A Machine Learning Analysis of Polarised Proton Substructure

A. J. Hasenack

A thesis submitted for the degree of
Master of Science

Supervisors:
Prof. Dr. J. Rojo
Dr. T.R. Rabemananjara
MSc. G. Magni

Examinors:
Prof. Dr. E.L.M.P. Laenen
Dr. U. Gürsoy

Date: Juli, 2024

Universiteit
Utrecht

Summary

We present the first NNPDFpol2.0 results, where we fit polarised parton distribution functions to deep inelastic scattering, W-production in Drell-Yan and (di)jet data including charm contributions at next-to-next leading order. The effect on singlet and non-singlet distributions turns out to be small, however the gluon pPDF is quite significantly altered. We continue by considering EIC pseudo-data and its implications on the distributions. Next to this, we attempt to lay a strong theoretical foundation of factorisation by using the framework of soft-collinear effective theory.

Acknowledgements

The author wishes to thank Prof. Rojo for the opportunity to work in the group, Dr. Rabemananjara and MSc. Magni, without whom none of the results would have been possible, and MSc. Jaarsma for the instructive discussions on effective field theory.

Contents

1 Introduction 3

2 Theory 6
- 2.1 Factorisation, Structure Functions and Interpretation . . . 6
  - 2.1.1 Hadronic Tensor . . . 6
  - 2.1.2 Soft Collinear Effective Theory . . . 6
- 2.2 Renormalisation and Evolution . . . 12
  - 2.2.1 Infrared Divergence and The DGLAP Equations . . . 12
- 2.3 Massive Corrections . . . 14

3 Methodology 16
- 3.1 NNPDF Architecture . . . 16
- 3.2 PDF Basis and Theoretical Constraints . . . 17
  - 3.2.1 Sum Rules . . . 17
  - 3.2.2 Integrability . . . 18
  - 3.2.3 Positivity . . . 18
- 3.3 Polarised Jets Code . . . 18
- 3.4 FK-tables . . . 19
- 3.5 Reweighting and Unweighting . . . 19
- 3.6 Data . . . 20
  - 3.6.1 DIS Data . . . 20
  - 3.6.2 Jet Data . . . 21
  - 3.6.3 W-data . . . 22
  - 3.6.4 EIC Pseudo-data . . . 22

4 Results 23
- 4.1 Reweighting . . . 23
- 4.2 Polarised jets . . . 27
- 4.3 pPDFs . . . 29
- 4.4 EIC Projection . . . 29

5 Conclusion 31

1 Introduction

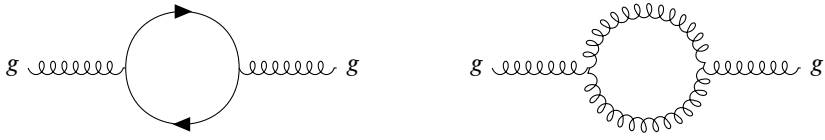
Quantum Chromodynamics (QCD) is the study of the strong force, responsible for the binding of the nuclei inside atoms. Just like QED, QCD has a field theoretic description (to be precise, a *Yang-Mills theory*), making use of *gluons* (a type of *boson*) to describe the force. We can model the strength of this force (*strong coupling*) for various energetic regimes. To lowest order, the vacuum polarisation contributes to its value. QCD – being a non-abelian theory – admits gluon loops, as can be seen in figure 1. We can regularise and renormalise the strong coupling to account for the divergences. However, because of the additional gluon loop contribution, we have two opposing effects: *screening* because the quark loop[1] and *anti-screening* because of the gluon loop. This effect can well be observed in the *running* of the strong coupling constant

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

with

$$\beta_0(Q^2) = \frac{11N_c - 2n_f(Q^2)}{12\pi} \tag{2}$$

where $11N_c$ comes from the gluon loop and scales with the number of colours (3) and $2n_f$ comes from the quark loop and scales with the number of flavours with mass smaller than Q. In the limit that $Q^2 \to \infty$, we see that $\beta_0 > 0$ and thus $\alpha_s(Q^2) \to 0$, a result which is known as *asymptotic freedom*. On the other hand, when the energy decreases, coupling increases and it becomes very expensive to split a closely bound group of quarks. The theory becomes non-perturbative in this regime, whence we must look for alternative descriptions [1] [2].



(a) quark loop correction (b) gluon loop correction

Figure 1: First order loop corrections to vacuum polarisation.

Analogously to QED, there is a charge related to the strong force, which is called *colour*. There are three such colours (typically red, green and blue) and it is a postulate of QCD that they are never observed individually, but in colourless combinations (a result called *confinement*). One of the allowed configurations, is the proton, which consists of three valence quarks (u, u, d) that have distinct colours. The sum of the charges of the quarks add up to $2 \times 2/3 - 1/3 = 1$, precisely the proton charge, and naively one would expect that the same would hold for the spin. However, it has been known for a while that this is not the case and the quarks only add up to about 34%

[1] Comparable to the charge screening in QED.

of the proton spin, $\hbar/2$ [3]. The question arises what other factors could contribute to the total. We can keep track of the contributions to the proton spin by introducing the Jaffe-Manohar decomposition

$$\frac{1}{2} = L_q + L_g + \frac{1}{2}\Delta\Sigma + \Delta G \tag{3}$$

where $L_q, L_g, \Delta\Sigma, \Delta G$ are the quark and gluon angular momenta and the quark and gluon spin respectively (the exact interpretation will follow in 2.1). It is already important to note here that the decomposition is not manifestly gauge invariant [4]. This also caused the search for different decompositions such as in [5], however, we will not further go into detail here. The bottom line is that the gluon contribution is of essential importance in describing the total proton momentum.

One of the processes that can help in pinpointing the quark and gluon spins, is called *Deep Inelastic Scattering* (DIS). DIS is a high-energy scattering process in which leptons are collided with multi-quark targets (hadrons), such that the initial bound state is no longer conserved. The shattered insides form new pairs or triplets in a stage which is called the *hadronisation stage*. The products that are formed in this latter phase are often denoted by X, reflecting our negligence to its further specification. In figure 2 such a process can be seen. The electron couples to one of the quarks inside the proton, carrying momentum fraction x and spin, via the photon with momentum q. The value of $q^2 = -Q^2$ is called the *virtuality* of the photon. Next to the possibility

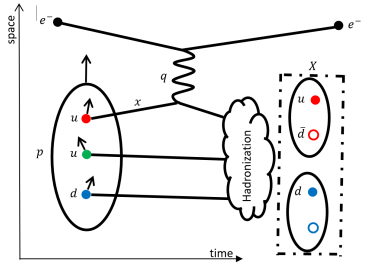


Figure 2: Schemetic illustration of DIS process.

