parton process in the five-flavour scheme is $gb \to bH$, drawn in Fig. 6.3 (c), and the leading order contribution of the four–flavor schemes enters into the five–flavor one at NLO. Finally, if one compares the theoretical prediction with the experimental cross–section for Higgs plus two b–jets ($p_T^{b,1}, p_T^{b,2} > p_T^{\rm cut}$; $|\eta^{b,1}|, |\eta^{b,2}| < \eta^{\rm cut}$), the two schemes include the same diagrams at the leading–order in perturbation theory (apart from the b–initiated contributions included only in the five–flavors scheme, which are however completely negligible).

[1]The parameter $\tan\beta = v_2/v_1$ is the ratio of the vacuum expectation values of the two Higgs fields generating the masses of up- and down-type particles in supersymmetric extensions of the SM.

[2]The quark–antiquark annihilation is not considered in the following discussion given that it is unrelated to the choice of the four versus five–flavor schemes and that its contribution is very small [213].

The case of the fully–inclusive cross–section is particularly interesting. In Fig. 6.4

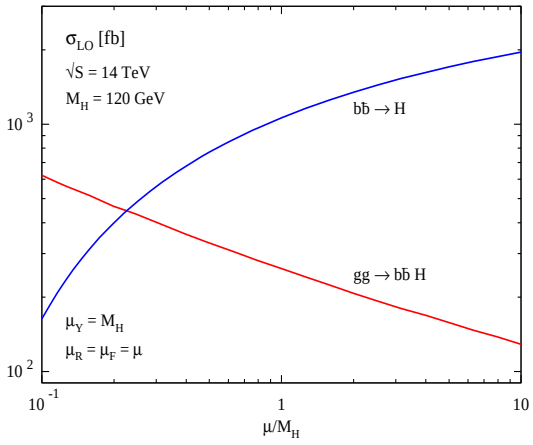


Figure 6.4: Scale variation of the LO inclusive Higgs production cross section prediction for $pp \to b\bar{b}H + X$ at the LHC in the four- and five flavour schemes. Taken from Ref. [213].

we display the LO prediction for the total $b\bar{b}H$ cross secion at the LHC in the two schemes as a function of the renormalisation and factorisation scales [213]. Both calculations exhibit a strong scale dependence. The scale dependence goes in opposite directions in the two cases, being driven by the α_s scale dependence in the four–flavor scheme and by the b PDF in the five–flavor one. The leading order scale analysis was confirmed by the explicit calculation of the NNLO in the massless scheme and NLO in the massive one, even though the discrepancy is slightly reduced. From the discussion above it is clear that at least for the fully–inclusive case, a “fair" comparison between the two schemes could only be done at NNLO for the massless and NLO for the massive one.

In the leading–order analysis, setting the renormalisation and factorisation scales to $\mu = M_H$, the five-flavour scheme prediction exceeds the four-flavour scheme prediction by more than a factor of 5. For this reason, in Ref. [215] the massive and massless approaches were combined by subtracting the double counting, according to the so called simplified ACOT scheme [47]. This matching is meant to keep the best characteristics of the two approaches. However this still does not clarify the origin of the large discrepancy observed in Fig. 6.4. In Ref. [216] it was suggested that the calculation of $b\bar{b} \to h$ may overestimate the inclusive cross section, due to crude approximations inherent in the kinematics, which give rise to large bottom-quark mass and phase-space effects. However in Ref. [217] the massless calculation is performed us-

ing the ACOT scheme which does not approximate the kinematics of the final bottom pair and the same discrepancy is observed. The matched and the massive calculations were then separately studied in more details in the same paper [217]. Looking at the collinear plateau of the five–flavor matched $b\bar{b} \to H$ calculation, it was argued that the choice of the factorisation scale $\mu_F = M_H$ is not the best choice given that the collinear plateau of the cross–section as a function of the collinearity $|t|$ drops much earlier. It was found instead that a choice such as $\mu_F = M_H/4$ leads to a reduced scale dependence in both massive and massless approaches and that the discrepancy between the four and five–flavor schemes was reduced by a factor of two, as one can also infer from Fig. 6.4. This suggests that the scale at which the gluon splits is softer than the scale of the hard process where the Higgs is produced. However the question about what is the dynamical origin of that is still unanswered.

In Ref. [218] a similar process was considered at next–to–leading order: the production of super symmetric charged Higgs boson in association with a top quark. In a four–flavor scheme the leading order process is $gg \to \bar{b}tH^-$ while in a five–flavor scheme the leading order process is $gb \to tH^-$. In this study, a heuristic method for identifying the dynamical scale of the collinear splitting of the gluon into a bottom pair was employed by analysing the distribution of the differential cross–section in the transverse momentum and in the virtuality of the bottom quark. More recently in Ref. [219], the NLO calculation in the four–flavor scheme has been fully carried out and a comparison between the results obtained at NLO in the four–flavor scheme have been compared to the NLO results in the five–flavor scheme. The result of this comparison is shown in Fig. 6.5. One sees that, even taking the scale uncertainty into account, the cross sections evaluated in the two scheme at NLO are barely consistent; independently on the charged Higgs boson mass, the central predictions in the five–flavor scheme are larger than those of the four–flavor one by approximately 40%. Recently, at the LHC Higgs working group, the same comparison was shown by using the MSTW2008 PDF input set [43] instead of the MRST2004 one [65], and the distance between the two predictions was slightly smaller, still remaining pretty significant. This requires a better understanding of its origin.

Processes whose final states are the W or Z boson plus one or two heavy quarks are also relevant at the Tevatron and the LHC. They not only provide the dominant background to the study of $t\bar{t}$, single top, as well as Higgs production, but also represent a useful benchmarks for the validation of the theoretical description of heavy quark jets at hadron colliders. The problem of the choice of the scheme is also present for this class of processes, since calculations may be performed in several ways and for some observables it is convenient to match the massive and massless calculations by adopting one of the variable flavor number schemes discussed in Chap. 2. In Table 6.1 we summarise the state of the art of these calculation indicating the scheme in which

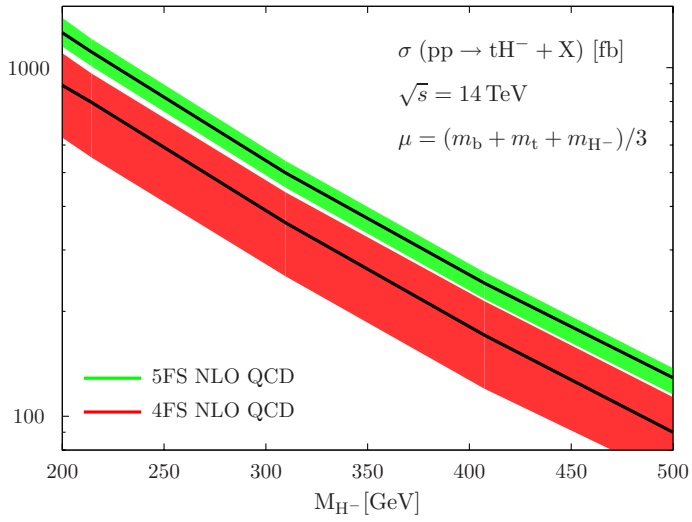


Figure 6.5: Left: Total NLO cross section for $pp \to tH^- + X$ at the LHC as a function of the Higgs mass in the four–flavor and five–flavor scheme. Central prediction and the scale dependence for $\mu_0/3 < \mu < 3\mu_0$, where $\mu_0 = (M_t + m_b + M_{H^-})/3$ in the four–flavor scheme, while $\mu_0 = (M_t + M_{H^-})/5$ in the five–flavor scheme. Input PDF set MRST2004 [65]. Taken from Ref. [219].

the NLO calculation has been performed and provide some references. Each calculation corresponds to a given experimental signature indicated on the last column of the table.

In the same class of processes one might consider the inclusive W production and the Wc associated production. The latter has been calculated in Ref. [227] using a three–flavor scheme where the charm is treated as a massive final state. The NLO result is found to be very sensitive to the choice of the factorisation and renormalisation scales by varying them about M_W. The authors interpreted this dependence as a sign that this might be attributed to the fact that in a massive approach the collinear logarithms resummed in the c PDF are neglected and they perform some analysis in order to assess the size of these logarithms. Since $c\bar{s}$ contributes to the W production rate amounts to 5% of the total W rate, this process is very interesting from the phenomenological point of view. The difference between the predictions obtained with the three–flavor scheme, where the leading order process is $cs \to W$, and the massive approach, where the leading order process is $sg \to Wc$, has been estimated to be of about 50% at the Tevatron energy if the factorisation scale is set to M_W, which leads to a difference in the prediction of the W production rate of about 2-3%, comparable to the error due to

Process	Scheme	Reference	Experimental signal
$pp \to ZQQ$ NLO	4F,5F ($m_Q = 0$)	[220]	Zjj with 2 Q–tags
$pp \to ZQQ$ NLO	4F	[221, 222]	Zjj with 2 Q–tags
$pp \to ZQ$ NLO	VF	[223]	Zj with 1 Q–tag
$pp \to ZQj$ NLO	5F	[224]	Zjj with 1 Q–tag
$pp \to Wbb$ NLO	4F ($m_b = 0$)	[220]	Wjj with 2 b–tags
$pp \to Wbb$ NLO	4F	[221, 222]	Wjj with 2 b–tags
$pp \to Wb$ NLO	VF	[225]	Wj with 1 b–tag
$pp \to Wbj$ NLO	5F	[226]	Wjj with 1 b–tag
$pp \to Wc$ NLO	3F	[227]	Wj with 1 c–tag
$pp \to Wc$ NLO	VF	[228]	Wj with 1 c–tag

Table 6.1: Available next–to-leading calculation of processes with a weak boson plus one or more heavy jets in the final states. For each process the scheme in which the calculation is performed, the reference and the experimental signatures are indicated. nF and VF stand for n–flavor and variable flavor number schemes respectively.

the scale variation. Analogously the total Z production rate receives a contribution of about 5% from the gluon–gluon fusion followed by bottom–anti-bottom annihilation $gg \to b\bar{b}Z$ [229]. Hence, to get a 1% accuracy on the total Z production, it must be under control at 20% level. In the next section we are going to consider explicitly the associated Wc process due to the simplicity of the calculation which allows us to gain a better analytical control.

Finally the choice of the heavy quark scheme in single top production was studied in details in Refs. [230, 231], where also the production of a heavy fourth generation quark t', b' was considered. In this study, it was shown that the central cross sections predicted by the $2 \to 2$, Fig. 6.1 (d), and $2 \to 3$ processes, Fig. 6.1 (c) differ by 5% or less, both at the Tevatron and at the LHC, for masses around the top quark, see Fig.6.6. At the Tevatron, the difference is well within the combined uncertainty from higher orders and PDFs, so they concluded that the two calculations are consistent. At the LHC (10 TeV) the consistency was found to be marginal. For larger masses, *i.e.* for t' production, the differences were found to be much larger. For a t' of mass of 1 TeV, the $2 \to 2$ prediction using the CTEQ6.6 PDF set is almost twice as large at the Tevatron and 20% larger at the LHC. Therefore, for such large top masses it could well be that the logarithm that is implicitly resummed in the bottom quark distribution function might become relevant or that an even smaller factorisation scale should be used.

Interestingly, the first comparison between the massless and the massive approach was performed long ago in a study of the charged Higgs production [232]. The six–flavor

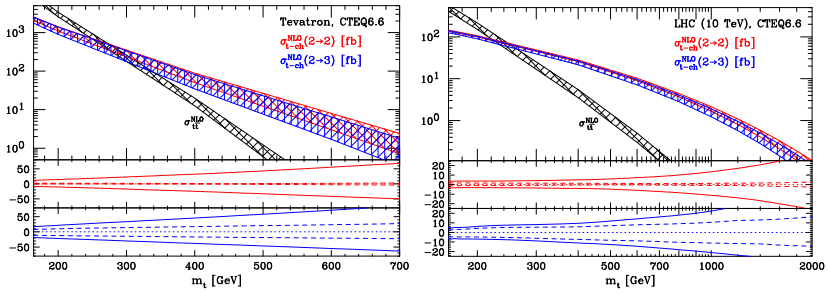


Figure 6.6: Cross sections (fb) at the Tevatron Run II (left) and LHC 10 GeV (right) for the sum of top and anti-top quark production in the t channel, as a function of the top mass obtained with the CTEQ6.6 PDF set and $V_{tb} = 1$ in the $2 \to 2$ and $2 \to 3$ schemes. Bands are the total uncertainty (scale+PDF). In the lower plots, dashed is scale uncertainty, solid is scale + PDF. Taken from Ref. [230].

calculation $b\bar{t} \to H^-$ was compared to a five–flavor calculation whose leading–order process is $bg \to H^- t$. A large difference between the two approaches was observed for the top masses lying in the range expected at those times, i.e. $m_t \in [3, 300]$ GeV. The size of the difference increased with the mass of the top.

All these studies raise the question about understanding what is the scale of the logarithms and what is their dynamical origin. This is what we investigate in the following analyses.

6.2 DIS heavy quarks production

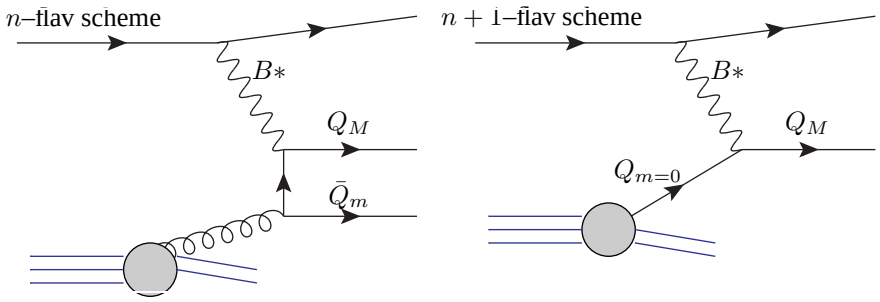
In this section we study the production of heavy quarks in DIS collisions. This class of processes includes both the production of heavy pairs of quark–antiquarks and the single top production. All calculations and results can be easily generalised to the hadron–hadron production of heavy quarks, as we will explicitly see in what follows.

We consider the general case where there are two heavy quarks whose masses M and m do not have to be equal, with $M \geq m$. In the Feynman diagrams below we draw the leading order process in the massive, n–flavor, and massless, $(n+1)$–flavor, schemes. In the massive scheme the heavy quark of mass m, Q_m, is treated as a massive final state and does not contribute to the proton wavefunction. At the leading order the process is

$$l(p_a) + g(p_b) \;\to\; B^*(q) \;\to\; l'(k_1) + Q_M(k_2) + \bar{Q}_m(k_3), \tag{6.2}$$

where B^* is the virtual vector boson which mediates the scattering. We stress that the leading order diagram in this scheme is a next-to-leading order contribution in the $(n+1)$–flavor scheme where the heavy quark Q_m is treated as a massless quark and appears as a parton in the initial state. In this case the leading order partonic contribution is given by the process

$$l(p_a) + Q_{m=0}(p_b) \;\rightarrow\; B^*(q) \;\rightarrow\; l'(k_1) + Q_M(k_2). \tag{6.3}$$



In the latter scheme the calculation is highly simplified and potentially high logarithms due to initial state collinear splitting of the gluon in $Q\bar{Q}$ are consistently resummed in the evolution of the heavy quark PDF. However, while at very high energy this scheme is perfectly sensible, at low or intermediate values of the scale characterising the process, this scheme does not treat correctly the kinematics and at leading order it ignores power suppressed contributions of $\mathcal{O}(m^2/\mu_F^2)$. In the former scheme, the computation is more complicated due to the addition of a massive final state; however the kinematic description of the heavy quark is correctly taken into account already at leading order. The scale which discriminates between the high–energy region, where a massless approach works well and the low–energy region, where instead the massive approach works better, is the scale of the collinear logarithm which is resummed in the DGLAP evolution of the $Q_{m=0}$ PDF. This scale is often identified with Q^2 [53, 41]. In what follows, we show that the scale is not exactly given by Q^2, rather by a dynamical scale that depends on the momenta of the final states.

As discussed in Chap. 2, the scattering of the incoming lepton can be described by the kinematical variables Q^2, x_B and y defined in Eqs. (1.1) and (1.2), such that $x_B = Q^2/(S_{\rm had} y)$, being $S_{\rm had}$ the hadronic centre–of–mass energy. In Eq. (1.6) the inclusive lepton–hadron scattering cross section was written as a product between the leptonic tensor $L_{\mu\nu}$, Eq. (1.9), describing the upper part of the diagram and the hadronic tensor $H^{\mu\nu}$. The latter assumes the generic form of Eq. (1.10) where, given that the quarks in the final states are not massless, the structure function F_4 and F_6 do not vanish. However, if one neglects the masses of the leptons, they do vanish when the hadronic tensor is contracted with the leptonic one. As a result, the total DIS cross

section can be expressed in terms of the F_2, F_L and F_3 heavy structure functions as

$$\sigma_{\rm tot} = \int_{y_{\rm min}}^{y_{\rm max}} dy \int_{Q^2_{\rm min}}^{Q^2_{\rm max}} dQ^2 \frac{2\pi\alpha_{lB}\alpha_{hB}}{y(M_B^2+Q^2)^2}\Big\{[1+(1-y)^2]F_2^Q(x,Q^2,M^2,m^2) - y^2F_L^Q(x,Q^2,M^2,m^2) + [1-(1-y)^2]F_3^Q(x,Q^2,M^2,m^2)\Big\}, \quad (6.4)$$

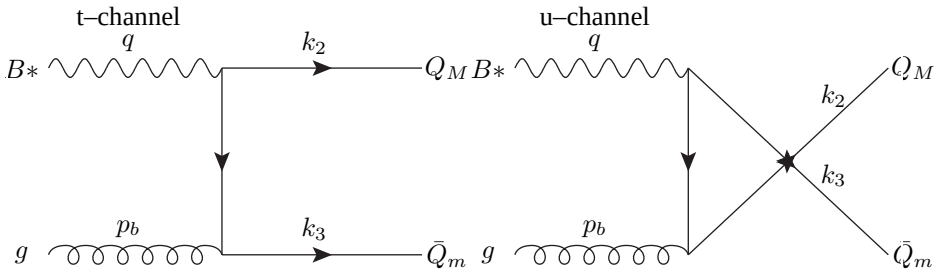
where by $F^Q_{2,L,3}$ we indicate the contribution to the structure functions coming from the heavy quark coupling to the virtual vector boson B^* (γ or Z boson, for the neutral current processes and $W^\pm$ for the charged current processes). The structure functions depend on the mass of the produced heavy quark Q_M and on the mass of Q_m. The dependence on m^2 is explicit in the coefficient functions in case of a massive calculation, while it is implicit in the definition of the $Q_{m=0}$ parton distribution in the massless calculation.

6.2.1 Massive calculation

The easiest way to evaluate the structure functions is to project the hadronic tensor in its transverse, longitudinal and axial components which are directly related to the amplitude $\mathcal{M}$ of the partonic subprocess by Eq. (1.20). From that, it is easy to derive the structure functions F_2, F_L and F_3 according to Eqs. (1.19) and (1.21). We can therefore consider only the subprocess as initiated by the B^* vector boson.