of coupling to one of three valence quarks, the photon can also interact with one of the *sea quarks*, like a strange or even charm quark and also with the gluons binding the proton. Hence, measuring the DIS process can give valuable insight in the (probabilistic) composition of the proton. Using the framework of QED and QCD, we can model the cross-section for this process

$$d\sigma \propto L^{\mu\nu} W_{\mu\nu} \tag{4}$$

Here, $L^{\mu\nu}$ and $W^{\mu\nu}$ are respectively the *leptonic* and *hadronic tensors*. In our efforts, the leptonic tensor will not be further specified, since it is fully perturbative in terms of the photon coupling and can be calculated term by term. For an overview we refer to [6]. The hadronic tensor however, is a more complicated matter and will turn out to depend on the quantities

$$F(x, Q^2) = \sum_b H_b(x, Q^2) \otimes f_{b/B}(x, Q^2) \tag{5}$$

where $f_{b/B}(x, Q^2)$ can be seen as the probability density to have the photon interacting with proton constituent b (quark or gluon) with momentum fraction x at photon energy scale Q^2. These density distributions are non-perturbative and cannot be analytically derived. On the other hand, $H_b(x, Q^2)$ is a perturbative amplitude of a photon interacting with a PDF. Separation of these non-perturbative – perturbative parts is referred to as *factorisation* and is a non-trivial fact that we will treat in detail in 2.1. To be more specific the density functions, $f_{b/B}(x, Q^2)$, come in variants that describe the number density, typically denoted by $u, \bar{u}, d$ etc., and spin densities $\Delta u, \Delta\bar{u}, \Delta d \ldots$, respectively referred to as *(polarised) parton distribution functions*. The calculation of the latter is the main scope of this work, as they take part in describing the missing spin contributions in equation 3. To this end, our formalism is based on the highly accurate NNPDF framework, of which at the current time the latest version is NNPDF4.0 [7] [8], and the approach of NNPDFpol1.0 [9].

To be able to probe the constituents of the proton to an even more precise extent, a novel collider will be built, called the *Electron Ion Collider* (EIC) [10]. The scope of the detector will partially be to narrow down quite significantly on the gluon spin, as can be seen in figure 3. Hence, for the NNPDF collaboration, it is useful to both make predictions for the EIC as well as study the impact on the PDFs. This will be part of the results section 4.

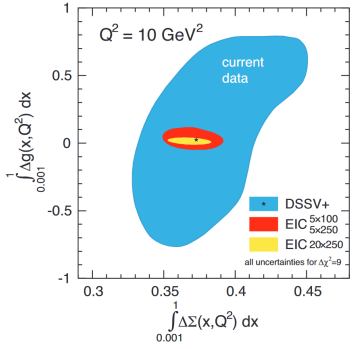


Figure 3: Projected reduction of the quark and gluon spin after conducting the EIC experiments. Image taken from the EIC whitepaper [10].

2 Theory

2.1 Factorisation, Structure Functions and Interpretation

In the following, we sketch a concise overview of how to obtain the DIS structure functions, F_1, F_2 and g_1 from first principles. In order to do so, we work in the SCET framework, in which we will relate field operators with PDFs.

2.1.1 Hadronic Tensor

The coupling of the photon field, A_μ, to the electromagnetic current field, J^μ, can be modelled by the interaction term

$$J^\mu(x)\, A_\mu(x) \tag{6}$$

In our case, the current is that of a quark and can be expressed as

$$J^\mu(x) = e_q \overline{\psi}(x)\gamma^\mu\psi(x) \tag{7}$$

where $\psi(x)$ are the quark spinor fields, γ^μ the usual gamma matrices and e_q the quark charge. We define the hadronic tensor as the squared expectation value of the (real) current flowing from the initial bound state, $|B(p)\rangle$, to the unspecified final product, $|X(p_X)\rangle$, under the assumption of an emitted photon with momentum q and summed over all the final product configurations

$$\begin{aligned} W^{\mu\nu} = \sum_{X,p_X} &\langle B(p)|J^\mu(0)|X(p_X)\rangle\times \\ &\langle X(p_X)|J^\nu(0)|B(p)\rangle\, \delta\left(p_X - p - q\right) \end{aligned} \tag{8}$$

After some rewriting and using the completeness of the sum over the final states, this is equivalent to

$$W^{\mu\nu}(p,q) = \int d^4y\, e^{iqy} \left\langle B(p)\right| \left[J^\mu(y), J^\nu(0)\right]_- \left| B(p)\right\rangle \tag{9}$$

[2]. In literature, the integration over the final product states $|X(p_X)\rangle$ is referred to as *fully inclusive DIS*. In figure 4, the squared amplitude can be seen for the DIS process, where the hadronic tensor we just defined (bottom part of diagram) couples to the leptonic tensor ($\mathcal{M}^2 \propto L^{\mu\nu}W_{\mu\nu}$). It is interesting to note that due to the separation of the leptonic and hadronic tensor, different multiple interactions between the electron and the bound state partons are essentially ignored, which is an inherent property of the framework.

2.1.2 Soft Collinear Effective Theory

In order to be able to find an effective expression for the hadronic tensor, one needs to apply a technique which is called *separation of scales*. This implies that some of the cross-sectional behaviour is captured by higher energy regimes (the *hard* part) and some in lower energy scales (the *soft* and *collinear* parts) and the full description is the product of the two. There are various ways to perform

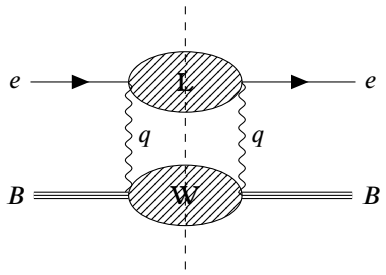


Figure 4: Feynman diagram for the squared amplitude of the DIS process where a photon emitted by an electron couples to the bound state B.

this separation of scales. In more traditional QCD literature we typically encounter one which is called the *operator product expansion* (*OPE*). However, OPE is not applicable in all processes. Thus, in recent years it has become increasingly popular to perform the relevant separation already at Lagrangian level with the so called *Soft Collinear Effective Theory* (*SCET*). In the following, we provide a concise overview of SCET, applied to deep inelastic scattering.

First, we define the light-cone coordinates in terms of lab-frame coordinates as

$$x^+ = n^\mu x_\mu, x^- = \overline{n}^\mu x_\mu, \vec{v} = (x_1, x_2) \tag{10}$$

with $\boldsymbol{n} = (1, 0, 0, 1), \overline{\boldsymbol{n}} = (1, 0, 0, -1)$ the *collinear* and *anti-collinear* directions respectively. Now, the idea behind SCET, is that, depending on their direction in terms of light-cone coordinates, various quantities scale with different powers of a small quantity, λ. In our case, λ is defined as the fraction between the scale above which perturbative QCD is applicable and the collision energy scale,

$$\lambda = \Lambda_{QCD}/Q \tag{11}$$

To illustrate, the bound state momentum behaves collinearly

$$\boldsymbol{p} \sim \boldsymbol{p}_c = (\lambda^2, 1, \lambda)Q \tag{12}$$

while the lepton momentum behaves anti-collinearly

$$\boldsymbol{p}_\ell \sim \boldsymbol{p}_{\overline{c}} = (1, \lambda^2, \lambda)Q \tag{13}$$

The energetically most suppressed type, is called *ultrasoft*, and scales like

$$\boldsymbol{p}_s = (\lambda^2, \lambda^2, \lambda^2)Q \tag{14}$$

and the most energetic is the hard mode

$$\boldsymbol{p}_h = (1, 1, 1)Q \tag{15}$$

We proceed by making some modifications to the regular theory, by which the less interesting modes (ultrasoft and hard) are separated from the quantities we would like to understand