The LO partonic subprocess $B(q)+g(p_b)\to Q(k_2,M)+\bar{Q}(k_3,m)$ is drawn in the diagram below. The amplitude receives a contribution from the t–channel and the u–channel where the Mandelstam variables t and u are defined as

$$\begin{aligned} t &= (p_b-k_3)^2 = m^2-2(p_b\cdot k_3), \\ u &= (p_b-k_2)^2 = M^2-2(p_b\cdot k_2). \end{aligned} \quad (6.5)$$



It is convenient to define another Mandelstam variable s, the centre of mass energy in the boson–gluon system (not to be confused with the lepton–parton centre–of–mass energy $\hat{s} = (p_a + p_b)^2 = \xi S_{\rm had}$ where ξ is the fraction of the momentum of the incoming hadron carried by the gluon), as

$$s = (p_b + q)^2 = 2\xi(P_H \cdot q) - Q^2 = \frac{Q^2\xi}{x_B}(1 - \frac{x_B}{\xi}) \tag{6.6}$$

$$= Q^2\frac{(1-z)}{z}, \qquad \text{where} \quad z \equiv \frac{x_B}{\xi}, \tag{6.7}$$

such that

$$s + t + u = M^2 + m^2 - Q^2. \tag{6.8}$$

We might also define the mass–subtracted Mandelstam variables

$$t_1 = t - m^2 \qquad u_1 = u - M^2 \qquad s1 = s + Q^2 = Q^2/z \tag{6.9}$$

such that

$$s_1 + t_1 + u_1 = 0. \tag{6.10}$$

The partonic tensor, related to the hadronic tensor by Eq. (1.19), is given by

$$\hat{H}_{\mu\nu} = \frac{1}{2s_1}\frac{1}{2}K_{g\gamma}\left[\overline{\sum}\mathcal{M}_\mu(\gamma^* g \to Q_M\bar{Q}_m)\mathcal{M}^*_\nu(\gamma^* g \to Q_M\bar{Q}_m)d\Gamma_2 + \mathcal{O}(\alpha_s^2)\right], \tag{6.11}$$

where $1/2s_1$ is the flux factor, $1/2$ comes from the average on the initial degrees of freedom, $K_{g\gamma}$ is the color average, $\mathcal{M}_\mu$ is the leading–order amplitude for the partonic process, $d\Gamma_2$ is the two–body phase space and $\overline{\sum}$ is the average over the initial and the

sum over the final polarisations. After integrating over the azimuthal angle between the plane containing the leptons and the plane containing the initial parton and the outgoing heavy quarks, the partonic tensor (6.11) can be written as

$$\begin{aligned}\hat{H}_{\mu\nu} &= d\hat{\sigma}_T\left(-g_{\mu\nu}+\frac{q_\mu q_\nu}{q^2}\right)+\left(p_{b\mu}-\frac{p_b\cdot q}{q^2}q_\mu\right)\left(p_{b\nu}-\frac{p_b\cdot q}{q^2}q_\nu\right)\\ &\cdot\left(\frac{-4q^2}{s_1^2}\right)\left(d\hat{\sigma}_G+\frac{3}{2}d\hat{\sigma}_L\right)-i\epsilon_{\mu\nu\rho\sigma}p_b^\rho q^\sigma\left(\frac{-2q^2}{s_1^2}\right)d\hat{\sigma}_3. \qquad (6.12)\end{aligned}$$

Using the projection tensors $g^{\mu\nu}$, $p_b^\mu p_b^\nu$ and $\varepsilon_{\mu\nu\sigma\rho}p_b^\sigma q^\rho$ one may derive the partonic cross sections $d\hat{\sigma}_T$, $d\hat{\sigma}_2 = (d\hat{\sigma}_G + 3/2 d\hat{\sigma}_L)$ and $d\hat{\sigma}_3$, and from them the structure functions. The partonic cross sections are directly proportional to the projected matrix elements according to

$$d\sigma_i = \frac{1}{2s_1}\frac{1}{2}K_{g\gamma}\mathcal{M}_i^g d\Gamma_2, \qquad i=T,L,3, \qquad (6.13)$$

where the latter are given by

$$\begin{aligned}\mathcal{M}_T^g(\gamma^*g\to Q_M\bar{Q}_m) &= -g^{\mu\nu}\sum\mathcal{M}_\mu(\gamma^*g\to Q_M\bar{Q}_m)\mathcal{M}_\nu^*(\gamma^*g\to Q_M\bar{Q}_m),\\ \mathcal{M}_L^g(\gamma^*g\to Q_M\bar{Q}_m) &= -\frac{4q^2}{(s_1)^2}p_b^\mu p_b^\nu\sum\mathcal{M}_\mu(\gamma^*g\to Q_M\bar{Q}_m)\mathcal{M}_\nu^*(\gamma^*g\to Q_M\bar{Q}_m),\\ \mathcal{M}_3^g(\gamma^*g\to Q_M\bar{Q}_m) &= -i\frac{2q^2}{(s_1)^2}\epsilon_{\rho\sigma}^{\mu\nu}p_b^\rho q^\sigma\sum\mathcal{M}_\mu(\gamma^*g\to Q_M\bar{Q}_m)\mathcal{M}_\nu^*(\gamma^*g\to Q_M\bar{Q}_m).\end{aligned}$$

If the generic coupling between the vector boson B^* and the heavy quark Q is given by

$$g_R\gamma^\mu\frac{1+\gamma_5}{2}+g_L\gamma^\mu\frac{1-\gamma_5}{2}\equiv g_R\gamma^\mu P_R+g_L\gamma^\mu P_L, \qquad (6.14)$$

then the leading–order amplitude $\mathcal{M}^\mu$ might be explicitly written as

$$\begin{aligned}\mathcal{M}^\mu(\gamma^*g\to Q_M\bar{Q}_m) = g_s t^a\bar{u}(k_2)\Bigg[&(g_R\gamma^\mu P_R+g_L\gamma^\mu P_L)\frac{\not{p}_b-\not{k}_3+m}{t-m^2}\gamma^\alpha+\\ &\gamma^\alpha\frac{\not{k}_2-\not{p}_b+M}{u-M^2}(g_R\gamma^\mu P_R+g_L\gamma^\mu P_L)\Bigg]v(k_3)\epsilon_\alpha(p_b) \qquad (6.15)\end{aligned}$$

In appendix D, the expressions for $M_{G,2,L}^g$ are written down explicitly and one can see the explicit symmetry in $t_1\to u_1$ ($M\to m$). The partonic cross–section $\hat{\sigma}_2$ diverges in $t\to 0$ in the limit $m\to 0$ and in $u\to 0$ in the limit $M\to 0$.
In the partonic centre–of–mass frame of the $B^*\gamma\to Q_M\bar{Q}_m$ process, the four-

momenta vectors are

$$\begin{aligned} p_b &= \left(\frac{s+Q^2}{2\sqrt{s}}, 0, 0, -\frac{s+Q^2}{2\sqrt{s}}\right), \\ q &= \left(\frac{s-Q^2}{2\sqrt{s}}, 0, 0, \frac{s+Q^2}{2\sqrt{s}}\right), \\ k_2 &= \left(E_2, 0, |\vec{k}|\sin\theta, |\vec{k}|\cos\theta\right), \\ k_3 &= \left(E_3, 0, -|\vec{k}|\sin\theta, -|\vec{k}|\cos\theta\right), \end{aligned}$$

where E_2, E_3 and $|\vec{k}|$ are determined by imposing the momentum conservation in the evaluation of the two–body phase space

$$\begin{aligned} d\Phi_2 &= \frac{d^3k_2}{(2\pi)^3 2E_2}\frac{d^3k_3}{(2\pi)^3 2E_3}(2\pi)^4\delta^{(4)}(q+p_b-k_2-k_3) \\ &= \frac{1}{(16\pi^2)}\frac{|\vec{k}|^2 d|\vec{k}|d\Omega}{E_2E_3}\delta\left(\sqrt{s}-\sqrt{|\vec{k}|^2+m^2}-\sqrt{|\vec{k}|^2+M^2}\right) \\ &= \frac{1}{8\pi}\frac{|\vec{k}|}{\sqrt{s}}d\cos(\theta) = \frac{1}{16\pi s_1}dt_1 \;=\; \frac{1}{16\pi s_1}du_1. \end{aligned} \tag{6.16}$$

From the energy–momentum conservation, we get

$$\begin{aligned} |\vec{k}| &= \frac{\Delta(s,M^2,m^2)}{2\sqrt{s}}, \\ E_2 &= \frac{(s+M^2-m^2)}{2\sqrt{s}}, \\ E_3 &= \frac{(s-M^2+m^2)}{2\sqrt{s}}, \end{aligned} \tag{6.17}$$

where $\Delta(a,b,c) = \sqrt{a^2+b^2+c^2-2ab-2ac-2bc}$ [109].

The explicit expressions for the partonic cross–sections in Appendix D show that the partonic cross–sections $d\hat{\sigma}_2$ and $d\hat{\sigma}_3$ have a pole in t_1 and a pole in u_1, which integrated over the phase space give rise to logarithms. In particular, the integration in t_1 gives rise to

$$L_t = \int_{|t_1|_{\min}}^{|t_1|_{\max}} \frac{dt_1}{t_1} = \log\left(\frac{s+M^2-m^2+\Delta(s,M^2,m^2)}{s+M^2-m^2-\Delta(s,M^2,m^2)}\right), \tag{6.18}$$

while the integration in u_1 gives rise to

$$L_u = \int_{|u_1|_{\min}}^{|u_1|_{\max}} \frac{du_1}{u_1} = \log\left(\frac{s - M^2 + m^2 + \Delta(s, M^2, m^2)}{s - M^2 + m^2 - \Delta(s, M^2, m^2)}\right). \tag{6.19}$$

In the limit $m^2 << M^2$, as in the case of single top production,

$$L_t \to \log\left[\frac{s}{m^2}\left(1 - \frac{M^2}{s}\right)^2\right] \qquad L_u \to \log\left(\frac{s}{M^2}\right). \tag{6.20}$$

In the special case $m^2 = M^2 \to 0$, as in the case of $b\bar{b}$ or $c\bar{c}$ production,

$$L_t = L_u \to \log\left(\frac{s}{m^2}\right), \tag{6.21}$$

i.e. the argument of the logarithm L_t does not have the phase space suppression on the numerator. We have seen explicitly that, in the collinear region $t \to 0$ and in the massless limit $m \to 0$, the scale associated to the splitting of the gluon into a heavy pair is associated to a scale which does not correspond to Q^2, rather to scales which are directly related to the final phase space configuration. In the next sections we analyse two processes corresponding to the each of the two limits presented above and investigate in details the collinear limits and its implication on the size of the collinear logarithms.

6.2.2 Bottom–Antibottom production

The bottom–antibottom production corresponds to the special case of the cross–section described in the previous subsection, where

$$Q_M = Q_m = b \qquad \text{and} \qquad M = m = m_b. \tag{6.22}$$

Moreover, if we consider only the scattering mediated by a virtual photon, i.e. we exclude the Z contribution, we have $g_R = g_L = e_b$. If θ ($0 \leq \theta \leq \pi$) is the angle between the incoming vector boson and the outgoing b quark in the partonic $\gamma^* g$ centre–of–mass frame, the mass-subtracted Mandelstam variables can be written as

$$t_1 = -\frac{s_1}{2}(1 - \beta\cos\theta) \qquad u_1 = -\frac{s_1}{2}(1 + \beta\cos\theta). \tag{6.23}$$

where β is the velocity of the heavy quark in the centre–of–mass frame, which tends to 1 in the $m_b \to 0$ limit

$$\beta = \frac{|\vec{k}|}{E_{2,3}} = \sqrt{1 - 4m_b^2/s}. \tag{6.24}$$

From the general expressions reported in Appendix D, we may write down the partonic differential cross sections, $d\hat{\sigma}_2$ and $d\hat{\sigma}_L$ as

$$\begin{aligned} \frac{d\hat{\sigma}_2}{d\cos\theta} &= \frac{\pi\alpha_e e_H^2 \alpha_s \beta}{Q^2\left(1-\beta^2\cos^2\theta\right)^2} z \Bigg[-\beta^4\cos^4\theta\left(6z^2-6z+1\right) \\ &\qquad +8\beta^2\cos^2\theta z(\varepsilon(4z-1)+z-1) - 2(4\varepsilon z+z)^2 + 2(4\varepsilon+1)z+1 \Bigg] \end{aligned} \tag{6.25}$$

and

$$\frac{d\hat{\sigma}_L}{d\cos\theta} = \frac{\pi\alpha_e e_H^2 \alpha_s \beta}{4Q^2\left(1-\beta^2\cos^2\theta\right)} z^2\left[(1-\beta^2\cos^2\theta)(z-1)+4\varepsilon z\right], \tag{6.26}$$

where

$$\varepsilon = \frac{m_b^2}{Q^2} \qquad z = \frac{x_B}{\xi}.$$

The differential cross-sections are symmetric in $\cos\theta$, reflecting the symmetry of the matrix elements upon the exchange of $u_1 \leftrightarrow t_1$. For $m_b \to 0$, the cross section diverges in $\cos\theta \to \pm 1$, displaying explicitly the collinear divergence associated to the splitting of a gluon in a pair of massless quarks. In order to study analytically the collinear limit of the partonic cross sections, we expand about $t_1 = 0$. The differential partonic cross section $d\hat{\sigma}_2/dt_1$ is easily deduced from Eq. (6.25) as

$$\begin{aligned} \frac{d\hat{\sigma}_2}{dt_1} &= \frac{2}{s_1\beta}\frac{d\hat{\sigma}_2}{d\cos\theta}\Bigg|_{\cos\theta=(s_1+2t_1)/(\beta s_1)=1/\beta+2t_1 z/(\beta Q^2)} \\ &= \frac{3\pi\alpha_e\alpha_s e_H^2 C_F}{4}\Bigg[\frac{z^2(1-2\varepsilon)\varepsilon}{2t_1^2} + \frac{2z\left(2\left(4\varepsilon^2+6\varepsilon-1\right)z^2+(2-4\varepsilon)z-1\right)}{Q^2 t_1} \\ &\quad -\frac{z^2\left(2\left(6\varepsilon^2+5\varepsilon+5\right)z^2-2(2\varepsilon+5)z+1\right)}{Q^4} + \mathcal{O}(t_1)\Bigg]. \end{aligned}$$

The same expression with $t_1 \to u_1$ would be found if we were expanding about $u_1 = 0$ due to the symmetry of the cross–section under the exchange $u_1 \leftrightarrow t_1$. If we

further expand about $\varepsilon = 0$, the above expression becomes

$$\frac{d\hat{\sigma}_2}{dt_1} = \frac{3\pi\alpha_e\alpha_s e_H^2 C_F}{4}\left[\frac{\mathcal{O}(\varepsilon)}{t_1^2} - \frac{4zP_{qg}(z)+\mathcal{O}(\varepsilon)}{Q^2 t_1} + \frac{z^2(-10z^2+10z-1)+\mathcal{O}(\varepsilon)}{Q^4} + \mathcal{O}(t_1)\right],$$

where P_{qg} is the Altarelli–Parisi splitting function introduced in Chap. 2. If we finally keep only the pole of the divergent part ignoring the $\mathcal{O}(\varepsilon)$ contributions, we end up with an expression for the collinear pole, which, integrated between $t_{\min} = t(\cos\theta = -1) = -s1(1+\beta)/2$ and $t_{\max} = t(\cos\theta = 1) = -s1(1-\beta)/2$ gives

$$\int_{t_{1,\min}}^{t_{1,\max}} dt_1 \frac{d\hat{\sigma}_2^{\rm coll}}{dt_1} = -\frac{3\pi\alpha_e\alpha_s e_b^2 C_F z}{Q^2} P_{qg}(z)\log\left(\frac{1-\beta}{1+\beta}\right) \propto \frac{\alpha_s}{2\pi}\sigma_2^{(0),5\rm F} P_{qg}(z)\log\frac{s}{m_b^2},$$

where $\sigma_2^{(0),5\rm F}$ is the leading–order cross–section in the five–flavor scheme. The integration of the collinear limit yields then a term proportional to

$$\frac{\alpha_s}{2\pi}P_{qg}(z)\log\left(\frac{s}{m_b^2}\right) = \frac{\alpha_s}{2\pi}P_{qg}(z)\log\left[\frac{Q^2}{m_b^2}\frac{(1-z)}{z}\right]. \tag{6.27}$$

The scale (s/m_b^2) corresponds to the scale of the splitting of the gluon into a collinear $b\bar{b}$ pair and it is proportional but not equal to Q^2

$$s = \frac{(1-z)}{z}Q^2 = M_{b\bar{b}}^2.$$

It is interesting to notice that, also in Ref. [233], in the context of the study of the soft gluon radiation in DIS and DY processes, the scale for the soft emission in the soft limit $z \to 1$ was found to be $Q^2(1-z)/z \to Q^2(1-z)$ and its origin was shown to be purely kinematic, i.e. purely due to the structure of the phase space.

The scale (6.27) is a dynamical scale which changes on an event–by–event basis depending on the momentum fraction carried by the gluon and on the kinematic invariants Q^2 and x_B. Therefore, to make any consideration about its size, one has to look at $d\hat{\sigma}/ds$. In other words, for a given collider energy and acceptance, one has to check what is the distribution of values of $(1-z)/z$ with respect to one. In the case it is found to be smaller than one, then the logarithms of (s/m_b^2) are not in fact large, even when $Q^2 \gg m_b^2$. If, on the other hand, it is close or even larger than one, logarithms should be resummed.

In order to investigate such behaviour, we run a modified version of the HVQDIS code [207] with two different settings: one for the HERA experiment and one that represents a typical relevant kinematical region for the LHC. For HERA we set the energies of the beams to the Run II energies $E_p = 920\,\text{GeV}$ and $E_e = 27.5\,\text{GeV}$ and we set the kinematical cuts in (y, Q^2) to $2 \le Q^2 \le 100\ \text{GeV}^2$ and $0.1 \le y \le 0.9$. At the LHC the incoming electron is replaced by an incoming quark. In order to identify

where the distribution of the fraction of momentum carried by the incoming quark is peaked, we study the process $ug \to ub\bar{b}$ and plot the distribution of the fraction of energy carried by an initial u quark. In Fig. 6.7 we see that, if we set the energy of the up quark to a fixed value given by the mean of the u momentum fraction distribution ($z = 0.15$, E_u=1.05 GeV), then the gluon momentum fraction distribution reproduces well the solid line histogram. Therefore, to reproduce a LHC-like kinematic, we set

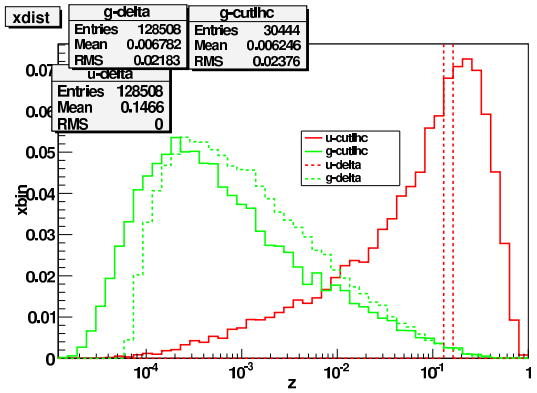


Figure 6.7: Solid lines: distribution of the fraction of momentum carried by the initial up quark and the gluon in the process $ug \to ub\bar{b}$, $s = (14\text{ TeV})^2$. Dashed lines: distribution of the fraction of momentum carried by the gluon in the process $ug \to ub\bar{b}$, if $E_g = 7$ TeV and $z_{\rm up} = 0.15$, E_u=1.05 TeV .

the energy of the incoming proton to a value of 7 TeV and the one of the incoming electron to $\sim 0.15 \cdot 7 = 1.05$ TeV so that $S_{\rm had} = 29.4\,\text{TeV}^2$. The value for the cuts in y and Q^2 is such that the transverse momentum of the outgoing electron is above 20 GeV, a typical experimental cut for the detection of light jets. Given that $p^e_{T,\rm min} \sim 20\text{GeV} = \sqrt{Q^2_{\rm min}(1-y_{\rm max})}$, we set

$$Q^2_{\rm min} = 2000\ \text{GeV}^2 \qquad y_{\rm max} = 0.8.$$

The leading–order results produced with the modified version of HVQDIS have been cross–checked against MadGraph v4 [234].

As from Tab. 6.2, the comparison shows that results are perfectly compatible within the uncertainty.

In Fig. 6.8 we plot the distribution of Q^2/m_b^2, s/m_b^2, and the factor $\frac{(1-z)}{z}$ in logarithmic scale for both kinematics. The distributions are very different in the two cases. In

cteq66	HVQDIS	MG
s = 101200 GeV^2 $2 \leq Q^2 \leq 100$ GeV^2 $0.1 \leq y \leq 0.9$	$\sigma_{\rm tot} = 398.749 \pm 0.097$ (pb)	$\sigma_{\rm tot} = 399.170 \pm 1.104$ (pb)
s = 29.4 TeV^2 $2{\cdot}10^3 \leq Q^2 \leq 10^6$ GeV^2 $0.001 \leq y \leq 0.8$	$\sigma_{\rm tot} = 46.704 \pm 0.072$ (pb)	$\sigma_{\rm tot} = 46.276 \pm 0.648$ (pb)

Table 6.2: Comparison between MG and HVQDIS total cross sections for HERA and LHC kinematics. Input PDF: CTEQ66 [42].

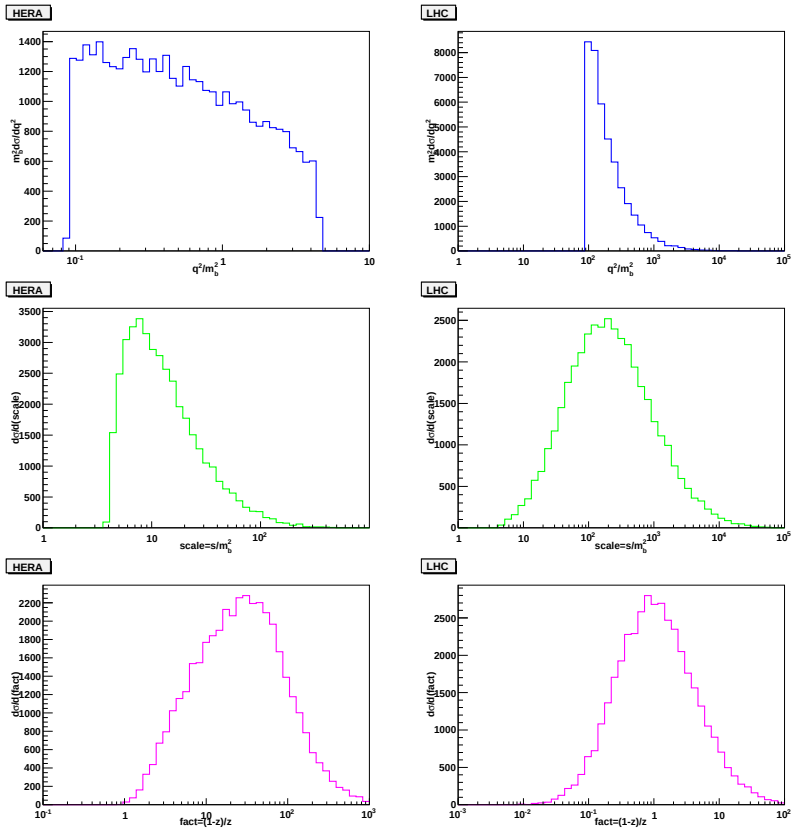


Figure 6.8: Q^2/m_b^2 (top), s/m_b^2 (middle) and $(1-z)/z$ (bottom) logarithmic distributions for $b\bar{b}$ production in the HERA (left – $2 < Q^2 < 100$, $0.1 < y < 0.9$, E_p= 920 GeV, E_e = 27.5 GeV) and LHC–like (right – $2 \cdot 10^3 < Q^2 < 10^6$, $0.001 < y < 0.8$, E_p= 7 TeV, E_e = 1.05 TeV) kinematics. Input PDF: CTEQ66 [42].

the HERA kinematics, the invariant mass of the $b\bar{b}$ pair has a peak at a higher scale with respect to Q^2. Indeed, the factor $(1-z)/z$ is peaked between 10 and 100. This means that the scale of the logarithm is actually larger than Q^2/m_b^2. Therefore, even if the experimental cuts are such that Q^2 lies in a region where $\log(Q^2/m_b^2)$ is very small, the effect of these logarithms might be enhanced by a scale which is effectively ten up to hundred times larger than Q^2. On the other hand, in the LHC–like kinematics the invariant mass distribution has a peak about a scale which is hundred times larger than m_b^2, consistently with the prefactor $(1-z)/z$ is peaked about one and its distribution is smoothly distributed between 0.1 and 100, with a slight dominance in the region on the right of the peak. This means that in a LHC–like kinematics, the scale associated to the collinear splitting is about Q^2 and that therefore the logarithms which are resummed in the bottom PDF are important as one would expect by looking at the Q^2 distribution.

Another interesting question is whether the shape of the pre–factor $(1-z)/z$ is due to the shape of the gluon PDF or to the energies and cuts associated to the considered kinematic configuration. To clarify this issue, we reproduced the same distribution using several different input PDFs for the LHC–like kinematics by ending up with the very same distributions as the ones shown in Fig. 6.8. We also run the code a toy gluon $x^{-a}(1-x)^b$ and looked at the distributions as varying the parameters a and b. We observed that, only if the gluon PDF were much steeper than the CTEQ66 gluon, i.e. only for unrealistic exponents $a \geq 5$, the peak of the $(1-z)/z$ distribution would be more pronounced and would shift to the left. For all other values of a, the distribution does not change. On the other hand, the high–x exponent b seems to be irrelevant. Therefore we concluded that for realistic PDFs the scale of the logarithms depends mostly on the kinematics rather than on the gluon shape.

We look then at the differential distributions of the cross–section in terms of the transverse momentum and pseudo-rapidity of the outgoing b quark. Their shapes change radically from the HERA to the LHC kinematics. In the HERA kinematics the transverse momentum is peaked in the small p_T region and the pseudo–rapidity is peaked about zero. In the LHC kinematics instead, the transverse momentum distribution exhibits a double peak, one at small p_T and one at $p_T \sim 50$ GeV, while the pseudo–rapidity distribution is centred in the backward region. In Fig. 6.9 we observe that the distribution of momenta of the less energetic quark is peaked in the small p_T, while the distribution of the most energetic quark is peaked about $p_T \sim 50$ GeV. One quark is soft and the other compensates the p_T of the outgoing electron. This picture is confirmed by looking at the pseudo–rapidity distribution. The difference of pseudo–rapidity for the three outgoing particles, electron, bottom and anti–bottom shows that the less energetic quark lies is in the forward region, compensating the electron pseudo–rapidity, while the hard quarks is in the central region. This observation tends to confirm what we found previously, namely that in the LHC kinematics the

logarithms are dominant and that therefore the five–flavor scheme where the leading order process is a $2 \to 2$ needs to be matched with the four–flavor one.

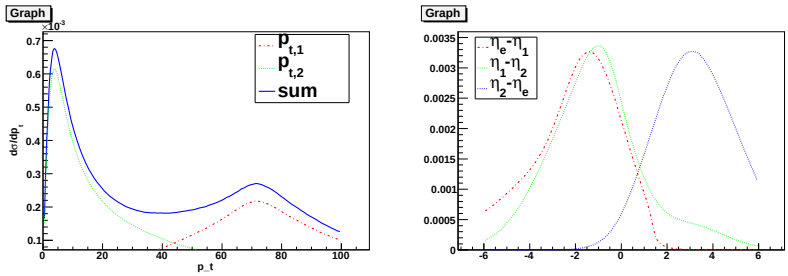


Figure 6.9: LO differential cross sections in p_T (left) and pseudo–rapidity η (right) for $b\bar{b}$ production in LHC kinematics, $Q^2_{min} = 2000$ GeV. The two dashed histograms represent the distributions relative to the most energetic quark (b_1) and for the less energetic one (b_2). PDF input set: CTEQ66 [42].

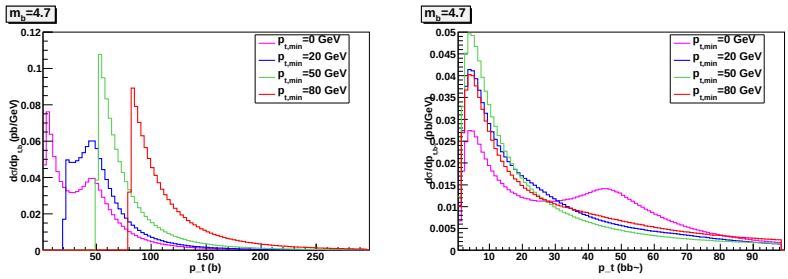


Figure 6.10: Comparison of normalised LO transverse momentum differential cross sections for bottom productions in the LHC kinematic. PDF input set: CTEQ66 [42]. Cut on the transverse momentum of the outgoing b quark.

We finally look at the transverse momentum distributions for b and $\bar{b}$ when a cut on their p_T is imposed. In absence of cuts on p_T, the b and $\bar{b}$ transverse momentum distributions look exactly the same. If instead a cut is applied on the transverse momentum of the outgoing bottom quark, leaving the other quark free, the distributions look obviously different. In Fig. 6.10 we notice that the shape of the transverse momentum of the $\bar{b}$ antiquark does not change much, even raising the cut on p_T^b from 20 to 80 GeV. This suggests that the scale associated to the collinear logarithms is softer when a cut on the p_T of the b is applied than when no cuts are imposed. Since the cut in p_T sup-

presses the u–channel contribution to the cross–section with respect to the t–channel one, this case relates to the one which is going to be discussed in the next section.

To conclude, we found that the scale of the logarithms is a dynamical scale which depends on the final state particles momenta. This scale is peaked at a scale higher than Q^2 in the HERA kinematics, while it is peaked about Q^2 for the LHC kinematics. This means that in both cases the size of the logarithms is as large or even larger than $\log(Q^2/m_b^2)$. This confirms the result of Ref. [209] about the size of the logarithms resummed in the massless approach and clarifies its dynamical origin.

6.2.3 Single top

This process corresponds to one special of the two special cases of the general process discussed in Sect. 7.2.1, where

$$M = M_t \qquad m = m_b \qquad g_R = 0 \qquad g_L = g_W/\sqrt{2}.$$

If we restrict the general expressions reported in Appendix D to this case, we may write the differential cross–section in t as

$$\begin{aligned}
\frac{d\sigma_2}{dt} \;=\; & \frac{3\alpha_s g_W^2 C_F}{128 s_1^4 (t-m_b^2)^2 (u-M_t^2)^2}\Bigg\{ 2m_b^2 M_t^2 s_1^4 + 2s_1^3 (M_t^4 t_1 + m_b^4 u_1) \\
& +12Q^2 t_1 u_1 \left[s_1 (t_1 M_t^2 + u_1 m_b^2) + u_1 t_1 (s_1 - Q^2) \right] \\
& -2s_1^2 \Bigg[m_b^2 u_1 (Q^2 s_1 - Q^2 t_1 + s_1 t_1) \\
& +M_t^2 t_1 (Q^2 s_1 - Q^2 u_1 + s_1 u_1) \Bigg] \\
& +s_1^2 t_1 u_1 \left(2Q^4 - 2Q^2 s_1 + t_1^2 + u_1^2 \right) \Bigg\}
\end{aligned} \tag{6.28}$$

It is apparent that the cross–section has a pole in $t = 0$ in the small m_b limit. If we expand the above expression about t_1=0, we get

$$\begin{aligned}\frac{d\sigma_2}{dt} &= \frac{3\alpha_s g_W^2 C_F}{64(s+Q^2)^2}\Bigg\{\frac{\mathcal{O}(m_b^4, m_b^2 M_t^2, m_b^2 Q^2)}{t_1^2} - \\ &\frac{+2M_t^4 + 4M_t^2 Q^2 - 2M_t^2 s_1 + 2Q^4 - 2Q^2 s_1 + s_1^2 + \mathcal{O}(m_b^4, m_b^2 M_t^2, m_b^2 Q^2)}{2(s+Q^2)t_1} - \\ &\frac{-4M_t^4 + 6M_t^2 Q^2 + 2M_t^2 s_1 + 10Q^4 - 10Q^2 s_1 + s_1^2 + \mathcal{O}(m_b^4, m_b^2 M_t^2, m_b^2 Q^2)}{2(s+Q^2)^2} \\ &+\mathcal{O}(t_1)\Bigg\} \end{aligned} \tag{6.29}$$

In the collinear limit $t \to 0$, if we ignore the terms proportional to m_b^2, we end up with an expression for the collinear pole in t_1, which, integrated between $t_{\min}$ and $t_{\max}$ yields

$$\begin{aligned}\int_{t_{\min}}^{t_{\max}} dt \frac{d\hat{\sigma}_2^{\rm coll}}{dt} &= \frac{3\alpha_s g_W^2 C_F}{128 s_1^3}\left(2M_t^4 + 4Q^2 M_t^2 - 2(s+Q^2)M_t^2 + 2Q^4 + s^2 + Q^4\right) \\ &\qquad \cdot \log\left[\frac{s}{m_b^2}\left(1 - \frac{M_t^2}{s}\right)^2\right] \\ &\propto \frac{\alpha_s}{2\pi}\sigma_2^{(0),5F} P\left(\frac{M_t^2}{s}\right)\log\left[\frac{M_t^2}{m_b^2}\frac{s}{M_t^2}\left(1 - \frac{M_t^2}{s}\right)^2\right], \quad m_b^2 \lesssim Q^2 \ll M_t^2,\end{aligned}$$

where $\sigma_2^{(0),5F}$ is the Born–level cross section for the leading order process in the five–flavor scheme. Notice that in the last equality we used $Q^2 \ll M_t^2$, which is justified by looking at the distribution of the cross section in Q^2/M_t^2 in Fig. 6.11. The scale corresponding to the collinear splitting then is not M_t^2, rather a dynamical scale proportional to the ratio M_t^2/m_b^2 by a factor

$$\frac{(1-z)^2}{z} \qquad \text{with} \quad z = \frac{M_t^2}{s}.$$

With respect to the previous case, there is an additional factor $(1 - z)$ due to the suppression of the final phase space related to the production of a heavy final state. In Fig. 6.11 we plot the differential cross section as a function of Q^2/M_t^2, of the scale of the logarithm $s(1 - M_t^2/s)^2/m_b^2$, of the factor $(1 - z)^2/z$ and its squared root, being $z = M_t^2/s$. The kinematics is the one of the LHC. We see that the distribution of the $(1-z)^2/z$ pre–factor is centred on a value close to one and it is broadly spread about it, being slightly enhanced on the left of the unity. This yields a distribution of the scale of the collinear logarithms which in average is slightly smaller than the scale that one

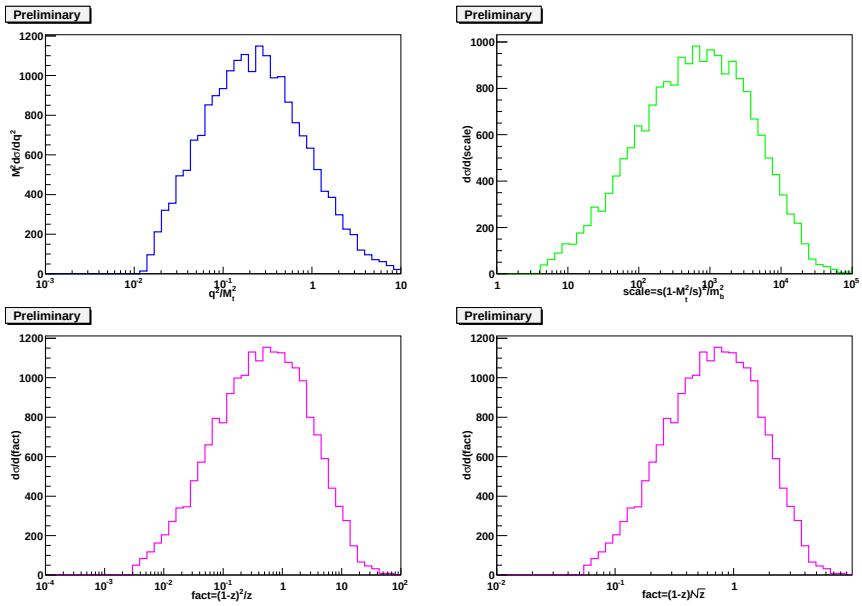


Figure 6.11: Distribution of q^2/M_t^2 (top left), the scale of the collinear logarithm (top right), the $(1-z)^2/z$ factor (bottom left) and its squared root (bottom right). Here the energies of the proton beams are at 7 TeV. $M_t = 174.3$ GeV, $m_b = 4.7$ GeV. Input PDF: CTEQ6L1 [59].

would expect by simply identifying the scale of the splitting with M_t^2. However this is less small than we would expect from the studies performed in Refs. [230, 231].

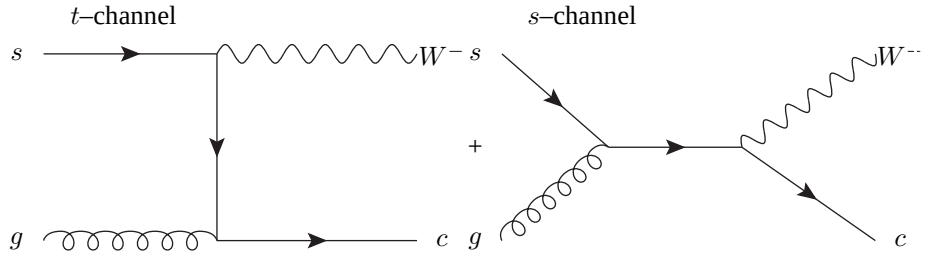
6.3 Associated W and charm production

We turn now to discuss another class of process where the large scale which makes the process perturbative is the invariant mass of the final states, rather than the high virtuality of the incoming particles, as it is in the DIS case. In particular we consider the Drell–Yan production of a W boson associated to the production of a charm quark. The advantage of this choice is that it is easier to gain analytic control and that, as discussed in Sect. 7.1, it might bear some phenomenological relevance.