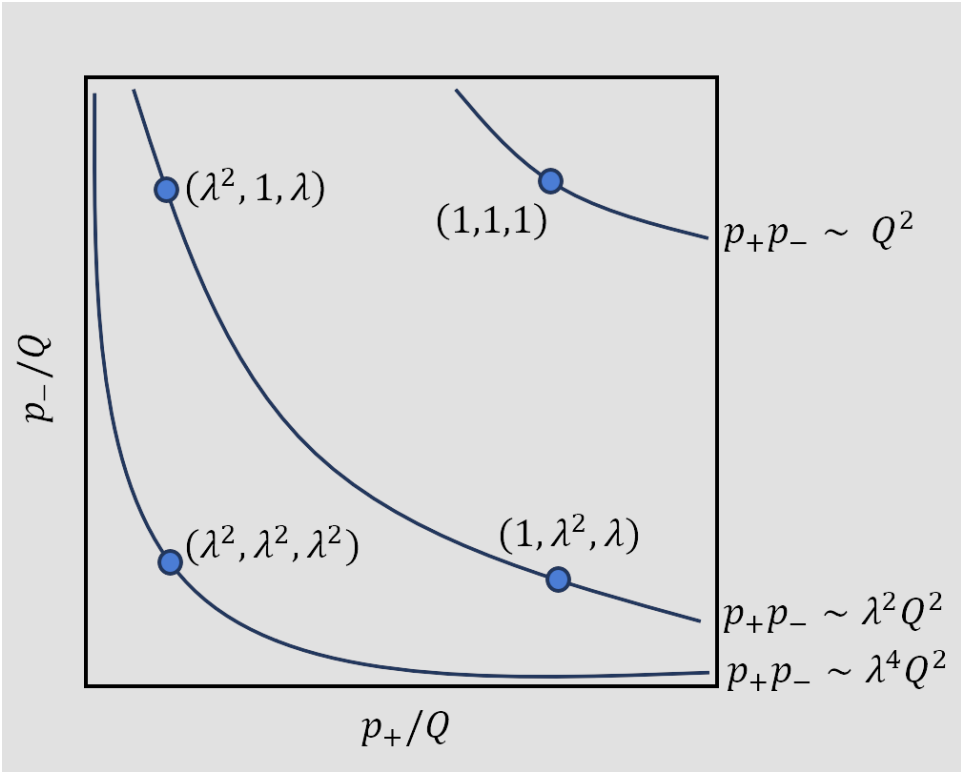


Figure 5: Overview of the different scales in SCETI factorisation.

(mostly collinear). The primary postulate – of which the type is quite common for effective field theorists – is that the quark and gluon fields can be written as a direct sum of the different regimes

$$\psi = \psi_c + \psi_{\bar{c}} + \psi_s \tag{16}$$

and

$$A^\mu = A_c^\mu + A_{\bar{c}}^\mu + A_s^\mu \tag{17}$$

The hard modes, on the other hand, will be integrated out and thus do not play a role in the underlying field theoretic description. The collinear field, ψ_c, is expected to have the most prominent contribution to the total. However, scaling of the quark fields also depends on their handedness. It can be shown that for the projectors $P_\pm = \gamma_\pm\gamma_\mp$[2] the collinear/anti-collinear spinors scale as

$$\xi_{c/\bar{c}} \equiv P_+\psi_{c/\bar{c}} \sim \lambda \tag{18}$$

[2]Here, $\gamma_+ = n^\mu\gamma_\nu$ and $\gamma_- = \bar{n}^\mu\gamma_\nu$.

whereas

$$\eta_{c/\bar{c}} \equiv P_- \psi_{c/\bar{c}} \sim \lambda^2 \tag{19}$$

Note here, that we have already excluded the hard contributions from the fields. These will later return in the form of coefficients (so-called *Wilson coefficients*) that can perturbatively calculated outside of the SCET framework. Next, we return to the hadronic tensor in 9 and we consider now only the operator responsible for its expectation value

$$W^{\mu\nu} = \int d^4y \, e^{iq\cdot y} \, \left[J^\mu(y), J^\nu(0) \right]_- \tag{20}$$

We now wish to find an expression in terms of gauge-invariant collinear building blocks that represents the same tensor. These covariant building blocks can then be used to construct operators in the SCET framework, which then form the total tensor. For this, we first transform the spinor field into a covariant version by invoking collinear *Wilson line*, W. This is defined as

$$W(x) = \exp\left[ig \int_{\mathscr{C}} dy_\mu \, A_c^\mu \, (x + y) \right] \tag{21}$$

with $\mathscr{C} = \{ u\bar{n} \mid u \in (-\infty, 0] \}$.

Intuition. Wilson lines can be understood both from a physical and mathematical perspective. Physically they correspond to a measure of the possible numbers of photons emitted over a line, weighed by some power of the coupling constant. This can be readily seen from the Taylor series in terms of the gluon field ([11]). The question arises why a straight line is considered. This has to do with the fact that due to the massive nature of the collinear particles, their paths can be taken to be approximately straight. This is called the *no-recoil approximation*. To study the Wilson lines mathematically on the other hand, we turn to the principal-bundle description, $\mathscr{B} = (\mathscr{M}, \mathscr{G})$, of the gauge-theory, where the base space, $\mathscr{M}$, is physical space and the fibers, $\mathscr{G}$, are copies of the $SU(3)$ group. A Wilson line, then corresponds to the solution of the parrallel-transport of the identity element on the line $\mathscr{C}$ with respect to the gauge covariant derivative. We see that if $W \in \mathscr{G}$ is such that

$$dW(x) = (\partial^\mu - igA^\mu(x)) \, W(x) \, dx_\mu = 0 \Leftrightarrow W(x) = \exp\left[ig \int_{\mathscr{C}} dy_\mu \, A^\mu \, (x) \right] \tag{22}$$

Using this definition, we claim that the replacements

$$\xi_c \to \chi \equiv W^\dagger \, \xi_c \tag{23}$$

and

$$A_c^\mu \to \mathscr{A}^\mu \equiv W^\dagger i D_c^\mu W \tag{24}$$

with $D_c = i\partial - gA_c$, the transform covariantly with respect to collinear gauge transformations. From the redefinition of the new gluon field, we also define the field strength tensor

$$ig \, \mathscr{G}^{\mu\nu} = [i\mathscr{D}^\mu, i\mathscr{D}^\nu] \tag{25}$$

with $\mathscr{D}^\mu = i\partial^\mu - g\mathscr{A}^\mu$. Moreover, we have that

$$\begin{cases} \mathscr{G}_\perp^\lambda \mathscr{G}_{\perp,\lambda} \\ \mathscr{G}^\lambda \tilde{\mathscr{G}}_\lambda \end{cases} \tag{26}$$

are the only collinear kinetic terms for the gluon which are not power supressed[3] [12], where

$$\begin{cases} \mathscr{G}^\lambda = \overline{n}_\rho \mathscr{G}^{\rho\lambda} \\ ig\,\mathscr{G}_\perp^\lambda = \overline{n}_\rho [i\mathscr{D}^\rho, i\mathscr{D}_\perp^\lambda] \\ ig\,\tilde{\mathscr{G}}^\lambda = \frac{1}{2}\epsilon^{\lambda\rho\mu\nu}\overline{n}_\rho [i\mathscr{D}_\mu, i\mathscr{D}_\nu] \end{cases} \tag{27}$$