6.3.1 Massive calculation

Here below we draw the LO diagram that contribute to the LO process for the Wc production:

$$g(p_1) + \bar{s}(p_2) \longrightarrow c(p_3) + W^-(p_4) \tag{6.30}$$



The Mandelstam invariants for this process are defined as

$$\begin{aligned}
\xi_1\xi_2 S_{\rm had} = s &= (p_1+p_2)^2 = 2(p_1\cdot p_2) = (p_3+p_4)^2 = m_c^2 + M_W^2 + 2(p_3\cdot p_4),\\
t &= (p_1-p_3)^2 = m_c^2 - 2(p_1\cdot p_3) = (p_2-p_4)^2 = M_W^2 - 2(p_2\cdot p_4),\\
u &= (p_1-p_4)^2 = M_W^2 - 2(p_1\cdot p_4) = (p_2-p_3)^2 = +m_c^2 - 2(p_2\cdot p_3).
\end{aligned}$$

As in the previous section, it is useful to define the Mandelstam mass subtracted variables as

$$t_1 = t - m_c^2 \qquad u_1 = u - M_W^2 \qquad s_1 = s - (m_c + M_W)^2. \tag{6.31}$$

The matrix element is the sum of the contributions coming from the t and the s channels

$$\begin{aligned}
\mathcal{M}(gs\to Wc) &= \mathcal{M}_t(gs\to Wc) + \mathcal{M}_s(gs\to Wc), \qquad (6.32)\\
\mathcal{M}_t(gs\to Wc) &= i\frac{g_s^2 g_W^2 t_A}{2\sqrt{2}}\varepsilon_\nu(p_2)\varepsilon^*_\mu(p_4)\bar{u}(p_3)\gamma^\nu\frac{(\not{p}_3-\not{p}_1)+m_c}{t-m_c^2}\gamma^\mu(1-\gamma_5)u(p_2),\\
\mathcal{M}_s(gs\to Wc) &= i\frac{g_s^2 g_W^2 t_A}{2\sqrt{2}}\varepsilon_\nu(p_2)\varepsilon^*_\mu(p_4)\bar{u}(p_3)\gamma^\mu(1-\gamma_5)\frac{(\not{p}_1+\not{p}_2)}{s}\gamma^\nu u(p_2).
\end{aligned}$$

It is the same as the one studied in the previous case with $u_1 \to s$, $M = m_c$, $m = 0$ and $Q^2 = -M_W^2$. The squared matrix element averaged over the initial colours and

polarisations and summed over the final ones reads

$$\begin{aligned}\bar{\Sigma}|\mathcal{M}|^2 &= \frac{g_s^2 g_W^2 |V_{cs}|^2}{24 M_W^2 s\,(t-m_c^2)^2}\Bigg[-m_c^8 + m_c^6(2s+t) \\ &\quad + m_c^4\left(2M_W^4 + 2M_W^2 t - (s+t)^2\right) \\ &\quad + m_c^2\Big(-4M_W^6 + 2M_W^4 t - 2M_W^2\left(s^2 - st + 2t^2\right) + t(s+t)^2\Big) \\ &\quad + 2M_W^2 t\left(2M_W^4 - 2M_W^2(s+t) + s^2 + t^2\right)\Bigg]. \end{aligned} \tag{6.33}$$

In the partonic centre of mass frame we can write the 4-moment of the particles as

$$\begin{aligned} p_1 &= (\sqrt{s}/2, 0, 0, \sqrt{s}/2), \\ p_2 &= (\sqrt{s}/2, 0, 0, -\sqrt{s}/2), \\ p_3 &= (E_3, 0, |\vec{p}|\sin\theta, |\vec{p}|\cos\theta), \\ p_4 &= (E_4, 0, -|\vec{p}|\sin\theta, -|\vec{p}|\cos\theta). \end{aligned}$$

where $|\vec{p}| = |\vec{p}_4| = |\vec{p}_3|$ and θ is the angle between the outgoing c quark and the colliding partons. The phase space in 4 dimensions is given by

$$\begin{aligned} d\Phi_2 &= \frac{d^3p_3}{(2\pi)^3 2E_3}\frac{d^3p_4}{(2\pi)^3 2E_4}(2\pi)^4\delta^{(4)}(p_1+p_2-p_3-p_4) \\ &= \frac{1}{8\pi}\frac{|\vec{p}|}{\sqrt{s}}d\cos(\theta) = \frac{1}{8\pi s}dt_1 = \frac{1}{8\pi s}\frac{p_T}{\sqrt{|\vec{p}|^2 - p_T^2}}dp_T, \end{aligned} \tag{6.34}$$

with

$$\begin{aligned} |\vec{p}| &= \frac{\Delta(s, M_W^2, m_c^2)}{2\sqrt{s}}, \\ E_3 &= \frac{s + m_c^2 - M_W^2}{2\sqrt{s}}, \\ E_4 &= \frac{s + M_W^2 - m_c^2}{2\sqrt{s}}. \end{aligned} \tag{6.35}$$

The total cross section is then obtained by integrating over the phase space and the parton distribution functions:

$$\sigma_{\rm tot}^{\rm LO,4F} = \int d\xi_1\, s(\xi_1,\mu_F^2) \int d\xi_2\, g(\xi_2,\mu_F^2) \int dX \frac{d\sigma}{dX}, \tag{6.36}$$

where

$$\frac{d\sigma}{dX} = \frac{1}{2s}\bar{\Sigma}|\mathcal{M}_{gs}^{(0)}|^2 \frac{d\Phi_2}{dX} \qquad \text{with} \quad X = \{t, t_1, p_T, \cos\theta, ...\}\,. \tag{6.37}$$

In Fig. 6.12 we plot the partonic cross section, without including PDFs, as a function of (s_1, t_1) and (s_1, p_T). We see that most of the events lie in the very small t_1 region, and in the relatively small s_1 region. The same for the p_T distribution, although there is also a peak around $p_T \sim p_{T,\rm max}$ due to the Jacobian which appears in the phase space, Eq. (6.34). In order to check how close to the threshold is the scale at which the partonic cross–section is damped, in the same figure we plot two transverse sections of the three–dimensional plots, one for t_1 close to 0 and one for $s_1 = 10\,{\rm GeV}^2$. We see that at very small $|t_1|$ the cross–section starts being damped at $s-(M_W+m_c)^2 \sim 2\,10^4$ GeV^2 which is about more than two units of $(m_c+M_W)^2$, therefore pretty far from the threshold region. On the other hand, looking at the plot for s_1 fixed, most of the events are in the region very close to the collinear limit $t_1 \sim 0$.

To draw the same plot by including the parton distributions, we perform the change of variables $(\xi_1,\xi_2) \longrightarrow (\tau, y)$, where

$$\xi_1 = \sqrt{\tau} e^{+y} \qquad \xi_2 = \sqrt{\tau} e^{-y}.$$

Under this transformation

$$\int_{\tau_0}^1 d\xi_1 \int_{\tau_0/\xi_1}^1 d\xi_2\, s(\xi_1,\mu_F^2) g(\xi_2,\mu_F^2) = \int_{\tau_0}^1 d\tau \int_{-\log\tau/2}^{+\log\tau/2} dy\, s(\sqrt{\tau}e^{+y},\mu_F^2) g(\sqrt{\tau}e^{-y},\mu_F^2),$$

where

$$\tau_0 = \frac{(m_c+M_W)^2}{S_{\rm had}}. \tag{6.38}$$

By integrating in the rapidity y the product of the PDFs, we obtain the parton luminosity as a function of τ

$$\mathcal{L}(\tau,\mu_F^2) = \int_{-\log\tau/2}^{+\log\tau/2} dy\, s(\sqrt{\tau}e^{+y},\mu_F^2) g(\sqrt{\tau}e^{-y},\mu_F^2) \tag{6.39}$$

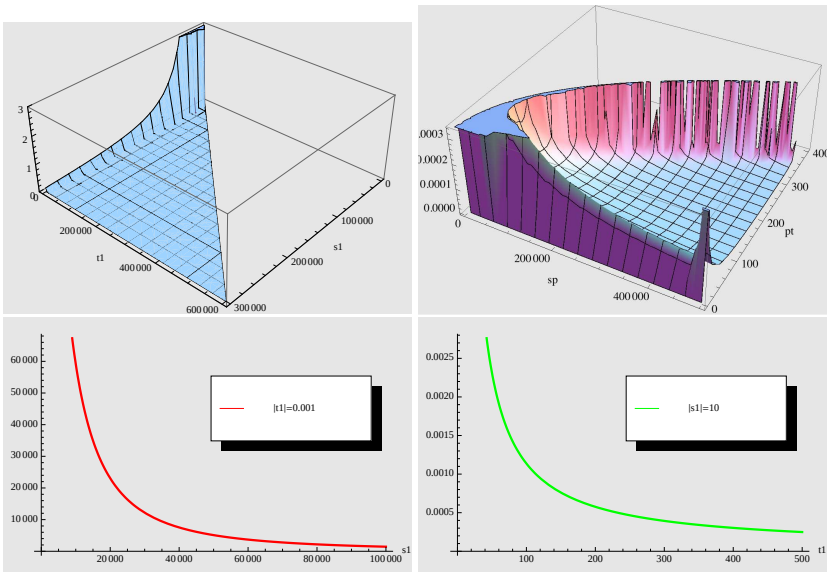


Figure 6.12: Partonic cross section plotted as a function of $(|t_1|, s_1)$ (top left) and (p_T, s_1) (top right). Transverse sections of the top left 3D plot for a fixed value of t_1 (bottom left) and s_1 (bottom right).

and we can write the total cross section as

$$\sigma_{\rm tot}^{\rm LO,4F} = \int_{\tau_0}^{1} d\tau \mathcal{L}(\tau, \mu_F^2) \int dX \frac{d\sigma}{dX}(\tau/S_{\rm had}, m_c^2, X, \alpha_s(\mu_R^2)). \tag{6.40}$$

In Fig. 6.13 we display the double differential partonic cross–section multiplied by the parton luminosity obtained with the MSTW08 [43] NLO parton set. The effect of the luminosity is to damp further the cross section in s_1. In order to assess how close to the threshold is the scale at which the PDFs further damp the cross–section, we plot on the same figure the MSTW08 parton luminosity as a function of $s_1 = s - (M_W + m_c)^2$ at LHC, where $S_{\rm had} = (10\ {\rm TeV})^2$. We see that the luminosity starts dropping at $s_1 \sim (100\ {\rm GeV})^2$, i.e. more than one unity in $(M_W + m_c)^2$, therefore the region enhanced by the luminosity is broader than the threshold region but narrower than the region which would be enhanced in absence of PDFs, see Fig. 6.12.

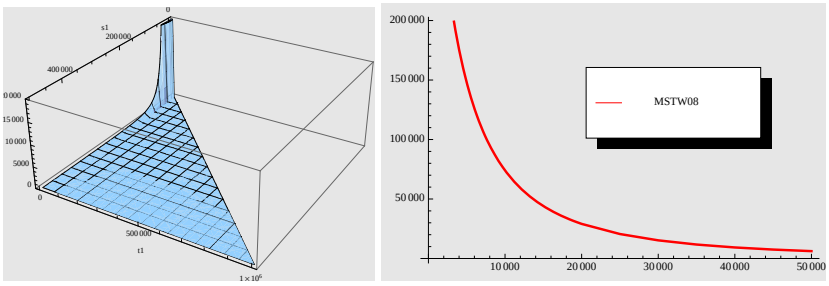


Figure 6.13: Partonic cross section multiplied by the parton luminosity plotted as a function of t_1 and s_1 (left) and parton luminosity as a function of s_1 (right) at the LHC $S = (10\text{TeV})^2$. Input PDFs: MSTW08 [43].

We now compute analytically the differential cross–section $d\hat{\sigma}/ds$, related to the total cross–section by

$$\sigma_{\text{tot}}^{\text{LO,4F}} = \int_{(m_c+M_W)^2}^{S_{\text{had}}} ds \mathcal{L}(s,\mu_F^2) \frac{d\hat{\sigma}}{ds}. \tag{6.41}$$

To evaluated $d\hat{\sigma}/ds$ we integrate the double differential cross section $d\hat{\sigma}/dsdt_1$ in t_1 between

$$\begin{aligned} t_{1,\max} &= -\frac{1}{2}\left(s - m_c^2 - M_W^2 + \Delta(s,m_c^2,M_W^2)\right) \\ t_{1,\min} &= -\frac{1}{2}\left(s - m_c^2 - M_W^2 - \Delta(s,m_c^2,M_W^2)\right) \end{aligned} \tag{6.42}$$

and obtain

$$\begin{aligned} \frac{d\hat{\sigma}}{ds} &= \int_{t_{1,\min}}^{t_{1,\max}} dt_1 \, \bar{\Sigma}|\mathcal{M}_{gs}(t_1,s,\alpha_s(\mu_R^2))|^2 \frac{1}{16\pi s^2} \qquad (6.43) \\ &= -\frac{\alpha_s G_F V_{cs}^2}{24\sqrt{2}s^3}\Bigg[2\left(m_c^2+2M_W^2\right)\left(2m_c^4+m_c^2\left(2s-4M_W^2\right)+2M_W^4-2M_W^2 s+s^2\right) \\ &\quad \log\left(\frac{s+m_c^2-M_W^2-\Delta(s,m_c^2,M_W^2)}{s+m_c^2-M_W^2+\Delta(s,m_c^2,M_W^2)}\right) \\ &\quad +\sqrt{m_c^4-2m_c^2\left(M_W^2+s\right)+\left(M_W^2-s\right)^2} \\ &\quad \cdot\left(7m_c^4+m_c^2\left(7M_W^2+3s\right)-2\left(7M_W^4+M_W^2 s\right)\right)\Bigg]. \end{aligned}$$

As we did in the precedent analyses, we study the collinear limit of the above expression. Expanding the integrand of Eq. (6.43) about $t_1 = 0$, we get

$$\begin{aligned}\frac{d\hat{\sigma}}{ds dt_1}(s, t_1, \frac{m_c^2}{M_W^2}, M_W^2) \quad &\propto \quad \frac{\mathcal{O}(\epsilon)}{t_1^2} + \frac{-2M_W^6 + 2sM_W^4 - s^2M_W^2 + \mathcal{O}(\epsilon)}{6\sqrt{2}s^3}\frac{1}{t_1} \\ &+ \left(\frac{M_W^4}{3\sqrt{2}s^3} + \mathcal{O}(\epsilon^1)\right) + \mathcal{O}(t_1), \qquad (6.44)\end{aligned}$$

where $\epsilon = m_c^2/M_W^2$. If we neglect term of $\mathcal{O}(\epsilon)$ and the non collinear term $\mathcal{O}(t_1^0)$, we are left with the pole in t_1

$$\frac{d\hat{\sigma}}{ds dt_1}(s, t_1, \frac{m_c^2}{M_W^2}, M_W^2) \sim \frac{-2M_W^6 + 2sM_W^4 - s^2M_W^2}{s^3}\frac{1}{t}. \qquad (6.45)$$

In approximation $m_c^2 \ll M_W^2$,

$$\begin{aligned}|t|_{\max} \quad &\sim \quad s - M_W^2 + \mathcal{O}(m_c^2), \\ |t|_{\min} \quad &\sim \quad \frac{m_c^2 s}{s - M_W^2} + \mathcal{O}(m_c^4), \qquad (6.46)\end{aligned}$$

and therefore the integration yields

$$\begin{aligned}\frac{d\hat{\sigma}}{ds} \quad &\sim \quad \frac{M_W^2}{s^3}(s^2 - 2M_W^2 s + 2M_W^4)\log\left[\frac{s}{m_c^2}\left(1 - \frac{M_W^2}{s}\right)^2\right], \\ &\propto \quad \frac{\alpha_s}{2\pi}\,\sigma^{(0),4F}\,P_{qg}\left(\frac{M^2}{s}\right)\log\left[\frac{s}{m_c^2}\left(1 - \frac{M_W^2}{s}\right)^2\right], \qquad (6.47)\end{aligned}$$

where $\sigma^{(0),4F}$ is the Born–level cross section for the leading–order process in the four–flavor scheme. The scale associated to the collinear splitting

$$\frac{M_W^2}{m_c^2}\frac{s}{M_W^2}\left(1 - \frac{M_W^2}{s}\right)^2 = \frac{M_W^2}{m_c^2}\frac{(1-z)^2}{z}, \qquad z = M_W^2/s \qquad (6.48)$$

is a dynamical scale which depends on the partonic centre–of–mass energy $s = \tau S$. The result is similar to the one found in the single top analysis.

In order to assess the size of this scale, in Fig. 6.14 we plot the distribution of the cross section in terms of s_1/M_W^2, $(1-z)^2/z$ and $(1-z)/\sqrt{z}$ with $z = M_W^2$/s and the scale of the logarithm $\frac{\sqrt{s}}{m_c}\left(1 - \frac{M_W^2}{s}\right)$. The scale is peaked about a value which is significantly smaller than M_W^2, as we expect by looking at the distribution of the

$(1-z)^2/z$ factor in the same Fig. Indeed, if the scale of the logarithm were M_W/m_c, the scale associated to the logarithm would be distributed about 60. Here the shift observed in its distribution with respect to M_W^2 is far more enhanced than in the single top case.

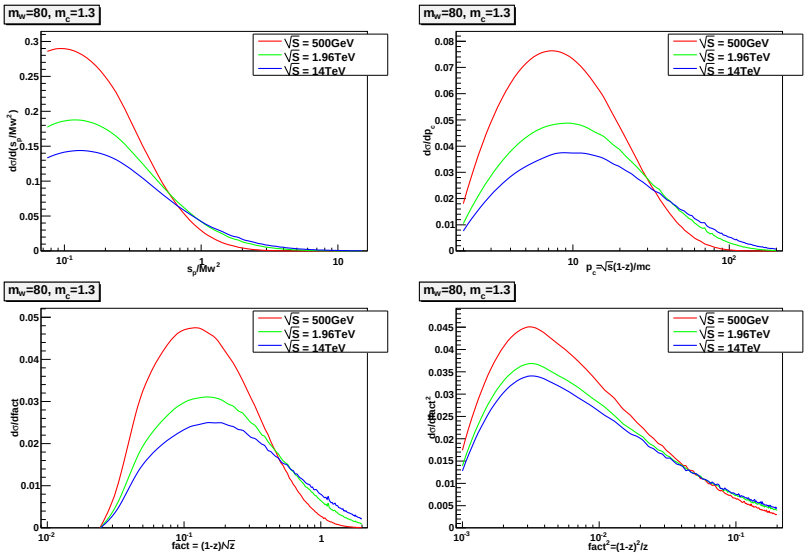


Figure 6.14: Distribution of the LO Wc cross–section as a function of s_1/M_W^2 (top left), the scale of the logarithm $\sqrt{s}\left(1-M_W^2/s\right)/m_c$ (top right), the factor multiplying M_W^2/m_c^2: $\frac{(1-z)^2}{z}$ with $z = M_W^2/s$ (bottom left) and its squared root (bottom right).

We may wonder what happens if the only the gluon PDF is turned on, while the strange PDFs is peaked as a delta function about the average fraction of momentum carried. To answer this question, we first plot the distributions of x_s and x_g for $m_c = 1.3$ GeV and $M_W = 80$ GeV for the three considered hadronic centre–of–mass energies. Then we set $s(x) = \delta(1 - x_{\max})$ and re-evaluate the distributions. The latter are displayed in Fig. 6.16. Their shape does not change significantly with respect to Fig. 6.14. We have also tried to substitute the initial strange by a initial down quark. Even though the distribution of ξ_2 is different when the initial quark is a down quark, as it is observe on the right–hand side plot in Fig. 6.15, the shape of the distributions does not change at all. This means that the distribution of the scale associated to the gluon splitting in a

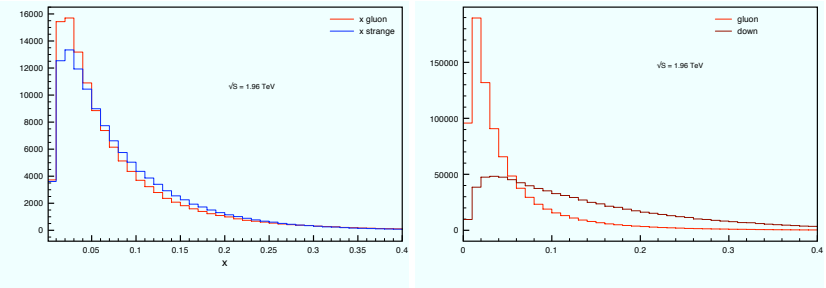


Figure 6.15: Plot of the momentum fraction carries by the gluon, strange and down PDFs for $S_{\rm had} = (14\,{\rm TeV})^2$. Here M_W=80 GeV^2, m_c= 1.3 GeV^2. The gluon and the strange distributions are plotted on the left and the gluon and the down distributions are plotted on the right.

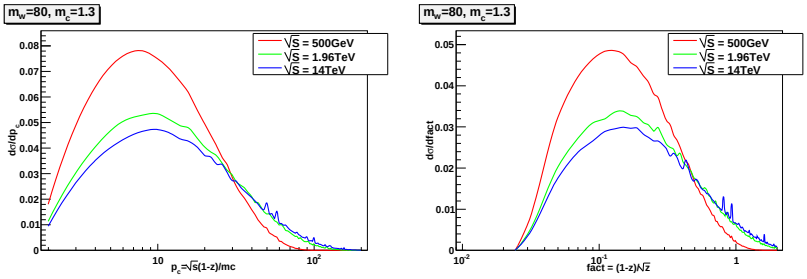


Figure 6.16: Distribution of the LO Wc cross–section as a function of the scale of the logarithm $s\left(1-M_W^2/s\right)^2/m_c^2$ (left) and the factor multiplying M_W^2/m_c^2: $\frac{(1-z)^2}{z}$ (right). Here $s(x) = \delta(1-x_{\rm avg})$

almost collinear c–$\bar{c}$ pair is rather independent of the PDFs. It seems to have a purely kinematic origin.

To conclude our discussion about this process, we integrate the double differential cross–section in s and look at the distribution of the differential cross section in the t_1

$$\frac{d\sigma}{dt_1} = \int_{\tau_{\rm min}}^{1} d\tau \mathcal{L}(\tau,\mu_f^2)\bar{\Sigma}|\mathcal{M}_{gs}(t,\tau S,\alpha_s(\mu_r^2))|^2 \frac{1}{16\pi s^2}, \tag{6.49}$$

where $\tau_{\rm min} > \tau_0$ is determined by inverting Eq. (6.42), as illustrated in Fig. 6.17, and obtaining

$$\tau_{\rm min} = \frac{m_c^2+M_W^2}{S} - \frac{m_c^2 M_W^2}{St} - \frac{t}{S}.$$

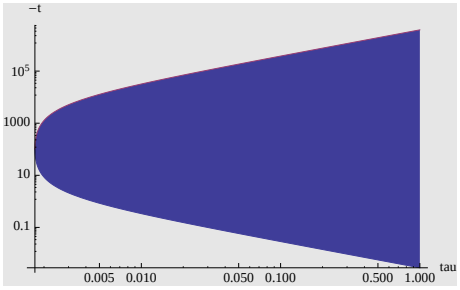


Figure 6.17: Region of integration of Eq. (6.43) in the plane $(\tau, |t|)$.

To perform analytically the integration in Eq. (6.49), we introduce a toy parton luminosity

$$\mathcal{L}^{\rm toy}(\tau) = \tau^{-l}.$$

and end up with the following result

$$\begin{aligned}
\frac{d\sigma}{dt_1} \;=\; & \frac{\alpha_s G_F V_{cs}^2}{12\sqrt{2}S^3\left(m_c^2-t\right)^2}\Bigg[-\frac{S^2(m_c^2+2M_W^2)(m_c^2-t)}{l} \\
& -\frac{\left(m_c^2+2M_W^2\right)\left(m_c^2-t\right)\left(m_c^4-2m_c^2M_W^2+2M_W^4-2M_W^2t+t^2\right)}{l+2} \\
& +S^2\Bigg(\frac{\left(m_c^2+2M_W^2\right)\left(m_c^2-t\right)}{l} \\
& +\frac{t^2\left(m_c^2+2M_W^2\right)\left(m_c^4-2m_c^2M_W^2+2M_W^4-2M_W^2t+t^2\right)}{(l+2)\left(m_c^2-t\right)\left(M_W^2-t\right)^2} \\
& -\frac{2t\left(m_c^6-m_c^4t+m_c^2t\left(M_W^2+t\right)-2M_W^4t\right)}{(l+1)\left(t-m_c^2\right)\left(M_W^2-t\right)}\Bigg)\left(\frac{\left(m_c^2-t\right)\left(t-M_W^2\right)}{St}\right)^{-l} \\
& +\frac{2S\left(m_c^6-m_c^4t+m_c^2t\left(M_W^2+t\right)-2M_W^4t\right)}{l+1}\Bigg],
\end{aligned} \tag{6.50}$$

which, in the $m_c^2 \to 0$ limit, tends to

$$\frac{d\sigma^{(m_c=0)}}{dt} = \frac{\alpha_s G_F V_{cs}^2}{6\sqrt{2}l\,(l^2+3l+2)}\frac{1}{S^3 t^2} \left[\left(-(l^2+l+2)M_W^4 + 2(l^2+2l+2)M_W^2 t - 2(l+1)^2 t^2\right)\left(\frac{S}{M_W^2 - t}\right)^{2+l} + \left(-2l(l+2)M_W^2 S + l(l+1)(2M_W^4 - 2M_W^2 t + t^2) + (l+1)(l+2)S^2\right)\right]. \tag{6.51}$$

In the limit $t \ll M_W^2$ and $m_c \ll |t|$ Eq. (6.50) simplifies to

$$\frac{d\sigma}{dt} \sim \frac{M_W^2}{t}\left[\log\left(\frac{S}{M_W^2}\right) + \left(1 - \frac{M_W^2}{S}\right)^2\right]. \tag{6.52}$$

The logarithm depends on the ratio S/M_W^2. Plotting Eqs.(6.50, 6.51) as a function of $\sqrt{|t|}/m_c$ in Fig. 6.18, we observe that the collinear plateau drops at a scale smaller than M_W^2. The scale where the bulk of the events is concentrated is in a region about $M_W/mc \sim 20$.

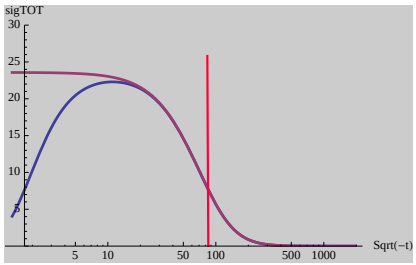


Figure 6.18: Plot of $d\sigma/dt$ as a function of $\sqrt{|t|}/m_c$ for l=2, $\sqrt{S} = 10$ TeV, $M_W = 80.398$ GeV for $m_c = 1.3$ GeV (pink) and $m_c = 0$ (blue).

In order to cross–check the result and assess whether the latter is modified by using a realistic input set of PDFs, we performed numerically the integration of Eq. (6.49). The code we used to perform the numerical integration has been cross–checked against other independent codes. The distribution of the events in t_1 is shown in Fig. 6.19, where the $t\frac{d\sigma}{dt}$ distribution is plotted as a function of $\sqrt{|t|}$ for $m_c = 0$ and for several values of $m_c \neq 0$. We see that the distribution presents a collinear plateau, damped near $t \sim 0$ for finite values of m_c. We observe that the scale at which the plateau drops is much smaller than M_W^2. This confirms the results obtained in the analytical

calculation and suggests that the scale associated to the splitting of the gluon is softer

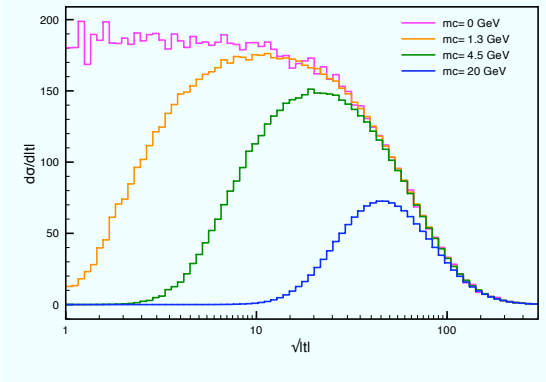


Figure 6.19: $\log(t)$ distribution for $\sqrt{S} = 1.96$ TeV, $M_W = 80.398$ GeV and various values of m_c. The distribution is not normalised

that the scale associated to the hard W production process, whose scale is naturally given by M_W^2. This implies that the factorisation scale that should be set for the bottom PDFs in the four–flavor scheme is smaller than M_W^2, which is what has been suggested in several studies presented in Sect. 7.1. Here we pointed out the reason why this choice is meaningful and we clearly showed its dynamical origin.

Conclusion

As pointed out several times in this thesis, the search for new physics at hadron colliders requires precise results in QCD phenomenology. The theoretical uncertainties in cross sections for hadron–hadron collisions is often dominated by the uncertainties of parton distribution functions. Their faithful estimation is therefore essential for providing reliable predictions at the LHC. Another source of uncertainty is related to the treatment of heavy quark masses in the calculation of processes involving heavy quarks. We have seen that there are several schemes for performing such calculations. Often they lead to predictions which differ by a significant amount, resulting in ambiguities in the theoretical description of this class of processes.

My research activity has focused on understanding and controlling the above sources of theoretical errors. On one side this thesis presents the results that I achieved within the NNPDF collaboration. The method, based on the combination of a Monte Carlo sampling of the probability measure in the space of PDFs with the use of neural networks as unbiased interpolating functions, provides a statistically–sound determination of parton densities. In this thesis I presented the progress that has been made during the last four years, the results obtained and the analyses carried out. They give an important contribution to the understanding of the role of partonic uncertainties in the high–energy predictions and help in clarifying several issues which cannot be dealt by mean of the traditional approaches. Moreover several phenomenological studies performed using the NNPDF method, have been presented and their relevance has been outlined. A related but somewhat independent result concerns the study of processes with heavy quarks in the initial states. I have introduced a method to assess the size of the potentially large logarithms of the mass of the heavy quarks which arise in massive calculations. The identification of the actual scale of these logarithms and the understanding of their dynamical origin enabled us to quantitatively estimate

their impact for several representative LHC processes. The analysis might be easily extended to any other process.

Parton densities determination

The NNPDF collaboration has developed a completely original method for determining parton distribution functions. Before the work carried out in this thesis, the NNPDF approach was only applied to the determination of a single parton density [67]. My main research project aimed at extending the method to the determination of a full set of PDFs. On top of the DIS data, upon which the first two fits were based [68, 70], inclusive jets and Drell–Yan data have recently been included, by yielding the first NNPDF global analysis [71]. In the latter, hadronic cross–sections are evaluated at NLO, without relying on any K–factor approximation thanks to the formulation of the FastKernel method. Moreover normalisations uncertainties are correctly taken into account by mean of the so–called t_0 prescription proposed in Ref. [118]. The faithful determination of PDFs uncertainties allowed us to perform a series of statistical analyses which cannot be easily addressed by other methods. In particular the dependence on the choice of the parametrisation of the PDFs has been assessed in a statistical way and satisfying answers when comparing the results obtained out of a reduced set of data with the global results have been obtained: while the uncertainty bands do increase in the regions where there are less data, and thus less information, the central values of the reduced and the full fits remain compatible within the uncertainty. Finally the compatibility between DIS and hadronic datasets has been studied in a quantitative way.

The most recent NNPDF global analysis however, still uses an approximate prescription for the treatment of heavy quarks masses, which might lead to systematic shifts in theoretical predictions. For this reason, the FONLL general mass VFN scheme [54] is currently under implementation and is going to be included in the forthcoming NNPDF2.1 release [105]. The inclusion of a better treatment of the heavy quark masses is particularly interesting within the unbiased NNPDF approach, since it allows us to disentangle its actual effect from other possible contributions due to PDFs parametrisation. Other features, such as the implementation of a NNLO analysis for precision studies, or the inclusion of small– and large–x resummation effects, are going to be explored in the near future. Finally, a suitable LO parton fit able to match the requirements of the leading–order Monte Carlo event generators is in preparation.

Related phenomenology

With the technique developed within the NNPDF approach, several phenomenological studies have been performed and presented in this thesis. The detailed study of the strange content of the nucleon showed that better control of the uncertainty of the strange and anti-strange parton distributions has a twofold significance. First, it enabled us to reassess the determination of electroweak parameters from the NuTeV dimuon data and to solve the well-known NuTeV anomaly, by making these results fully consistent with the precision electroweak data. Secondly, it yielded the most precise direct determination of the CKM matrix element $|V_{cs}|$ [70]. It would be interesting to perform such analysis with the most recent NNPDF parton set [71] and check whether the accuracy in the CKM matrix elements determination may further improve. The same kind of analysis might be performed by varying α_s and obtaining a direct determination from the parton fit. In performing such analysis, all details presented in this thesis about the correlation between α_s and PDFs must be carefully taken into account.

Another application of the Monte Carlo representation of the NNPDF results is the so–called Bayesian Reweighting technique. In this thesis the method has been explained in detail and the technique has been applied to the inclusion of the D0 W–lepton asymmetry data into the NNPDF fit without need of refitting. The analysis deserves to be extended to more W lepton asymmetry datasets, by including the most recent D0 muon data and the separate electron bins data that have not been considered in this preliminary study. This would enable us to compare the results to those obtained by the MSTW and CTEQ collaborations, where a tension is found between these data and some fixed target DIS data. Furthermore the technique might be easily applied to the inclusion of future experiments pseudo–data and used to study how they might further constrain the parton content of the proton.

On top of what has been presented in this thesis, the NNPDF method has a much wider range of applications, thanks to the generality of its features. It might be implemented whenever a function or a set of functions cannot be derived by first principles and must be inferred from a set of experimental data and theoretical assumptions, for instance in cosmology problems or in saturation models.

Heavy quark processes

As mentioned several times in this thesis, there are two complementary approaches in performing a QCD calculation of processes initiated by a heavy quark. The heavy

quark is either considered a massless parton in the initial state; or it does not contribute to the proton wavefunction and it is generated as a massive final state. In the latter scheme, possibly large logarithms developing in the initial states due to heavy quarks collinear splittings are not resummed into the heavy quark PDF. The sizeable difference observed between the two computations for several key processes at the LHC requires a better understanding of the size of the scale associated to the collinear logarithms and its dynamical origin. This was the aim of the preliminary analysis presented in this thesis.

By considering three representative processes, the heavy quark pair production, the single top and the associate W and c productions, I formulated a general method for gaining a better understanding of the difference between massless and massive computations. I showed that the scale of the logarithms arising due to the collinear splitting of heavy quarks is a dynamical scale which depends on the final state particles momenta. In the case of DIS heavy quarks production, the dynamical scale is actually peaked about the expected Q^2 scale, while in the other cases, when the phase space of final particles is suppressed by the production of a massive particle, such as the top quark or the W boson, this scale is peaked at a softer scale than the expected hard m_t or M_W scale. My analysis of the distributions not only reinforce the idea that the origin of leading–order discrepancies between the two schemes may be attributed to a choice of factorisation scales which is too large with respect to the actual scale associated to the splitting process, but it gives a quantitative explanation for that.

The simplicity of the analysis presented in this thesis allows us to extend the results obtained to relevant processes that were mentioned in the phenomenological review, such as the Higgs production induced by bottom quarks, or the analogous Z production. In this way the generality of the method proposed may be assessed and more final conclusions may be drawn.

Appendix A

Notation

The classical Lagrangian corresponding to QCD, is given by the Yang–Mill Lagrangian

$$\mathcal{L}_{\text{class}} = \sum_{\text{flavors}} \bar{\Psi}_a \left(i\gamma_\mu D^\mu - m\right)_{ab} \Psi_b - \frac{1}{4}\text{Tr}\, G^A_{\mu\nu} G^{\mu\nu}_A \,, \tag{A.1}$$

where Ψ_a are the quark fields, $G^A_{\mu\nu}$ is the field strenght tensor derived from the gluon field A^A

$$G^A_{\mu\nu} = \left[\partial_\mu A^A_\nu - \partial_\nu A^A_\mu - g f^{ABC} A^B_\mu A^C_\nu\right], \tag{A.2}$$

and D^μ is the covariant derivative:

$$(D^\mu)_{ab} = \partial^\mu \delta_{ab} + ig(t^C A^C_\alpha)_{ab}, \qquad (D^\mu)_{AB} = \partial^\mu \delta_{AB} + ig(T^C A^C_\alpha)_{AB}, \tag{A.3}$$

where t^C and T^C are the SU(3) generator matrices in the fundamental and adjoint representations, respectively:

$$[t^A, t^B] = if^{ABC}t^C, \qquad [T^A, T^B] = if^{ABC}T^C, \qquad (T^A)_{BC} = -if^{ABC}. \tag{A.4}$$

By convention, the normalization of the $SU(N)$ matrices is chosen to be

$$\text{Tr}\, t^A t^B = T_R \delta^{AB}, \qquad T_R = \frac{1}{2}. \tag{A.5}$$

With this choice, the color matrices obey the following relations:

$$\begin{aligned} \sum_A t^A_{ab} t^A_{bc} &= C_F \delta_{ab} \quad C_F = \frac{N^2-1}{2N} \\ \mathrm{Tr}\, T^C T^D &= \sum_{A,B} f^{ABC} f^{ABD} = C_A \delta^{CD} \quad C_A = N. \end{aligned} \tag{A.6}$$

For the specific case of $SU(3)$ the structure constants are: $C_F = 4/3, C_A = 3$.

Appendix B

Statistical Estimators

B.1 Monte Carlo statistical estimators

Here I define the Monte Carlo statistical estimators used in the analyses presented in this thesis.

- Central value of the i-th experimental point

$$\left\langle F_i^{(\mathrm{net})} \right\rangle_{\mathrm{rep}} = \frac{1}{N_{\mathrm{rep}}} \sum_{k=1}^{N_{\mathrm{rep}}} F_i^{(\mathrm{net})(k)} \, . \tag{B.1}$$

- Variance of the i-th experimental point

$$\sigma_i^{(\mathrm{net})} = \sqrt{\left\langle \left(F_i^{(\mathrm{net})} \right)^2 \right\rangle_{\mathrm{rep}} - \left\langle F_i^{(\mathrm{net})} \right\rangle_{\mathrm{rep}}^2} \, . \tag{B.2}$$

- Covariance and correlation of the i-th experimental point

$$\rho_{ij}^{(\mathrm{net})} = \frac{\left\langle F_i^{(\mathrm{net})} F_j^{(\mathrm{net})} \right\rangle_{\mathrm{rep}} - \left\langle F_i^{(\mathrm{net})} \right\rangle_{\mathrm{rep}} \left\langle F_j^{(\mathrm{net})} \right\rangle_{\mathrm{rep}}}{\sigma_i^{(\mathrm{net})} \sigma_j^{(\mathrm{net})}} \, . \tag{B.3}$$

$$\mathrm{cov}_{ij}^{(\mathrm{net})} = \rho_{ij}^{(\mathrm{net})} \sigma_i^{(\mathrm{net})} \sigma_j^{(\mathrm{net})} . \tag{B.4}$$

- Mean variance and percentage error on central values over the $N_{\rm dat}$ data points.

$$\left\langle V\left[\left\langle F^{(\rm net)}\right\rangle_{\rm rep}\right]\right\rangle_{\rm dat} = \frac{1}{N_{\rm dat}}\sum_{i=1}^{N_{\rm dat}}\left(\left\langle F_i^{(\rm net)}\right\rangle_{\rm rep} - F_i^{(\rm exp)}\right)^2 , \quad \text{(B.5)}$$

$$\left\langle PE\left[\left\langle F^{(\rm net)}\right\rangle_{\rm rep}\right]\right\rangle_{\rm dat} = \frac{1}{N_{\rm dat}}\sum_{i=1}^{N_{\rm dat}}\left[\frac{\left\langle F_i^{(\rm net)}\right\rangle_{\rm rep} - F_i^{(\rm exp)}}{F_i^{(\rm exp)}}\right] . \quad \text{(B.6)}$$

- $\left\langle V\left[\left\langle\sigma^{(\rm net)}\right\rangle_{\rm rep}\right]\right\rangle_{\rm dat}$, $\left\langle V\left[\left\langle\rho^{(\rm net)}\right\rangle_{\rm rep}\right]\right\rangle_{\rm dat}$, $\left\langle V\left[\left\langle{\rm cov}^{(\rm net)}\right\rangle_{\rm rep}\right]\right\rangle_{\rm dat}$ $\left\langle PE\left[\left\langle\sigma^{(\rm net)}\right\rangle_{\rm rep}\right]\right\rangle_{\rm dat}$, $\left\langle PE\left[\left\langle\rho^{(\rm net)}\right\rangle_{\rm rep}\right]\right\rangle_{\rm dat}$, $\left\langle PE\left[\left\langle{\rm cov}^{(\rm net)}\right\rangle_{\rm rep}\right]\right\rangle_{\rm dat}$ relative to errors, correlations and covariances are defined in the same way. They indicate how close are the averages over generated data and the experimental values.