Using the previous fact, the most general, non-local SCET quark and gluon operators in momentum space, can be written as

$$\begin{cases} J_q^j(\boldsymbol{k}) = \sum\limits_x \int_x^1 d\xi\, C_q^j(\frac{x}{\xi})\, \overline{\chi}_q(\boldsymbol{k})\, \overline{\boldsymbol{n}}\, \chi_q(\frac{\xi}{x} Q\boldsymbol{n} - \boldsymbol{k}) \\ K_q^j(\boldsymbol{k}) = \sum\limits_x \int_x^1 d\xi\, \tilde{C}_q^j(\frac{x}{\xi})\, \overline{\chi}_q(\boldsymbol{k})\, \overline{\boldsymbol{n}}\gamma_5\, \chi_q(\frac{\xi}{x} Q\boldsymbol{n} - \boldsymbol{k}) \\ L_g^j(\boldsymbol{k}) = \sum\limits_x \int_x^1 d\xi\, C_g^j(\frac{x}{\xi})\, \mathscr{G}_\perp^\lambda(\boldsymbol{k}) \mathscr{G}_{\perp,\lambda}(\frac{\xi}{x} Q\boldsymbol{n} - \boldsymbol{k}) \\ M_g^j(\boldsymbol{k}) = \sum\limits_x \int_x^1 d\xi\, \tilde{C}_g^j(\frac{x}{\xi})\, \mathscr{G}^\lambda(\boldsymbol{k}) \tilde{\mathscr{G}}_\lambda(\frac{\xi}{x} Q\boldsymbol{n} - \boldsymbol{k}) \end{cases} \tag{28}$$

where q is the quark type, $x \equiv \frac{Q^2}{2\boldsymbol{p}\cdot\boldsymbol{q}} = \frac{Q}{\overline{\boldsymbol{n}}\cdot\boldsymbol{p}}$ is called the *Bjorken scale* and the C's, are the Wilson coefficients. Note that the integration bounds are such that the spinor and vector momentum is between that of the photon and of the proton. A useful notational trick is the *Mellin product*

$$(\alpha \otimes \beta)(x) = \int_x^1 d\xi\, \alpha\left(\frac{x}{\xi}\right) \beta(\xi) \tag{29}$$

Then, the operator matrix elements can be obtained

$$\begin{cases} \langle B(\boldsymbol{p})|J_q^j(\boldsymbol{k})|B(\boldsymbol{p})\rangle = \sum\limits_x C_q^j \otimes f_{q/B} \\ \langle B(\boldsymbol{p})|K_q^j(\boldsymbol{k})|B(\boldsymbol{p})\rangle = \sum\limits_x \tilde{C}_q^j \otimes \tilde{f}_{q/B} \\ \langle B(\boldsymbol{p})|L_g^j(\boldsymbol{k})|B(\boldsymbol{p})\rangle = \sum\limits_x C_g^j \otimes f_{g/B} \\ \langle B(\boldsymbol{p})|M_g^j(\boldsymbol{k})|B(\boldsymbol{p})\rangle = \sum\limits_x \tilde{C}_g^j \otimes \tilde{f}_{g/B} \end{cases} \tag{30}$$

where we identify the so-called *parton distribution functions (PDFs)* with

$$\begin{cases} f_{q/B}(\xi) \equiv \int ds\, e^{-2i\overline{\boldsymbol{n}}\cdot\boldsymbol{p}\xi s} \langle B(\boldsymbol{p})|\overline{\chi}_q(s\boldsymbol{n})\, \overline{\boldsymbol{n}}\, \chi_q(s\boldsymbol{n})|B(\boldsymbol{p})\rangle \\ \tilde{f}_{q/B}(\xi) \equiv \int ds\, e^{-2i\overline{\boldsymbol{n}}\cdot\boldsymbol{p}\xi s} \langle B(\boldsymbol{p})|\overline{\chi}_q(s\boldsymbol{n})\, \overline{\boldsymbol{n}}\gamma^5\, \chi_q(s\boldsymbol{n})|B(\boldsymbol{p})\rangle \\ g(\xi) \equiv f_{g/B}(\xi) \equiv \int ds\, e^{-2i\overline{\boldsymbol{n}}\cdot\boldsymbol{p}\xi s} \langle B(\boldsymbol{p})|\mathscr{G}_\perp^\lambda(s\boldsymbol{n}) \mathscr{G}_{\perp,\lambda}(s\boldsymbol{n})|B(\boldsymbol{p})\rangle \\ \Delta g(\xi) \equiv \tilde{f}_{g/B}(\xi) \equiv \int ds\, e^{-2i\overline{\boldsymbol{n}}\cdot\boldsymbol{p}\xi s} \langle B(\boldsymbol{p})|\mathscr{G}^\lambda(s\boldsymbol{n}) \tilde{\mathscr{G}}_\lambda(s\boldsymbol{n})|B(\boldsymbol{p})\rangle \end{cases} \tag{31}$$

[3] Here, the perpendicular parts in the first term come from the presumed light-cone gauge ($A^+ = 0$).

Because we are still considering the light-cone coordinates, the hadronic tensor will not directly depend on the quark PDFs but rather on symmetric and anti-symmetric combinations with its charge-conjugated counterpart. The lab-frame quark number density is given by

$$\begin{aligned} q(\xi) =& f_{q/B}(\xi) + f^C_{q/B}(\xi) \\ =& \int ds\, e^{-2i\bar{\boldsymbol{n}}\cdot p\,\xi s} \langle B(\boldsymbol{p})|\chi^\dagger_q(s\boldsymbol{n})\chi_q(s\boldsymbol{n})|B(\boldsymbol{p})\rangle \end{aligned} \tag{32}$$

while the z-directed spin density is given by[4]

$$\begin{aligned} \Delta q(\xi) =& \tilde{f}_{q/B}(\xi) - \tilde{f}^C_{q/B}(\xi) \\ =& \int ds\, e^{-2i\bar{\boldsymbol{n}}\cdot p\,\xi s} \langle B(\boldsymbol{p})|\chi^\dagger_q(s\boldsymbol{n})(I\otimes\sigma_z)\chi_q(s\boldsymbol{n})|B(\boldsymbol{p})\rangle \end{aligned} \tag{33}$$

The gluon spin term, Δg, is written in a quite complicated fashion. However, in the case that $A^+ = 0$, the total gluon spin reduces to the z-component of the gluon spin that we are familiar with [13]

$$\Delta G = \int_x^1 d\xi\, \Delta g(\xi) = \int ds \left(\vec{E}\times\vec{B}\right)^3 \tag{34}$$