- Scatter correlation:

$$r\left[F^{(\rm net)}\right] = \frac{\left\langle F^{(\rm exp)}\left\langle F^{(\rm net)}\right\rangle_{\rm rep}\right\rangle_{\rm dat} - \left\langle F^{(\rm exp)}\right\rangle_{\rm dat}\left\langle\left\langle F^{(\rm net)}\right\rangle_{\rm rep}\right\rangle_{\rm dat}}{\sigma_s^{(\rm exp)}\sigma_s^{(\rm net)}} , \quad \text{(B.7)}$$

  where the scatter variances are defined as

$$\sigma_s^{(\rm exp)} = \sqrt{\left\langle\left(F^{(\rm exp)}\right)^2\right\rangle_{\rm dat} - \left(\left\langle F^{(\rm exp)}\right\rangle_{\rm dat}\right)^2} , \quad \text{(B.8)}$$

$$\sigma_s^{(\rm net)} = \sqrt{\left\langle\left(\left\langle F^{(\rm net)}\right\rangle_{\rm rep}\right)^2\right\rangle_{\rm dat} - \left(\left\langle\left\langle F^{(\rm net)}\right\rangle_{\rm rep}\right\rangle_{\rm dat}\right)^2} . \quad \text{(B.9)}$$

- $r\left[\sigma^{(\rm net)}\right]$ $r\left[\rho^{(\rm net)}\right]$ $r\left[{\rm cov}^{(\rm net)}\right]$ are defined in the same way.
  The scatter correlation indicates the size of the spread of data around a straight line. Specifically $r\left[\sigma^{(\rm net)}\right] = 1$ implies that $\left\langle\sigma_i^{(\rm net)}\right\rangle$ is proportional to $\sigma_i^{(\rm exp)}$.

- Average variance:

$$\left\langle\sigma^{(\rm net)}\right\rangle_{\rm dat} = \frac{1}{N_{\rm dat}}\sum_{i=1}^{N_{\rm dat}}\sigma_i^{(\rm net)} . \quad \text{(B.10)}$$

- $\left\langle\rho^{(\rm net)}\right\rangle_{\rm dat}$ $\left\langle{\rm cov}^{(\rm net)}\right\rangle_{\rm dat}$ are defined in the same way.

B.2 Distances

Given a set of $N_{\rm rep}^{(k)}$ replicas $q_i^{(k)}$ of some quantity q, the estimator for the expected (true) value of q is the mean

$$\langle q^{(k)} \rangle_{(i)} = \frac{1}{N_{\rm rep}^{(i)}} \sum_{i=1}^{N_{\rm rep}^{(i)}} q_i^{(k)}. \tag{B.11}$$

The square distance between the two estimates of the expected value obtained from sets $q_i^{(1)}$, $q_i^{(2)}$ is then

$$d^2\left(\langle q^{(1)} \rangle, \langle q^{(2)} \rangle\right) = \frac{\left(\langle q^{(1)} \rangle_{(1)} - \langle q^{(2)} \rangle_{(2)}\right)^2}{\sigma_{(1)}^2[\langle q^{(1)} \rangle] + \sigma_{(2)}^2[\langle q^{(2)} \rangle]} \tag{B.12}$$

where the variance of the mean is given by

$$\sigma_{(i)}^2[\langle q^{(i)} \rangle] = \frac{1}{N_{\rm rep}^{(i)}} \sigma_{(i)}^2[q^{(i)}] \tag{B.13}$$

in terms of the variance $\sigma_{(i)}^2[q^{(i)}]$ of the variables $q^{(i)}$ (which a priori could come from two distinct probability distributions).

We estimate the variance of the mean from the variance of the replica sample as

$$\sigma_{(i)}^2[q^{(i)}] = \frac{1}{N_{\rm rep}^{(i)} - 1} \sum_{k=1}^{N_{\rm rep}^{(i)}} \left(q_k^{(i)} - \langle q^{(i)} \rangle\right)^2, \tag{B.14}$$

with $\langle q^{(i)} \rangle$ given by Eq. (3.34).

Given a set of $N_{\rm rep}^{(k)}$ replicas $q_i^{(k)}$ of some quantity q, the estimator for the square uncertainty of q is the variance of the replica sample given by Eq. (B.14). The distance between the two estimates of the square uncertainty obtained from sets $q_i^{(1)}$, $q_i^{(2)}$ is then

$$d^2(\sigma_{(1)}^2, \sigma_{(2)}^2) = \frac{\left(\bar{\sigma}_{(1)}^2 - \bar{\sigma}_{(2)}^2\right)^2}{\sigma_{(1)}^2[\bar{\sigma}_{(1)}^2] + \sigma_{(2)}^2[\bar{\sigma}_{(2)}^2]} \tag{B.15}$$

where for brevity we have defined

$$\bar{\sigma}_{(i)}^2 \equiv \sigma_{(i)}^2[q^{(i)}]. \tag{B.16}$$

The variances $\sigma^2_{(i)}[\bar{\sigma}^2_{(i)}]$ of the square uncertainties could also be estimated from the replica sample, by computing the variance from various subsets of the given replica sample, and then the variance of these resulting variances as the subset is varied; for finite number of replicas this may lead to loss of statistical accuracy. For simplicity here we use instead the expression [189]

$$\sigma^2_{(i)}[\bar{\sigma}^2_{(i)}] = \frac{1}{N^{(i)}_{\rm rep}} \left[m_4[q^{(i)}] - \frac{N^{(i)}_{\rm rep}-3}{N^{(i)}_{\rm rep}-1} \left(\bar{\sigma}^2_{(i)} \right)^2 \right], \tag{B.17}$$

where as above $\bar{\sigma}^2_{(i)}$ is estimated using Eq. (B.14), while the fourth moment m_4 of the probability distribution is estimated from the corresponding moment of the replica sample (which provides an estimate of it which is only asymptotically unbiased):

$$m_4[q^{(i)}] = \frac{1}{N^{(i)}_{\rm rep}} \sum_{k=1}^{N^{(i)}_{\rm rep}} \left(q^{(i)}_k - \langle q^{(i)} \rangle \right)^4. \tag{B.18}$$

In practice, for small–sized replica samples the distances defined in Eq. (B.12) and Eq. (B.15) display sizable statistical fluctuations. In order to stabilise the result, distances are determined as follows: we randomly pick $N^{(i)}_{\rm rep}/2$ out of the $N^{(i)}_{\rm rep}$ replicas for each of the two subsets. The computation of the square distance Eq. (3.35) or Eq. (B.15) is then repeated for $N_{\rm part} = 100$ (randomly generated) choices of $N^{(i)}_{\rm rep}/2$ replicas, and the result is averaged: this is sufficient to bring the statistical fluctuations of the distance at the level of a few percent.

Appendix C

PDFs uncertainty computation

Within the NNPDF approach the expectation value of any function $\mathcal{O}[\{q\}]$ which depends on the PDFs is computed as an average over the ensemble of PDFs, using the master formula

$$\langle \mathcal{O}[\{q\}] \rangle = \frac{1}{N_{\rm set}} \sum_{k=1}^{N_{\rm set}} \mathcal{O}[\{q\}], \tag{C.1}$$

where $N_{\rm set} = N_{\rm rep}$ is the number of sets of PDFs in the ensemble, equal to the number of replicas. The associated uncertainty is found as the standard deviation of the sample, according to the usual formula

$$\begin{aligned} \sigma_{\mathcal{O}} &= \left(\frac{N_{\rm set}}{N_{\rm set}-1} \left(\langle \mathcal{O}[\{q\}]^2 \rangle - \langle \mathcal{O}[\{q\}] \rangle^2 \right) \right)^{1/2} \\ &= \left(\frac{1}{N_{\rm set}-1} \sum_{k=1}^{N_{\rm set}} \left(\mathcal{O}[\left\{ q^{(k)} \right\}] - \langle \mathcal{O}[\{q\}] \rangle \right)^2 \right)^{1/2}. \end{aligned} \tag{C.2}$$

These formulae may also be used for the determination of central values and uncertainties of the parton distribution themselves, in which case the functional $\mathcal{O}$ is identified with the parton distribution $q : \mathcal{O}[\{q\}] \equiv q$.

Here we briefly compare the procedures we use to determine central values and errors with those used for the various other PDFs available through HEPDATA. Available methods for the determination of PDF uncertainties fall broadly into two distinct categories, which we shall refer to as the HEPDATA method (used as a default in the PDF

server at the HEPDATA database) and the Monte Carlo method. In both methods sets of PDFs with uncertainties are given as an ensemble of $N_{\rm set}$ sets of PDFs,

$$\left\{q^{(k)}\right\} , \qquad k = 0, \ldots, N_{\rm set}. \tag{C.3}$$

Conventionally the PDF set $q^{(0)}$ corresponds to a "central" set.

In the HEPDATA method, the central set is a best fit set of PDFs, which thus provides the central value for PDFs themselves. The central value of any quantity $\mathcal{O}[\{q\}]$ is obtained in this method by evaluating it as a function of the central set:

$$\mathcal{O}^{(0)} = \mathcal{O}[\left\{q^{(0)}\right\}]. \tag{C.4}$$

In the Monte Carlo method, the central values of any quantity $\mathcal{O}[\{q\}]$ is instead given by Eq. (C.1). However, for any quantity $\mathcal{O}[\{q\}]$ which depends nonlinearly on the PDFs

$$\langle \mathcal{O}[\{q\}] \rangle \neq \mathcal{O}[\left\{q^{(0)}\right\}]. \tag{C.5}$$

Hence, Eq. (C.1) must be used for the determination of the central value, and use of the set $q^{(0)}$ is not recommended. However, for a quantity that does depend linearly on the PDFs, such as a DIS structure function, Eq. (C.4) with the central PDFs $q^{(0)}$ gives the same result as Eq. (C.1), and thus it may be used also with the Monte Carlo method. Note that set $q^{(0)}$ should not be included when computing an average with Eq. (C.1), because it is itself already an average.

The determination of uncertainties with the HEPDATA method is based on the idea that sets $q^{(k)}$ with $k > 0$ provide upper and lower variations (for even and odd values of k) away from the central set $q^{(0)}$ which correspond to eigenvectors in parameter space. The one-σ uncertainty is then found by adding in quadrature these variations:

$$\sigma_{\mathcal{O}}^{\rm hepdata} = \frac{1}{2C_{90}} \left(\sum_{k=1}^{N_{\rm set}/2} \left(\mathcal{O}[\left\{q^{(2k-1)}\right\}] - \mathcal{O}[\left\{q^{(2k)}\right\}] \right)^2 \right)^{1/2} , \tag{C.6}$$

where the factor

$$C_{90} \equiv \sqrt{2}{\rm Erf}^{-1}[0.90] = 1.64485 \tag{C.7}$$

accounts for the fact that the upper and lower parton sets correspond to 90% confidence levels rather than to one-σ uncertainties. This method should be used with the CTEQ and MRST/MSTW sets.

A slightly different application of the HEPDATA method is required for the Alekhin/ABKM PDF sets Ref. [73, 74, 75, 102]. With these PDFs, sets $q^{(k)}$ with $k > 0$ each provide the uncertainty limits from the central set, with upper and lower PDFs symmetrical by construction and already corresponding to one-σ uncertainties. So for these PDFs

$$\sigma_{\mathcal{F}}^{\rm hepdata} = \left(\sum_{k=1}^{N_{\rm set}} \left(\mathcal{F}[\left\{ q^{(k)} \right\}] - \mathcal{F}[\left\{ q^{(0)} \right\}] \right)^2 \right)^{1/2} . \tag{C.8}$$

Appendix D

$\mathcal{O}(\alpha_s)$ matrix elements for heavy quark production

In this appendix we provide the explicit expression for the matrix elements defined in Eq. (6.14). They refer to leading–order process

$$g(p_b) + B^*(q) \;\rightarrow\; Q_M(k_2) + \bar{Q}_m(k_3),$$

where Q_M and Q_m are heavy quarks of masses M and m respectively whose coupling with the generic vector boson B^* is given by

$$g_R\gamma^\mu\frac{1+\gamma_5}{2} + g_L\gamma^\mu\frac{1-\gamma_5}{2} \equiv g_R\gamma^\mu P_R + g_L\gamma^\mu P_L.$$

The transverse component is given by

$$\begin{aligned}
\mathcal{M}_T^g \;=\; & \frac{4}{t_1^2u_1^2}\Bigg\{g_L^2\Bigg[+2m^2M^2s_1^2 - 2s_1\left[m^4u_1 + M^4t_1 + (m^2+M^2)u_1t_1\right] \\
& +t_1u_1\left(2Q^4 - 2s_1Q^2 + t_1^2 + u_1^2\right) + 2Q^2\left[m^2u_1(2t_1+u_1) + M^2t_1(2u_1+t_1)\right]\Bigg] \\
& +16g_Lg_RmM\Big[s_1(u_1m^2 + t_1M^2) + u_1t_1(s_1 - Q^2)\Big] \\
& +g_R^2\Bigg[+2m^2M^2s_1^2 - 2s_1\left[m^4u_1 + M^4t_1 + (m^2+M^2)u_1t_1\right] \\
& +t_1u_1\left(2Q^4 - 2s_1Q^2 + t_1^2 + u_1^2\right) + 2Q^2\left[m^2u_1(2t_1+u_1) + M^2t_1(2u_1+t_1)\right]\Bigg]\Bigg\}.
\end{aligned} \tag{D.1}$$

The longitudinal component is given by

$$\mathcal{M}_L^g = \frac{16Q^2\left(g_L^2+g_R^2\right)}{s_1^2 t_1 u_1}\left[s_1(u_1 m^2 + t_1 M^2) + u_1 t_1 (s_1 - Q^2)\right]. \tag{D.2}$$

The axial component, which vanishes if $g_L = g_R$ is given by

$$\begin{aligned}\mathcal{M}_3^g \quad = \quad & -\frac{4Q^2\left(g_L^2-g_R^2\right)}{s_1 t_1^2 u_1^2}\Big\{(-2m^4 s_1^2 u_1 - m^2 s_1 \\ & \left(2M^2 s_1(s_1+2t_1) - u_1\left(2Q^2(s_1+3t_1) - 2s_1^2 + 2s_1\hat{s}y - 5s_1 t_1 + 3\hat{s}t_1 y\right)\right) \\ & + t_1\Big(2M^4 s_1^2 + M^2 s_1\left(4Q^2 s_1 + 6Q^2 t_1 - 3s_1^2 + s_1\hat{s}y - 5s_1 t_1 + 3\hat{s}t_1 y\right) \\ & + \left(s_1^2 + 3s_1 t_1 + 2t_1^2\right)\left(2Q^4 - 3Q^2 s_1 + Q^2\hat{s}y + s_1^2\right)\Big)\Big\}\end{aligned} \tag{D.3}$$

Bibliography

[1] Murray Gell-Mann, "Symmetries of baryons and mesons", *Phys. Rev.*, vol. 125, pp. 1067–1084, 1962.

[2] Murray Gell-Mann, "A Schematic Model of Baryons and Mesons", *Phys. Lett.*, vol. 8, pp. 214–215, 1964.

[3] G. Zweig, "AN SU(3) MODEL FOR STRONG INTERACTION SYMMETRY AND ITS BREAKING. 2", CERN-TH-412.

[4] Y. Dothan, Murray Gell-Mann, and Yuval Ne'eman, "SERIES OF HADRON ENERGY LEVELS AS REPRESENTATIONS OF NONCOMPACT GROUPS", *Phys. Lett.*, vol. 17, pp. 148–151, 1965.

[5] Richard P. Feynman, "Very high-energy collisions of hadrons", *Phys. Rev. Lett.*, vol. 23, pp. 1415–1417, 1969.

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

[7] D. J. Gross and Frank Wilczek, "Asymptotically Free Gauge Theories. 1", *Phys. Rev.*, vol. D8, pp. 3633–3652, 1973.

[8] Steven Weinberg, "Nonabelian Gauge Theories of the Strong Interactions", *Phys. Rev. Lett.*, vol. 31, pp. 494–497, 1973.

[9] Michael Edward Peskin and Daniel V. Schroeder, "An Introduction to quantum field theory", Reading, USA: Addison-Wesley (1995) 842 p.

[10] G. Dissertori, I. G. Knowles, and M. Schmelling, "High energy experiments and theory", Oxford, UK: Clarendon (2003) 538 p.

[11] R. Keith Ellis, W. James Stirling, and B. R. Webber, "QCD and collider physics", *Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol.*, vol. 8, pp. 1–435, 1996.

[12] Siegfried Bethke, "The 2009 Wolrd Average of $\alpha_s(M_Z)$", *Eur. Phys. J.*, vol. C64, pp. 689–703, 2009.

[13] D. J. Gross and Frank Wilczek, "ULTRAVIOLET BEHAVIOR OF NON-ABELIAN GAUGE THEORIES", *Phys. Rev. Lett.*, vol. 30, pp. 1343–1346, 1973.

[14] H. David Politzer, "RELIABLE PERTURBATIVE RESULTS FOR STRONG INTERACTIONS?", *Phys. Rev. Lett.*, vol. 30, pp. 1346–1349, 1973.

[15] Richard P. Feynman, "Very high-energy collisions of hadrons", *Phys. Rev. Lett.*, vol. 23, pp. 1415–1417, 1969.

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

[17] M. Klein and R. Yoshida, "Collider Physics at HERA", 2008.

[18] Nicola Cabibbo, "Unitary Symmetry and Leptonic Decays", *Phys. Rev. Lett.*, vol. 10, pp. 531–533, 1963.

[19] Makoto Kobayashi and Toshihide Maskawa, "CP Violation in the Renormalizable Theory of Weak Interaction", *Prog. Theor. Phys.*, vol. 49, pp. 652–657, 1973.