Next, we turn to the calculation of the hadronic tensor in terms of the previously found PDFs. By invoking charge conservation ($(n_\mu - \bar{n}_\mu)W^{\mu\nu} = 0$) and parity conservation, the general form of a SCET rank-two tensor up to twist-2 that is both dependent on the current and proton spin can be written as

$$\begin{aligned} W^{\mu\nu}_{SCET}(n,s) = \frac{1}{x}\Bigg[& \eta^{\mu\nu}_\perp \left(\sum_q e_q^2 C_q^1 \otimes q + \frac{C_g^1}{Q}\otimes g\right) \\ & + (n^\mu + \bar{n}^\mu)(n^\nu + \bar{n}^\nu)\left(\sum_q e_q^2 C_q^2 \otimes q + \frac{C_g^2}{Q}\otimes g\right) \\ & + i\epsilon^{\mu\nu\rho\sigma}(n_\rho - \bar{n}_\rho)s_\sigma\left(\sum_q e_q^2 \tilde{C}_q^1 \otimes \Delta q + \frac{\tilde{C}_g^1}{Q}\otimes\Delta g\right)\Bigg] \end{aligned} \tag{35}$$

Similarly, the form of the QCD hadronic tensor is equal to [9] [14]

$$\begin{aligned} W^{\mu\nu}_{QCD}(q,s) =& 2\left(-\eta^{\mu\nu} + \frac{q^\mu q^\nu}{q^2}\right)F_1(x,Q^2) \\ & + \frac{2}{p\cdot q}\left(p^\mu + \frac{q^\mu}{2x}\right)\left(p^\nu + \frac{q^\nu}{2x}\right)F_2(x,Q^2) \\ & + \frac{2Q}{p\cdot q} i\epsilon^{\mu\nu\rho\sigma} q_\rho s_\sigma g_1(x,Q^2) \end{aligned} \tag{36}$$

[4] Where $\otimes$ refers to the tensor product here.

Here, F_1, F_2, g_1 are called *structure functions*[5]. It must be noted that there are more structure functions but that here, $g_2 = F_4 = F_5 = 0$ by mass suppression and $F_3 = g_3 = g_4 = g_5 = 0$ by restricting ourselves to the electromagnetic force[6].By making the comparison between the SCET and QCD hadronic tensor, an expression of the structure functions in terms of the parton distribution functions can be obtained

$$\begin{cases} F_1(x,Q^2) = \sum_q e_q^2 \, H_q^1 \otimes q + \frac{H_g^1}{Q} \otimes g = -\frac{1}{2x}\left(\sum_q e_q^2 \, C_q^1 \otimes q + \frac{C_g^1}{Q} \otimes g\right) \\ F_2(x,Q^2) = \sum_q e_q^2 \, H_q^2 \otimes q + \frac{H_g^2}{Q} \otimes g \\ \qquad\qquad = 2\left(4\sum_q e_q^2 \, C_q^2 \otimes q + 4\frac{C_g^2}{Q} \otimes g - \sum_q e_q^2 \, C_q^1 \otimes q - \frac{C_g^1}{Q} \otimes g\right) \\ g_1 = \frac{1}{2}\left(\sum_q e_q^2 H_{\Delta q}^1 \otimes \Delta q + \frac{H_{\Delta g}^1}{Q} \otimes \Delta g\right) = \frac{1}{2}\left(\sum_q e_q^2 \tilde{C}_q \otimes \Delta q + \frac{\tilde{C}_g^1}{Q} \otimes \Delta g\right) \end{cases} \tag{37}$$

We refer to [15] for an in-depth analysis of the collinear redefinitions and for the effective Lagrangian and to [16] for the specific application to QCD. It must be noted that the contributions that we have found here are not the full story. Once we go to higher powers of $1/Q$, there will be additional PDFs (for example dependent on $i\sigma^{\mu\nu}$ and $i\gamma^5\sigma^{\mu\nu}$) and additional structure functions, but treating these is outside of the scope of this work.

2.2 Renormalisation and Evolution

The factorisation from last section partially solves the issue of non-perturbative behaviour at low Q^2. Still there remain some divergences which need to be treated.

2.2.1 Infrared Divergence and The DGLAP Equations

If the PDFs, $\{f_{a/B}\}$ (standing for both polarised and unpolarised PDFs), are strictly interpreted as probability distributions, we would like to give an answer to the question:

"What is the probability of measuring a parton, given all the processes that could have lead to the production of this parton, up to some power of α_s?"

In this regard, a really helpful definition, is that of the *splitting functions*, $P_{a/b}(z)$, which give the probability that a parton of species a is created from b by emitting a gluon or quark (where in our analysis a and b are quarks and gluons), retaining collinear momentum fraction z. It is not hard

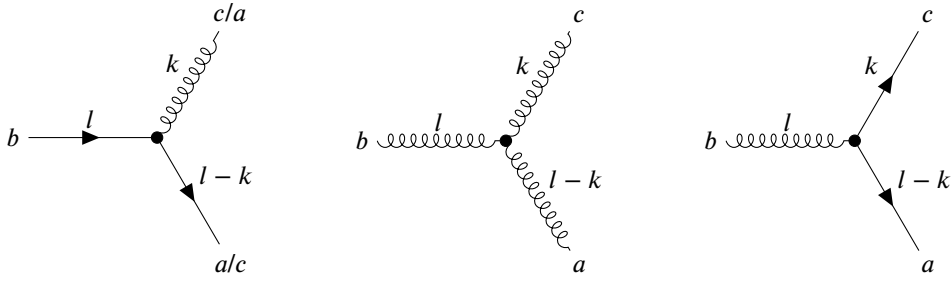
[5] The terms in front of g_1 remind us of the Pauli-Lubansky pseudo-vector form.

[6] A similar DIS process is possible for the weak force, where the photon is replaced by a Z-boson.

to see that up to order α_s, the chance of splitting parton a from parton b is approximately [17]

$$\begin{aligned} f_{a/b}(z) &= \delta_{ab}\,\delta(1-z) + \alpha_s \int_{Q_0}^{Q} \frac{dk_\perp^2}{k_\perp^2} P_{a/b}(z) + \mathcal{O}(\alpha_s^2) \\ &= \delta_{ab}\,\delta(1-z) + \alpha_s P_{a/b}(z) \ln \frac{Q^2}{Q_0^2} + \mathcal{O}(\alpha_s^2) \end{aligned} \tag{38}$$

where $k_\perp$ is the transverse gluon momentum which is unrestricted. Because of this fact, in the limit $Q_0 \to 0$, this leads to *infrared divergence*. We will solve this issue by introducing the *fac-*



(a) quark/gluon correction (b) gluon correction (c) quark correction

Figure 6: First order corrections to parton distribution functions where parton a is created from b by emission of c.

torisation scale, μ_F. The resulting correction for measuring parton a split off from B, given any intermediate parton b is equal to

$$\begin{aligned} \sum_b f_{a/b} \otimes f_{b/B} &\approx f_{a/B} + \alpha_s \ln \frac{Q^2}{Q_0^2} \sum_b P_{a/b} \otimes f_{b/B} \\ &= f_{a/B} + \alpha_s \left[\ln\left(\frac{Q^2}{\mu_F^2}\right) + \ln\left(\frac{\mu_F^2}{Q_0^2}\right) \right] \sum_b P_{a/b} \otimes f_{b/B} \end{aligned} \tag{39}$$

The part dependent on Q will be absorbed in H_a, while that dependent on Q_0 will be cancelled by the redefinition of the parton distribution functions. Note that this approach makes sense, since the PDFs are designed to capture all (non-perturbative) behaviour at small energy scales. Hence, let $f_{a/B}(x, Q^2) \to f_{a/B}(x, Q^2, \mu_F^2)$ under the constraint that the correction from above is independent of the factorisation scale

$$\frac{\partial f_{a/B}(x, Q_0^2, \mu_F^2)}{\partial \ln \mu_F^2} = \alpha_s \sum_b P_{a/b} \otimes f_{a/B}(x, Q_0^2, \mu_F^2) + \mathcal{O}(\alpha_s^2) \tag{40}$$

This system of coupled PDEs is called the *DGLAP equations*. The solution to the set of equations is given by

$$f_{a/B}(x, Q_0^2, \mu_F^2) = \sum_b \Gamma_{ab}(x, Q_0^2, \mu_F^2) \otimes f_{b/B}(x, Q_0^2) \tag{41}$$

where $\{f_{b/B}(x, Q_0^2)\}$ is the set of PDFs at the initial scale, Q_0, and $\{\Gamma_{ab}(x, \mu_F^2, Q_0^2)\}$ is called the set of *evolution kernel operators* from the initial scale to the factorisation scale. Note also that

$$\forall Q^2 \,:\, \Gamma_{ab}(x, Q^2, Q^2) = \delta_{ab}\delta(x) \tag{42}$$

since there is nothing to evolve. It turns out that the PDF basis[7]

$$\begin{cases} \varsigma = \sum_q (q + \overline{q}) & \text{(the } \textit{singlet} \text{ contribution)} \\ g & \\ q_{\text{NS}} & \text{(the } \textit{non-singlet} \text{ contributions)} \end{cases} \tag{43}$$

is such that the tuple (ς, g) decouples from q_{NS} and that we have

$$\begin{cases} q_{\text{NS}}(x, Q_0^2, \mu_F^2) = \Gamma_{\text{NS}}(x, Q_0^2, \mu_F^2) \otimes q_{\text{NS}}(x, Q_0^2) \\ \begin{pmatrix} \varsigma \\ g \end{pmatrix}(x, Q_0^2, \mu_F^2) = \begin{pmatrix} \Gamma_{qq} & 2n_f\Gamma_{qg} \\ \Gamma_{gq} & \Gamma_{gg} \end{pmatrix}(x, Q_0^2, \mu_F^2) \otimes \begin{pmatrix} \varsigma \\ g \end{pmatrix}(x, Q_0^2) \end{cases} \tag{44}$$

Here, the non-singlet, q_{NS} stands for a family of contributions. A few examples, some of which we will later need, are

$$\begin{cases} V = (u - \overline{u}) + (d - \overline{d}) + (s - \overline{s}) \\ V_3 = (u - \overline{u}) + (d - \overline{d}) \\ T_3 = (u + \overline{u}) - (d + \overline{d}) \\ T_8 = (u + \overline{u}) + (d + \overline{d}) - 2(s + \overline{s}) \end{cases} \tag{45}$$

For a more detailed overview, we refer to [18]. The polarised counterparts come in comparable form and carry a Δ in front. We will refer to this basis as the *(polarised) evolution basis*.

2.3 Massive Corrections

In recent studies[19], heavier quark contributions have been included in the framework. To this end, next to the lighter quarks (u, d, s), massive contributions from the charm quark are no longer neglected. There are several ways to model the contributions from the charm quarks. In the first one, charm corrections are included in the parton distribution functions. The scheme which makes use of this is called the *Zero Mass, Variable Flavour Number Scheme (ZM-VFNS)* and works by varying the amount of PDFs for different energetic regimes

$$n_f = \begin{cases} n_\ell & \forall Q^2 \in [0, m_c^2] \\ n_\ell + 1 & \forall Q^2 \in (m_c^2, \infty) \end{cases} \tag{46}$$

[7] We have chosen for the symbol ς instead of Σ to be able to make a distinction between the density and the total, i.e. $\Sigma = \int dx\, \varsigma$.

where n_ℓ is the number of light quarks. In the second model, charm corrections are absorbed in the Wilson coefficients in a scheme called *Fixed Flavour Number Scheme (FFNS)* and we have

$$H_a = H_a\left(x, Q^2, \frac{Q^2}{m_c^2}\right) \tag{47}$$

Both methods have their disadvantages. ZM-VFNS is less accurate in the $Q^2 \ll m_c^2$ region, while FFNS is less accurate in the $Q^2 \gg m_c^2$ region. This has to do with the factorisation theorem, where particles are either regarded as hard or collinear, but certainly not both, while for $Q^2 \sim m_c^2$ this distinction is not that clear. A better model, however, is already achieved by combining the two previous schemes in a framework which is called FONNL, where for every structure function, we have

$$F^{\text{FONLL}} = F^{n_\ell+1} + F^{n_\ell}\left(\frac{Q^2}{m_c^2}\right) - \lim_{m_c \to 0} F^{n_\ell}\left(\frac{Q^2}{m_c^2}\right) \tag{48}$$

Still, in the region where $Q^2 \sim m_c^2$ – the so-called *threshold region* –, the charm quark is both present and not light. This makes that continuity is not necessarily achieved and we need other tools [20].

3 Methodology

In order to describe the non-perturbative (p)PDFs, one must choose the right methodology. Since analytic derivations have not been found, the step towards numerical methods is soon made. There are many groups working on such numerical models, such as HERA [21], ABM [22] and CTEQ [23]. In the following section, we describe the framework of the NNPDF project, more specifically, the extension to include polarised PDFs. Our next-to-next leading order determination of pPDFs is following up NLO results from NNPDFpol1.0 [24] and DSSV [25] among others.

3.1 NNPDF Architecture

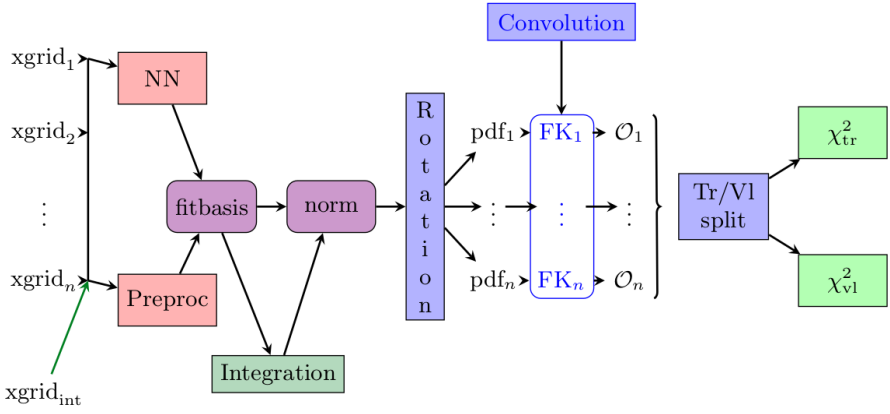


Figure 7: NNPDF neural network architecture [26].

We use a single neural network that produces the basis from section 3.2. The model learns based on gradient-descent and uses hyperparameter optimisation to specify model parameters. To furthermore speed up the fitting procedure, the neural network is additionally scaled in small-x and high-x regimes according to the expected limiting behaviour and normalised by the sum rules that we treat in section 3.2.1.

$$xf_{a/B}(x, Q_0^2) = A_a x^{1-\alpha_a}(1-x)^{\beta_a}\text{NN}_a(x) \tag{49}$$

with A_a the normalisation and α_a and β_a variable during the fitting procedure. To propagate errors coming from the experimental data, the central values of the fitting data is statistically fluctuated and an ensemble of replicas is generated.