[20] T. Kinoshita, "Mass singularities of Feynman amplitudes", *J. Math. Phys.*, vol. 3, pp. 650–677, 1962.

[21] T. D. Lee and M. Nauenberg, "Degenerate Systems and Mass Singularities", *Phys. Rev.*, vol. 133, pp. B1549–B1562, 1964.

[22] Enrico C. Poggio and Helen R. Quinn, "The Infrared Behavior of Zero-Mass Green's Functions and the Absence of Quark Confinement in Perturbation Theory", *Phys. Rev.*, vol. D14, pp. 578, 1976.

[23] G. Sterman, "Kinoshita's Theorem in Yang-Mills Theories", *Phys. Rev.*, vol. D14, pp. 2123–2125, 1976.

[24] 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*, vol. 46, pp. 641–653, 1977.

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

[26] Guido Altarelli and G. Parisi, "Asymptotic Freedom in Parton Language", *Nucl. Phys.*, vol. B126, pp. 298, 1977.

[27] John C. Collins and Davison E. Soper, "The Theorems of Perturbative QCD", *Ann. Rev. Nucl. Part. Sci.*, vol. 37, pp. 383–409, 1987.

[28] W. Furmanski and R. Petronzio, "Singlet Parton Densities Beyond Leading Order", *Phys. Lett.*, vol. B97, pp. 437, 1980.

[29] S. Moch, J. A. M. Vermaseren, and A. Vogt, "Next-to-next-to-leading order QCD corrections to the photon's parton structure", *Nucl. Phys.*, vol. B621, pp. 413–458, 2002.

[30] S. Moch, J. A. M. Vermaseren, and A. Vogt, "The three-loop splitting functions in QCD: The non-singlet case", *Nucl. Phys.*, vol. B688, pp. 101–134, 2004.

[31] A. Vogt, S. Moch, and J. A. M. Vermaseren, "The three-loop splitting functions in QCD: The singlet case", *Nucl. Phys.*, vol. B691, pp. 129–181, 2004.

[32] A. Vogt, "Efficient evolution of unpolarized and polarized parton distributions with QCD-PEGASUS", *Comput. Phys. Commun.*, vol. 170, pp. 65–92, 2005.

[33] Andrzej J. Buras, "Asymptotic Freedom in Deep Inelastic Processes in the Leading Order and Beyond", *Rev. Mod. Phys.*, vol. 52, pp. 199, 1980.

[34] Scott Willenbrock, "QCD CORRECTIONS TO p anti-p $\rightarrow$ W+ + X: A CASE STUDY", Presented at Theoretical Adv. Summer Inst. (TASI), Boulder, CO, Jun 4-30, 1989.

[35] John C. Collins, Frank Wilczek, and A. Zee, "Low-Energy Manifestations of Heavy Particles: Application to the Neutral Current", *Phys. Rev.*, vol. D18, pp. 242, 1978.

[36] John C. Collins, "RENORMALIZATION. AN INTRODUCTION TO RENORMALIZATION, THE RENORMALIZATION GROUP, AND THE OPERATOR PRODUCT EXPANSION", Cambridge, Uk: Univ. Pr. (1984) 380p.

[37] John C. Collins, "Hard-scattering factorization with heavy quarks: A general treatment", *Phys. Rev.*, vol. D58, pp. 094002, 1998.

[38] J. Pumplin, H. L. Lai, and W. K. Tung, "The charm parton content of the nucleon", *Phys. Rev.*, vol. D75, pp. 054029, 2007.

[39] M. Buza, Y. Matiounine, J. Smith, and W. L. van Neerven, "Charm electroproduction viewed in the variable-flavour number scheme versus fixed-order perturbation theory", *Eur. Phys. J.*, vol. C1, pp. 301–320, 1998.

[40] K. G. Chetyrkin, Bernd A. Kniehl, and M. Steinhauser, "Strong coupling constant with flavour thresholds at four loops in the MS-bar scheme", *Phys. Rev. Lett.*, vol. 79, pp. 2184–2187, 1997.

[41] W. K. Tung et al., "Heavy quark mass effects in deep inelastic scattering and global QCD analysis", *JHEP*, vol. 02, pp. 053, 2007.

[42] Pavel M. Nadolsky et al., "Implications of CTEQ global analysis for collider observables", 2008.

[43] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, "Parton distributions for the LHC", *Eur. Phys. J.*, vol. C63, pp. 189–285, 2009.

[44] B. W. Harris and J. Smith, "Heavy quark correlations in deep inelastic electroproduction", *Nucl. Phys.*, vol. B452, pp. 109–160, 1995.

[45] Pavel M. Nadolsky and Wu-Ki Tung, "Improved Formulation of Global QCD Analysis with Zero-mass Matrix Elements", *Phys. Rev.*, vol. D79, pp. 113014, 2009.

[46] M. A. G. Aivazis, John C. Collins, Fredrick I. Olness, and Wu-Ki Tung, "Leptoproduction of heavy quarks. 2. A Unified QCD formulation of charged and neutral current processes from fixed target to collider energies", *Phys. Rev.*, vol. D50, pp. 3102–3118, 1994.

[47] Michael Kramer, 1, Fredrick I. Olness, and Davison E. Soper, "Treatment of heavy quarks in deeply inelastic scattering", *Phys. Rev.*, vol. D62, pp. 096007, 2000.

[48] Wu-Ki Tung, Stefan Kretzer, and Carl Schmidt, "Open heavy flavor production in QCD: Conceptual framework and implementation issues", *J. Phys.*, vol. G28, pp. 983–996, 2002.

[49] S. Kretzer, H. L. Lai, F. I. Olness, and W. K. Tung, "CTEQ6 parton distributions with heavy quark mass effects", *Phys. Rev.*, vol. D69, pp. 114005, 2004.

[50] Fredrick Olness and Ingo Schienbein, "Heavy Quarks: Lessons Learned from HERA and Tevatron", *Nucl. Phys. Proc. Suppl.*, vol. 191, pp. 44–53, 2009.

[51] R. S. Thorne and R. G. Roberts, "An ordered analysis of heavy flavour production in deep inelastic scattering", *Phys. Rev.*, vol. D57, pp. 6871–6898, 1998.

[52] R. S. Thorne, "A variable-flavour number scheme for NNLO", *Phys. Rev.*, vol. D73, pp. 054019, 2006.

[53] Matteo Cacciari, Mario Greco, and Paolo Nason, "The p(T) spectrum in heavy-flavour hadroproduction", *JHEP*, vol. 05, pp. 007, 1998.

[54] Stefano Forte, Eric Laenen, Paolo Nason, and Juan Rojo, "Heavy quarks in deep-inelastic scattering", *Nucl. Phys.*, vol. B834, pp. 116–162, 2010.

[55] J. R. Andersen et al., "The SM and NLO multileg working group: Summary report", 2010.

[56] R. McElhaney and S. F. Tuan, "Some consequences of a modified Kuti Weisskopf quark parton model", *Phys. Rev.*, vol. D8, pp. 2267–2272, 1973.

[57] M. Gluck and E. Reya, "Operator Mixing and Scaling Deviations in Asymptotically Free Field Theories", *Phys. Rev.*, vol. D14, pp. 3034–3044, 1976.

[58] A. J. Buras and K. J. F. Gaemers, "Simple Parametrizations of Parton Distributions with q**2 Dependence Given by Asymptotic Freedom", *Nucl. Phys.*, vol. B132, pp. 249, 1978.

[59] J. Pumplin et al., "New generation of parton distributions with uncertainties from global QCD analysis", *JHEP*, vol. 07, pp. 012, 2002.

[60] J. Huston, J. Pumplin, D. Stump, and W. K. Tung, "Stability of NLO global analysis and implications for hadron collider physics", *JHEP*, vol. 06, pp. 080, 2005.

[61] J. F. Owens et al., "The Impact of new neutrino DIS and Drell-Yan data on large-x parton distributions", *Phys. Rev.*, vol. D75, pp. 054030, 2007.

[62] H. L. Lai et al., "The Strange Parton Distribution of the Nucleon: Global Analysis and Applications", *JHEP*, vol. 04, pp. 089, 2007.

[63] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, "Uncertainties of predictions from parton distributions. I: Experimental errors. ((T))", *Eur. Phys. J.*, vol. C28, pp. 455–473, 2003.

[64] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, "Uncertainties of predictions from parton distributions. II: Theoretical errors", *Eur. Phys. J.*, vol. C35, pp. 325–348, 2004.

[65] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, "Physical gluons and high-E(T) jets", *Phys. Lett.*, vol. B604, pp. 61–68, 2004.

[66] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, "Update of Parton Distributions at NNLO", *Phys. Lett.*, vol. B652, pp. 292–299, 2007.

[67] Luigi Del Debbio, Stefano Forte, Jose I. Latorre, Andrea Piccione, and Joan Rojo, "Neural network determination of parton distributions: the nonsinglet case", *JHEP*, vol. 03, pp. 039, 2007.

[68] Richard D. Ball et al., "A determination of parton distributions with faithful uncertainty estimation", *Nucl. Phys.*, vol. B809, pp. 1–63, 2009.

[69] Juan Rojo et al., "Update on Neural Network Parton Distributions: NNPDF1.1", 2008.

[70] Richard D. Ball et al., "Precision determination of electroweak parameters and the strange content of the proton from neutrino deep-inelastic scattering", *Nucl. Phys.*, vol. B823, pp. 195–233, 2009.

[71] Richard D. Ball et al., "A first unbiased global NLO determination of parton distributions and their uncertainties", 2010.

[72] Sergey I. Alekhin, "Global fit to the charged leptons DIS data: alpha(s), parton distributions, and high twists", *Phys. Rev.*, vol. D63, pp. 094022, 2001.

[73] Sergey Alekhin, "Parton distributions from deep-inelastic scattering data", *Phys. Rev.*, vol. D68, pp. 014002, 2003.

[74] S. Alekhin, "Parton distribution functions from the precise NNLO QCD fit", *JETP Lett.*, vol. 82, pp. 628–631, 2005.

[75] Sergey Alekhin, Kirill Melnikov, and Frank Petriello, "Fixed target Drell-Yan data and NNLO QCD fits of parton distribution functions", *Phys. Rev.*, vol. D74, pp. 054033, 2006.

[76] L. W. Whitlow, E. M. Riordan, S. Dasu, Stephen Rock, and A. Bodek, "Precise measurements of the proton and deuteron structure functions from a global analysis of the SLAC deep inelastic electron scattering cross-sections", *Phys. Lett.*, vol. B282, pp. 475–482, 1992.

[77] A. C. Benvenuti et al., "A High Statistics Measurement of the Proton Structure Functions F(2) (x, Q**2) and R from Deep Inelastic Muon Scattering at High Q**2", *Phys. Lett.*, vol. B223, pp. 485, 1989.

[78] A. C. Benvenuti et al., "A HIGH STATISTICS MEASUREMENT OF THE DEUTERON STRUCTURE FUNCTIONS F2 (x, Q**2) AND R FROM DEEP INELASTIC MUON SCATTERING AT HIGH Q**2", *Phys. Lett.*, vol. B237, pp. 592, 1990.

[79] M. Arneodo et al., "Measurement of the proton and deuteron structure functions, F2(p) and F2(d), and of the ratio sigma(L)/sigma(T)", *Nucl. Phys.*, vol. B483, pp. 3–43, 1997.

[80] M. Arneodo et al., "Accurate measurement of F2(d)/F2(p) and R(d)-R(p)", *Nucl. Phys.*, vol. B487, pp. 3–26, 1997.

[81] M. R. Adams et al., "Proton and deuteron structure functions in muon scattering at 470-GeV", *Phys. Rev.*, vol. D54, pp. 3006–3056, 1996.

[82] Sarah Boutle, "Heavy quark production at HERA", *Fizika*, vol. B17, pp. 77–84, 2008.

[83] F. D. Aaron et al., "Measurement of the Proton Structure Function F_L at Low x", *Phys. Lett.*, vol. B665, pp. 139–146, 2008.

[84] F. D. Aaron et al., "Combined Measurement and QCD Analysis of the Inclusive ep Scattering Cross Sections at HERA", *JHEP*, vol. 01, pp. 109, 2010.

[85] S. Chekanov et al., "Measurement of high-Q**2 neutral current deep inelastic $e^- p$ scattering cross sections with a longitudinally polarised electron beam at HERA", *Eur. Phys. J.*, vol. C62, pp. 625–658, 2009.

[86] Un-Ki Yang et al., "Measurements of F_2 and $xF_3^\nu - xF_3^{\bar{\nu}}$ from CCFR $\nu_\mu-$Fe and $\bar{\nu}_\mu-$Fe data in a physics model independent way", *Phys. Rev. Lett.*, vol. 86, pp. 2742–2745, 2001.

[87] A. O. Bazarko et al., "Determination of the strange quark content of the nucleon from a next-to-leading order QCD analysis of neutrino charm production", *Z. Phys.*, vol. C65, pp. 189–198, 1995.

[88] G. Onengut et al., "Measurement of nucleon structure functions in neutrino scattering", *Phys. Lett.*, vol. B632, pp. 65–75, 2006.

[89] M. Goncharov et al., "Precise measurement of dimuon production cross-sections in nu/mu Fe and anti-nu/mu Fe deep inelastic scattering at the Tevatron", *Phys. Rev.*, vol. D64, pp. 112006, 2001.

[90] G. Moreno et al., "Dimuon production in proton - copper collisions at $\sqrt{s}$ = 38.8-GeV", *Phys. Rev.*, vol. D43, pp. 2815–2836, 1991.

[91] P. L. McGaughey et al., "Cross-sections for the production of high mass muon pairs from 800-GeV proton bombardment of H-2", *Phys. Rev.*, vol. D50, pp. 3038–3045, 1994.

[92] J. C. Webb et al., "Absolute Drell-Yan dimuon cross sections in 800-GeV/c p p and p d collisions", 2003.

[93] Jason C. Webb, "Measurement of continuum dimuon production in 800-GeV/c proton nucleon collisions", 2003.

[94] R. S. Towell et al., "Improved measurement of the anti-d/anti-u asymmetry in the nucleon sea", *Phys. Rev.*, vol. D64, pp. 052002, 2001.

[95] T. Aaltonen et al., "Direct Measurement of the W Production Charge Asymmetry in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV", *Phys. Rev. Lett.*, vol. 102, pp. 181801, 2009.

[96] V. M. Abazov et al., "Measurement of the shape of the boson rapidity distribution for $p\bar{p} \rightarrow Z/gamma^* \rightarrow e^+e^-$ + X events produced at $\sqrt{s}$ of 1.96-TeV", *Phys. Rev.*, vol. D76, pp. 012003, 2007.

[97] T. Aaltonen et al., "Measurement of $d\sigma/dy$ of Drell-Yan e^+e^- pairs in the Z Mass Region from $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV", 2009.

[98] Joachim Meyer, "Jets in deep inelastic scattering at HERA".

[99] A. Abulencia et al., "Measurement of the Inclusive Jet Cross Section using the $\boldsymbol{k_T}$ algorithmin$\boldsymbol{p\overline{p}}$ Collisions at$\boldsymbol{\sqrt{s}}$ = 1.96 TeV with the CDF II Detector", *Phys. Rev.*, vol. D75, pp. 092006, 2007.

[100] V. M. Abazov et al., "Measurement of the inclusive jet cross-section in $p\bar{p}$ collisions at $s^{91/2)}$ =1.96-TeV", *Phys. Rev. Lett.*, vol. 101, pp. 062001, 2008.

[101] M. Dittmar et al., "Parton distributions: Summary report", 2005.

[102] S. Alekhin, J. Blumlein, S. Klein, and S. Moch, "The 3-, 4-, and 5-flavor NNLO Parton from Deep-Inelastic- Scattering Data and at Hadron Colliders", *Phys. Rev.*, vol. D81, pp. 014032, 2010.

[103] H. L. Lai and W. K. Tung, "Charm production and parton distributions", *Z. Phys.*, vol. C74, pp. 463–468, 1997.

[104] H. L. Lai et al., "Global QCD analysis of parton structure of the nucleon: CTEQ5 parton distributions", *Eur. Phys. J.*, vol. C12, pp. 375–392, 2000.

[105] Richard D. Ball et al., "An unbiased determination of parton distributions with heavy quark mass effects".

[106] Jon Pumplin, "Parametrization dependence and Delta Chi-squared in parton distribution fitting", 2009.

[107] Stanley J. Brodsky and Glennys R. Farrar, "Scaling Laws at Large Transverse Momentum", *Phys. Rev. Lett.*, vol. 31, pp. 1153–1156, 1973.

[108] F. D. Aaron et al., "A Precision Measurement of the Inclusive ep Scattering Cross Section at HERA", *Eur. Phys. J.*, vol. C64, pp. 561–587, 2009.

[109] R. Michael Barnett et al., "Review of particle physics. Particle Data Group", *Phys. Rev.*, vol. D54, pp. 1–720, 1996.

[110] Walter T. Giele and Stephane Keller, "Implications of hadron collider observables on parton distribution function uncertainties", *Phys. Rev.*, vol. D58, pp. 094023, 1998.

[111] Walter T. Giele, Stephane A. Keller, and David A. Kosower, "Parton distribution function uncertainties", 2001.

[112] J. Pumplin et al., "Uncertainties of predictions from parton distribution functions. 2. The Hessian method", *Phys. Rev.*, vol. D65, pp. 014013, 2001.

[113] John M. Campbell, J. W. Huston, and W. J. Stirling, "Hard Interactions of Quarks and Gluons: A Primer for LHC Physics", *Rept. Prog. Phys.*, vol. 70, pp. 89, 2007.

[114] John C. Collins and Jon Pumplin, "Tests of goodness of fit to multiple data sets", 2001.

[115] Amanda M. Cooper-Sarkar, "Uncertainties on parton distribution functions from the ZEUS NLO QCD fit to data on deep inelastic scattering", *J. Phys.*, vol. G28, pp. 2669–2678, 2002.

[116] J. Pumplin, D. R. Stump, and W. K. Tung, "Multivariate fitting and the error matrix in global analysis of data", *Phys. Rev.*, vol. D65, pp. 014011, 2001.

[117] G. D'Agostini, "Bayesian reasoning in data analysis: A critical introduction", New Jersey, USA: World Scientific (2003) 329 p.

[118] Richard D. Ball et al., "Fitting Experimental Data with Multiplicative Normalization Uncertainties", 2009.

[119] Alan D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, "MRST2001: Partons and α_s from precise deep inelastic scattering and Tevatron jet data", *Eur. Phys. J.*, vol. C23, pp. 73–87, 2002.

[120] S. Chekanov et al., "Measurement of the neutral current cross section and F2 structure function for deep inelastic e+ p scattering at HERA", *Eur. Phys. J.*, vol. C21, pp. 443–471, 2001.

[121] C. Adloff et al., "Deep-inelastic inclusive e p scattering at low x and a determination of alpha(s)", *Eur. Phys. J.*, vol. C21, pp. 33–61, 2001.

[122] M. Dittmar et al., "Parton Distributions", 2009.

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

[124] C. Adloff et al., "Measurement of neutral and charged current cross-sections in positron proton collisions at large momentum transfer", *Eur. Phys. J.*, vol. C13, pp. 609–639, 2000.

[125] C. Adloff et al., "Measurement of neutral and charged current cross sections in electron proton collisions at high Q**2", *Eur. Phys. J.*, vol. C19, pp. 269–288, 2001.

[126] W. Giele et al., "The QCD / SM working group: Summary report", 2002.

[127] Heli Honkanen and Simonetta Liuti, "New approach to the Parton Distribution Functions: Self- Organizing Maps", *PoS*, vol. LC2008, pp. 022, 2008.

[128] Carsten Peterson and Thorsteinn Rognvaldsson, "An Introduction to artificial neural networks", Lectures given at 1991 CERN School of Computing, Ystad, Sweden, Aug 23 - Sep 2, 1991.

[129] Warren S. McCulloch and Walter Pitts, "A logical calculus of the ideas immanent in nervous activity", *Bull. Math. Biophys.*, vol. 5, pp. 115–133, 1943.

[130] G. Cybenko, "Approximation by superpositions of a sigmoidal function", *Math. Control Signals Systems*, vol. 2, no. 4, pp. 303–314, 1989.

[131] Youli Andreev Kanev, "Application of neural networks and genetic algorithms in high-energy physics", UMI-99-05968.

[132] G. Cowan, "Statistical data analysis", Oxford, UK: Clarendon (1998) 197 p.

[133] Christopher M. Bishop, *Neural Networks for Pattern Recognition*, Oxford University Press, Oxford, 1995.

[134] Guido Altarelli, Stefano Forte, and Giovanni Ridolfi, "On positivity of parton distributions", *Nucl. Phys.*, vol. B534, pp. 277–296, 1998.

[135] Stefano Forte, Lluis Garrido, Jose I. Latorre, and Andrea Piccione, "Neural network parametrization of deep-inelastic structure functions", *JHEP*, vol. 05, pp. 062, 2002.

[136] Luigi Del Debbio, Stefano Forte, Jose I. Latorre, Andrea Piccione, and Joan Rojo, "Unbiased determination of the proton structure function f2(p) with faithful uncertainty estimation", *JHEP*, vol. 03, pp. 080, 2005.

[137] M. R. Whalley, D. Bourilkov, and R. C. Group, "The Les Houches Accord PDFs (LHAPDF) and Lhaglue", 2005.

[138] J. Breitweg et al., "Measurement of high-Q**2 charged-current e+ p deep inelastic scattering cross sections at HERA", *Eur. Phys. J.*, vol. C12, pp. 411–428, 2000.

[139] S. Chekanov et al., "Measurement of high-Q**2 e- p neutral current cross sections at HERA and the extraction of xF3", *Eur. Phys. J.*, vol. C28, pp. 175, 2003.

[140] S. Chekanov et al., "Measurement of high-Q**2 charged current cross sections in e- p deep inelastic scattering at HERA", *Phys. Lett.*, vol. B539, pp. 197–217, 2002.

[141] S. Chekanov et al., "High-Q**2 neutral current cross sections in e+ p deep inelastic scattering at s**(1/2) = 318-GeV", *Phys. Rev.*, vol. D70, pp. 052001, 2004.

[142] S. Chekanov et al., "Measurement of high-Q**2 charged current cross sections in e+ p deep inelastic scattering at HERA", *Eur. Phys. J.*, vol. C32, pp. 1–16, 2003.

[143] S. Chekanov et al., "Measurement of charged current deep inelastic scattering cross sections with a longitudinally polarised electron beam at HERA", *Eur. Phys. J.*, vol. C61, pp. 223–235, 2009.

[144] David Alexander Mason, "Measurement of the strange - antistrange asymmetry at NLO in QCD from NuTeV dimuon data", FERMILAB-THESIS-2006-01.

[145] V. M. Abazov et al., "Measurement of the muon charge asymmetry from W boson decays", *Phys. Rev.*, vol. D77, pp. 011106, 2008.

[146] V. M. Abazov et al., "Measurement of the electron charge asymmetry in $p\bar{p} \to W + X \to e\nu + X$ events at $\sqrt{s}$ = 1.96-TeV", *Phys. Rev. Lett.*, vol. 101, pp. 211801, 2008.

[147] Darin E. Acosta et al., "Measurement of the forward-backward charge asymmetry from $W \to e\nu$ production in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV", *Phys. Rev.*, vol. D71, pp. 051104, 2005.

[148] P. M. Nadolsky, J. Huston, H. L. Lai, J. Pumplin, and C. P. Yuan, "Progress in CTEQ/TEA global QCD analysis", 2009.

[149] T. Aaltonen et al., "Measurement of the Inclusive Jet Cross Section at the Fermilab Tevatron p-pbar Collider Using a Cone-Based Jet Algorithm", *Phys. Rev.*, vol. D78, pp. 052006, 2008.

[150] Mrinal Dasgupta, Lorenzo Magnea, and Gavin P. Salam, "Non-perturbative QCD effects in jets at hadron colliders", *JHEP*, vol. 02, pp. 055, 2008.

[151] Matteo Cacciari, Juan Rojo, Gavin P. Salam, and Gregory Soyez, "Quantifying the performance of jet definitions for kinematic reconstruction at the LHC", *JHEP*, vol. 12, pp. 032, 2008.

[152] Gavin P. Salam and Gregory Soyez, "A practical Seedless Infrared-Safe Cone jet algorithm", *JHEP*, vol. 05, pp. 086, 2007.

[153] A. Accardi et al., "New parton distributions from large-x and low-Q^2 data", *Phys. Rev.*, vol. D81, pp. 034016, 2010.

[154] Jon Pumplin, "Experimental consistency in parton distribution fitting", *Phys. Rev.*, vol. D81, pp. 074010, 2010.

[155] Jon Pumplin et al., "Collider Inclusive Jet Data and the Gluon Distribution", *Phys. Rev.*, vol. D80, pp. 014019, 2009.

[156] Federico Demartin, Stefano Forte, Elisa Mariani, Juan Rojo, and Alessandro Vicini, "The impact of PDF and alphas uncertainties on Higgs Production in gluon fusion at hadron colliders", 2010.

[157] S. Alekhin, "NNLO parton distributions from deep-inelastic scattering data", 2003.

[158] J. Abate and P. Valko, "Multi-precision laplace transform inversion", *International Journal for Numerical Methods in Engineering*, vol. 60, pp. 979–993, 2003.

[159] Gavin P. Salam and Juan Rojo, "A Higher Order Perturbative Parton Evolution Toolkit (HOPPET)", *Comput. Phys. Commun.*, vol. 180, pp. 120–156, 2009.

[160] Howard Georgi and H. David Politzer, "Freedom at moderate energies: Masses in color dynamics", *Phys. Rev.*, vol. D14, pp. 1829, 1976.

[161] Tancredi Carli, Gavin P. Salam, and Frank Siegert, "A posteriori inclusion of PDFs in NLO QCD final-state calculations", 2005.

[162] T. Kluge, K. Rabbertz, and M. Wobisch, "Fast pQCD calculations for PDF fits", 2006.

[163] Tancredi Carli et al., "A posteriori inclusion of parton density functions in NLO QCD final-state calculations at hadron colliders: The APPLGRID Project", 2009.

[164] T. Gehrmann, "QCD corrections to double and single spin asymmetries in vector boson production at polarized hadron colliders", *Nucl. Phys.*, vol. B534, pp. 21–39, 1998.

[165] T. Gehrmann, "QCD corrections to the longitudinally polarized Drell-Yan process", *Nucl. Phys.*, vol. B498, pp. 245–266, 1997.

[166] Giovanni Ridolfi and Alessandro Vicini, "Private communication", 2009.

[167] G. P. Zeller et al., "A precise determination of electroweak parameters in neutrino nucleon scattering", *Phys. Rev. Lett.*, vol. 88, pp. 091802, 2002.

[168] S. Davidson, S. Forte, P. Gambino, N. Rius, and A. Strumia, "Old and new physics interpretations of the NuTeV anomaly", *JHEP*, vol. 02, pp. 037, 2002.

[169] M. Tzanov et al., "Precise measurement of neutrino and anti-neutrino differential cross sections", *Phys. Rev.*, vol. D74, pp. 012008, 2006.

[170] A. Kayis-Topaksu et al., "Leading order analysis of neutrino induced dimuon events in the CHORUS experiment", *Nucl. Phys.*, vol. B798, pp. 1–16, 2008.

[171] D. Mason et al., "Measurement of the Nucleon Strange-Antistrange Asymmetry at Next-to-Leading Order in QCD from NuTeV Dimuon Data", *Phys. Rev. Lett.*, vol. 99, pp. 192001, 2007.

[172] F. Olness et al., "Neutrino dimuon production and the strangeness asymmetry of the nucleon", *Eur. Phys. J.*, vol. C40, pp. 145–156, 2005.

[173] Stefan Kretzer et al., "The parton structure of the nucleon and precision determination of the Weinberg angle in neutrino scattering", *Phys. Rev. Lett.*, vol. 93, pp. 041802, 2004.

[174] S. Alekhin, S. Kulagin, and R. Petti, "Determination of Strange Sea Distributions from Neutrino- Nucleon Deep Inelastic Scattering", 2008.

[175] D. de Florian and R. Sassot, "Nuclear parton distributions at next to leading order", *Phys. Rev.*, vol. D69, pp. 074028, 2004.

[176] M. Hirai, S. Kumano, and T. H. Nagai, "Determination of nuclear parton distribution functions and their uncertainties at next-to-leading order", *Phys. Rev.*, vol. C76, pp. 065207, 2007.

[177] Alberto Guffanti, Juan Rojo, and Maria Ubiali, "The NNPDF1.2 parton set: implications for the LHC", 2009.

[178] T. Aaltonen et al., "First measurement of the production of a W boson in association with a single charm quark in $p\bar{p}$ collisions at $\sqrt{s}$ = 1.96-TeV", *Phys. Rev. Lett.*, vol. 100, pp. 091803, 2008.

[179] *MCFM*.

[180] John M. Campbell and R. Keith Ellis, "Radiative corrections to Z b anti-b production", *Phys. Rev.*, vol. D62, pp. 114012, 2000.

[181] R. S. Thorne and W. K. Tung, "PQCD Formulations with Heavy Quark Masses and Global Analysis", 2008.

[182] Henning Flacher et al., "Gfitter - Revisiting the Global Electroweak Fit of the Standard Model and Beyond", *Eur. Phys. J.*, vol. C60, pp. 543–583, 2009.

[183] S. Heinemeyer, W. Hollik, and G. Weiglein, "Electroweak precision observables in the minimal supersymmetric standard model", *Phys. Rept.*, vol. 425, pp. 265–368, 2006.

[184] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, "Parton distributions incorporating QED contributions", *Eur. Phys. J.*, vol. C39, pp. 155–161, 2005.

[185] Janet M. Conrad, Michael H. Shaevitz, and Tim Bolton, "Precision measurements with high-energy neutrino beams", *Rev. Mod. Phys.*, vol. 70, pp. 1341–1392, 1998.

[186] Michelangelo L. Mangano et al., "Physics at the front-end of a neutrino factory: A quantitative appraisal", 2001.

[187] H. Abramowicz et al., "Experimental Study of Opposite Sign Dimuons Produced in Neutrino and anti-neutrinos Interactions", *Z. Phys.*, vol. C15, pp. 19, 1982.

[188] Tim Bolton, "Determining the CKM parameter V(cd) from nu N charm production", 1997.

[189] Claude Amsler et al., "Review of particle physics", *Phys. Lett.*, vol. B667, pp. 1, 2008.

[190] Donald E. Groom et al., "Review of particle physics", *Eur. Phys. J.*, vol. C15, pp. 1–878, 2000.

[191] *UTFit*.

[192] *CKMfitter*.

[193] Andreas Hocker, H. Lacker, S. Laplace, and F. Le Diberder, "A New approach to a global fit of the CKM matrix", *Eur. Phys. J.*, vol. C21, pp. 225–259, 2001.

[194] P. Abreu et al., "Measurement of |V(cs)| using W decays at LEP2", *Phys. Lett.*, vol. B439, pp. 209–224, 1998.

[195] A. Ealet, "W W cross sections and |V(cs)| measurement", *Nucl. Phys. Proc. Suppl.*, vol. 115, pp. 249–254, 2003.

[196] W. James Stirling, "Progress in Parton Distribution Functions", 2008.

[197] Matteo Cacciari, Stefano Frixione, Michelangelo L. Mangano, Paolo Nason, and Giovanni Ridolfi, "Updated predictions for the total production cross sections of top and of heavier quark pairs at the Tevatron and at the LHC", *JHEP*, vol. 09, pp. 127, 2008.

[198] Nikolaos Kidonakis and Ramona Vogt, "The Theoretical top quark cross section at the Tevatron and the LHC", *Phys. Rev.*, vol. D78, pp. 074005, 2008.

[199] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, "Uncertainties on alpha$_S$ in global PDF analyses and implications for predicted hadronic cross sections", *Eur. Phys. J.*, vol. C64, pp. 653–680, 2009.

[200] Hung-Liang Lai et al., "Uncertainty induced by QCD coupling in the CTEQ-TEA global analysis of parton distributions", 2010.

[201] R. S. Thorne, A. D. Martin, W. J. Stirling, and G. Watt, "The effects of combined HERA and recent Tevatron W - > lepton neutrino charge asymmetry data on the MSTW PDFs", 2010.

[202] Tancredi Carli et al., "A posteriori inclusion of parton density functions in NLO QCD final-state calculations at hadron colliders: The APPLGRID Project", *Eur. Phys. J.*, vol. C66, pp. 503–524, 2010.

[203] Mika Vesterinen, "A measurement of the muon charge asymmetry in W boson events", 2010.

[204] Stefano Catani, Leandro Cieri, Giancarlo Ferrera, Daniel de Florian, and Massimiliano Grazzini, "Vector boson production at hadron colliders: a fully exclusive QCD calculation at NNLO", *Phys. Rev. Lett.*, vol. 103, pp. 082001, 2009.

[205] M. Buza, Y. Matiounine, J. Smith, and W. L. van Neerven, "An ordered analysis of heavy flavour production in deep inelastic scattering", *Eur. Phys. J.*, vol. C1, pp. 301, 1998.

[206] M. Gluck and E. Reya, "Duality Predictions for the Production of Heavy Quark Systems in QCD", *Phys. Lett.*, vol. B79, pp. 453, 1978.

[207] J. Smith and B. W. Harris, "Heavy-quark correlations in deep inelastic scattering", *Nucl. Phys. Proc. Suppl.*, vol. 51C, pp. 188–194, 1996.

[208] R. Vogt and S. J. Brodsky, "QCD and intrinsic heavy quark predictions for leading charm and beauty hadroproduction", *Nucl. Phys.*, vol. B438, pp. 261–277, 1995.

[209] W. L van Neerven, "Production of heavy quarks in deep inelastic lepton hadron scattering", 2001.

[210] J. Breitweg et al., "Measurement of open beauty production in photoproduction at HERA", *Eur. Phys. J.*, vol. C18, pp. 625–637, 2001.

[211] A. Chuvakin, J. Smith, and B. W. Harris, "Variable flavor number schemes versus fixed order perturbation theory for charm quark electroproduction", *Eur. Phys. J.*, vol. C18, pp. 547–553, 2001.

[212] Risto Raitio and Walter W. Wada, "HIGGS BOSON PRODUCTION AT LARGE TRANSVERSE MOMENTUM IN QCD", *Phys. Rev.*, vol. D19, pp. 941, 1979.

[213] Michael Kramer, 1, "Associated Higgs production with bottom quarks at hadron colliders", *Nucl. Phys. Proc. Suppl.*, vol. 135, pp. 66–70, 2004.

[214] R. Michael Barnett, Howard E. Haber, and Davison E. Soper, "Ultraheavy Particle Production from Heavy Partons at Hadron Colliders", *Nucl. Phys.*, vol. B306, pp. 697, 1988.

[215] Duane A. Dicus and Scott Willenbrock, "Higgs Boson Production from Heavy Quark Fusion", *Phys. Rev.*, vol. D39, pp. 751, 1989.

[216] David L. Rainwater, Michael Spira, and Dieter Zeppenfeld, "Higgs boson production at hadron colliders: Signal and background processes", 2002.

[217] F. Maltoni, Z. Sullivan, and S. Willenbrock, "Higgs-boson production via bottom-quark fusion", *Phys. Rev.*, vol. D67, pp. 093005, 2003.

[218] Eduard Boos and Tilman Plehn, "Higgs-boson production induced by bottom quarks", *Phys. Rev.*, vol. D69, pp. 094005, 2004.

[219] Stefan Dittmaier, Michael Kramer, Michael Spira, and Manuel Walser, "Charged-Higgs-boson production at the LHC: NLO supersymmetric QCD corrections", 2009.

[220] John M. Campbell and R. Keith Ellis, "Radiative corrections to Z b anti-b production", *Phys. Rev.*, vol. D62, pp. 114012, 2000.

[221] F. Febres Cordero, L. Reina, and D. Wackeroth, "NLO QCD corrections to $Zb\bar{b}$ production with massive bottom quarks at the Fermilab Tevatron", *Phys. Rev.*, vol. D78, pp. 074014, 2008.

[222] F. Febres Cordero, L. Reina, and D. Wackeroth, "W- and Z-boson production with a massive bottom-quark pair at the Large Hadron Collider", *Phys. Rev.*, vol. D80, pp. 034015, 2009.

[223] John M. Campbell, R. Keith Ellis, F. Maltoni, and S. Willenbrock, "Associated production of a Z Boson and a single heavy quark jet", *Phys. Rev.*, vol. D69, pp. 074021, 2004.

[224] John M. Campbell, R. Keith Ellis, F. Maltoni, and S. Willenbrock, "Production of a Z boson and two jets with one heavy- quark tag", *Phys. Rev.*, vol. D73, pp. 054007, 2006.

[225] John M. Campbell et al., "Associated Production of a W Boson and One b Jet", *Phys. Rev.*, vol. D79, pp. 034023, 2009.

[226] John M. Campbell, R. Keith Ellis, F. Maltoni, and S. Willenbrock, "Production of a W boson and two jets with one b^- quark tag", *Phys. Rev.*, vol. D75, pp. 054015, 2007.

[227] Walter T. Giele, Stephane Keller, and Eric Laenen, "QCD corrections to W boson plus heavy quark production at the Tevatron", *Phys. Lett.*, vol. B372, pp. 141–149, 1996.

[228] John M. Campbell, R. Keith Ellis, and Francesco Tramontano, "Single top production and decay at next-to-leading order", *Phys. Rev.*, vol. D70, pp. 094012, 2004.

[229] J. Baines et al., "Heavy quarks (Working Group 3): Summary report", 2006.

[230] John M. Campbell, Rikkert Frederix, Fabio Maltoni, and Francesco Tramontano, "NLO predictions for t-channel production of single top and fourth generation quarks at hadron colliders", *JHEP*, vol. 10, pp. 042, 2009.

[231] John M. Campbell, Rikkert Frederix, Fabio Maltoni, and Francesco Tramontano, "Next-to-Leading-Order Predictions for t-Channel Single-Top Production at Hadron Colliders", *Phys. Rev. Lett.*, vol. 102, pp. 182003, 2009.

[232] Fredrick I. Olness and Wu-Ki Tung, "When Is a Heavy Quark Not a Parton? Charged Higgs Production and Heavy Quark Mass Effects in the QCD Based Parton Model", *Nucl. Phys.*, vol. B308, pp. 813, 1988.

[233] Stefano Forte and Giovanni Ridolfi, "Renormalization group approach to soft gluon resummation", *Nucl. Phys.*, vol. B650, pp. 229–270, 2003.

[234] Johan Alwall et al., "MadGraph/MadEvent v4: The New Web Generation", *JHEP*, vol. 09, pp. 028, 2007.