3.2 PDF Basis and Theoretical Constraints

Since at this point we only possess DIS, W, and (di)jet data, we have to limit the amount of basis pPDFs to fit. The set that is used is

$$\{\Delta\varsigma, \Delta g, \Delta T_3, \Delta T_8, \Delta V, \Delta V_3\} \tag{50}$$

at discrete (x_i, Q_0^2). The $\Delta\varsigma$, ΔT_3, ΔT_8 can be fitted because of the available DIS data (section 3.6.1), which consists of proton and deuteron targets, since at leading order

$$\begin{cases} g_1^p \propto \frac{1}{9}\Delta\varsigma + \frac{1}{12}\Delta T_3 + \frac{1}{36}\Delta T_8 \\ g_1^d \propto \frac{1}{9}\Delta\varsigma + \frac{1}{36}\Delta T_8 \end{cases} \tag{51}$$

Moreover, because of the charged current property of the W-data that we will discuss in section 3.6.3, the negative PDF combinations of the V and V_3 contributions can be determined. In order to limit the space of all pPDF configurations to be searched by the neural network, we need to impose constraints a priori. These stem from a combination of theoretical and experimental considerations. After the individual contributions, a cost function incorporating the total can be defined.

3.2.1 Sum Rules

There are sum rules for both PDFs and pPDFs. A few examples in the unpolarised case, are that the number of valence quarks should be conserved

$$\begin{cases} \int_0^1 dx\,(u(x) - \overline{u}(x)) = 2 \\ \int_0^1 dx\,(d(x) - \overline{d}(x)) = 1 \end{cases} \tag{52}$$

and that the momentum should be conserved

$$\int_0^1 dx\,(\varsigma(x) + g(x)) = 1 \tag{53}$$

In our new framework, we also have sum rules. In the early days of QCD, it was believed that the quarks u, d, s exhibited $SU(3)$ *flavour symmetry*. Later, it turned out that this was not the case and that it is an approximate symmetry. We can measure the extent to which the symmetry is broken in baryon decays by considering the two constants

$$\int_0^1 dx\,\Delta T_3(x, Q^2) = a_3 \tag{54}$$

$$\int_0^1 dx\,\Delta T_8(x, Q^2) = a_8 \tag{55}$$

with

$$\begin{cases} a_3 = 1.2701 \pm 0.0025 \\ a_8 = 0.585 \pm 0.025 \end{cases} \tag{56}$$

and require that the pPDFs obey these constraints. Where we refer to equation 45. This strategy was also imposed in earlier works of NNPDF and it is interesting to note that even when the latter constraints are varied by similar neural network architectures, similar values can be achieved [24].

3.2.2 Integrability

Another property that a properly defined probability distribution should have, is that it is integrable. It is a common fact that for t^α, $\alpha = -1$ is the transition point for $\int_0^a dt\, t^\alpha$ to be integrable. Hence, the normalised condition

$$\frac{I(x_1, x_2)}{I(x_1', x_2)} < \frac{\ln \frac{x_2}{x_1}}{\ln \frac{x_2}{x_1'}} \tag{57}$$

for suitably chosen $x_1 = 1e-5$, $x_1' = 2e-5$, $x_2 = 1e-4$ exactly encapsulates this criterion for our model. A comparable strategy was also used in NNPDF4.0.

3.2.3 Positivity

The pPDFs are allowed to take on both positive and negative values. This makes restricting them somewhat trickier than for unpolarised PDFs. However, we can impose a constraint on the absolute value. Using the additional headroom of one standard deviation, we can consider the difference

$$\mathscr{C}_{a/B} = f_{a/B}(x_i) + \sigma_{a/B}(x_i) - |\Delta f_{a/B}(x_i)| \tag{58}$$

Then, using an activation function, the new positivity contribution to the cost function is

$$\chi^2_{\text{POS}} = \Lambda_{\text{POS}} \sum_{k=1}^{8} \sum_{i=1}^{n_i} \text{ReLu}(-\mathscr{C}_{a/B}(x_i, Q^2)) \tag{59}$$

3.3 Polarised Jets Code

Accept for fitting to DIS data, the pPDFs can be used to generate predictions for proton-proton collisions. To this end, we use a preexisting code [27] that takes in a set of PDFs and pPDFs and produces the double spin asymmetry for a range of kinematic variables (see section 3.6.2), for both inclusive and exclusive jets. The code has been written in Fortran and since the publication in '98 has been modernised for NNPDFpol1.0 and now again for NNPDFpol2.0.

3.4 FK-tables

Since the analytic calculation of the structure functions 37 is quite computationally heavy and involves solving the DGLAP equations 40, we can introduce a discretisation procedure via the generation of *Fast Kernel (FK) tables*. The substitution that is performed in the numerical framework is, for every relevant observable, F,

$$F(x_i, Q^2) = \sum_{a,i} \mathrm{FK}_a(x_i, Q_0^2, Q^2) \otimes f_{a/B}(x_i, Q_0^2) \tag{60}$$

where dependence on the factorisation scale is suppressed by taking $\mu_F^2 = Q^2$ and we have

$$\mathrm{FK}_a(x_i, Q_0^2, Q^2) = \sum_b H_b(x_i, Q^2) \otimes \Gamma_{ba}(x_i, Q_0^2, Q^2) \tag{61}$$

for some discrete set $\{x_i\}$. Using the defined tables, we can interpolate for x with respect to the descretisation. Since all the elements that constitute the FK coefficients are fully perturbative in nature, the tables can be calculated à priori. This makes generating predictions depending on the PDFs and thus the fitting procedure a matter of multiplying precomputed tensors with parametrised functions.

3.5 Reweighting and Unweighting

Since the framework of NNPDF currently accounts for a fixed set of processes (DIS, Drell-Yan, etc.) it is highly beneficial to have a mechanism of modifying PDFs in another way. Suppose for one, that we have an algorithm, which takes in PDFs and generates functions,

$$A : f \mapsto Af \in \mathbb{R}^* \tag{62}$$

and additionally a set of N datapoints and their corresponding observables

$$D = (x, y) \in \mathbb{R}^N \times \mathbb{R}^N \tag{63}$$

then we can compare the datapoints and predictions by use of the error metric

$$\chi^2(y, f) = (y - Af(x))\, \sigma^{-1}\, (y - Af(x)) \tag{64}$$

with σ the experimental covariance matrix. Given now an ensemble of PDFs, $\{f\}$, we can incorporate the error of the additional experimental data into the weighted mean by a procedure which is called *Bayesian reweighting*. This procedure, under assumption of a Gaussian likelihood around the mean of the data, states that the weights of the new ensemble must be chosen that

$$w_i \propto \chi^{N-1} \exp\left(-\frac{1}{2}\chi^2(f_i, y)\right) \tag{65}$$

where i refers to the replica number. The weights that are found that are found this way are decimal numbers and we would prefer to change back to an integer for each replica, u_i, that in limit

$$N'_{\mathrm{rep}} = \sum_i u_i \to \infty \tag{66}$$

we have

$$\frac{u_i}{\sum_i u_i} \to \frac{w_i}{\sum_i w_i} \tag{67}$$

To do so, first define

$$p_i = \Big(\sum_i w_i \Big)^{-1} w_i \tag{68}$$

and

$$P_i = \sum_{j=1}^{i} p_j \tag{69}$$

Then, we choose the unweighted weights according to [Ball2011]

$$u_i = \sum_{j=1}^{N'_{\text{rep}}} \theta\left(\frac{j}{N'_{\text{rep}}} - P_{i-1} \right) \theta\left(P_i - \frac{j}{N'_{\text{rep}}} \right) \tag{70}$$

The choice of N'_{rep} is a subtle issue. On the one hand, the value should not be too small, since then a lot of information will be thrown away. On the other hand, it should not be too large, since in this case we would add non-existent information to our system. To quantify the amount of information, we introduce the effective amount of replicas [28]

$$N_{\text{eff}} = \exp\left\{ -\frac{1}{N} \sum_{i=1}^{N} w_i \log(w_i/N) \right\} \tag{71}$$

and demand that $N'_{\text{rep}} \lesssim N_{\text{eff}}$. We can moreover visualise the effect of the choice of N'_{rep} by considering the information entropy (Kullback-Leibler divergence) of the unweighted weights compared to the reweighted weights

$$H = D_{\text{KL}}(u \| w) = \sum_i u_i \log \frac{u_i}{w_i} \tag{72}$$

3.6 Data

Below, the various types of data and their physical background can be found. The tables with the exact experiments and their properties can be found in the appendices.

3.6.1 DIS Data

The main contribution to the data is that coming from DIS experiments. The form in which these data points are delivered is

$$P = (x, Q^2, G) \tag{73}$$

with x and Q^2 the independent variables and for DIS data the dependent variable, G, is either g_1 or g_1/F_1. In order to make predictions for these structure functions, the FK-tables can be

generated from the kinematics (x, Q^2). For g_1/F_1, more than one FK-table is needed, since we want to calculate the quantity

$$\frac{g_1}{F_1}(x, Q^2) = \frac{\sum_a \mathrm{FK}_a^{g_1}(x, x_i, Q^2, Q_0^2) \otimes f_{a/B}(x_i^2, Q_0^2)}{\sum_b \mathrm{FK}_b^{F_1}(x, x_i, Q^2, Q_0^2) \otimes f_{b/B}(x_i^2, Q_0^2)} \tag{74}$$

The complete list of DIS datasets used for the fits can be found in table 1.

3.6.2 Jet Data

Due to the universality of the PDF sets, the predictive power of our model is not constrained to only DIS. Another process which we can study is (di)jet production resulting from proton-proton collisions. The form in which the data is delivered is either

$$P = (p_T, A_{LL}) \tag{75}$$

for single jets (*inclusive jet*) and

$$P = (m_{jj}, A_{LL}) \tag{76}$$

for dijets (*exclusive jet*), where p_T is the transverse proton momentum, m_{jj} the invariant dijet mass and A_{LL} the longitudinal double spin asymmetry. The latter variable is dependent on the polarised cross section

$$\sigma^{\pm\pm} = \sum_{a,b} \hat{\sigma}_{ab}^{\pm\pm} \otimes f_{a/B} \otimes f_{b/B} \tag{77}$$

in the following way

$$A_{LL} = \frac{\sigma^{++} - \sigma^{+-}}{\sigma^{++} + \sigma^{+-}} \tag{78}$$

Here, $\hat{\sigma}_{ab}^{\pm\pm}$ is the hard cross section, serving a comparable role as the hard coefficient functions, H_a. The complete list of jet and dijet datasets used for the unweighting can be found in table 2.

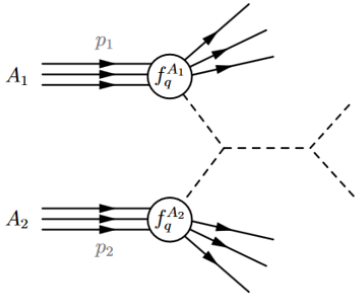


Figure 8: Typical configuration for (di)jet production in proton-proton collisions.

3.6.3 W-data

Next to the DIS and (di)jet processes, polarised PDFs play a role in $W^{\pm}$ production in the Drell-Yann charged current process. Precise treatment of this process is outside the scope of this work, but for a recent study we refer to [29] [30]. The relevant observable is the longitudinal single spin asymmetry

$$A_L = \frac{\sigma^+ - \sigma^-}{\sigma^+ + \sigma^-} \tag{79}$$

where $\sigma^{\pm}$ is the cross section for producing a W-boson when the helicity of the longitudinally polarised proton beam is positive/negative. The data can be found in table 3.

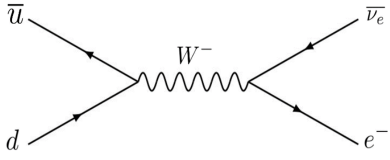


Figure 9: Typical configuration for W-production in charged current Drell-Yan processes.

3.6.4 EIC Pseudo-data

The EIC pseudo-data (being a DIS process) looks like the aforementioned DIS data and comes in two forms, the first being

$$P = (x, Q^2, A_{LL}) \tag{80}$$

where in this case we can approximate $A_{LL} \approx g_1/F_1$. The second is

$$P = (x, Q^2, g_1^c/F_1^c) \tag{81}$$

which needs some explanation. In the massive scheme, where $n_f = 3$ (see FFNS in 2.3), the presence of the charm quark has to be deduced from non-charm PDFs. This system is hence referred to as *exclusive charm*, to be compared with the mass corrections in 2.3. The leading term ($\mathcal{O}(\alpha_s)$) that accounts for its presence, involves the gluon PDF (see figure 6c). To this extent, we can define the (un)polarised charm structure functions as [31]

$$F_1^c(x, Q^2, m_c^2) = e_c^2 H_g^c \otimes g \tag{82}$$

and

$$g_1^c(x, Q^2, m_c^2) = e_c^2 H_{\Delta g}^c \otimes \Delta g \tag{83}$$

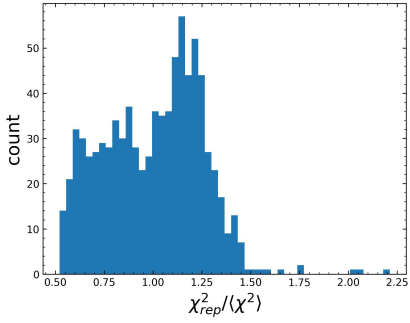
The complete overview of the EIC projection data can be found in table 4.

4 Results

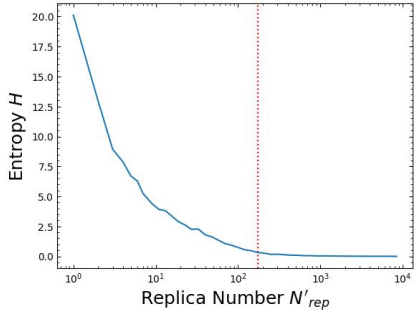
Using the NNPDF framework, PDFs were obtained by fitting to, first (1) DIS-only data, which was used for creating EIC projections. Later, also (2) W-data was included. This was used to unweight and compare to a (3) DIS + W + Jet fit, all at NNLO.

4.1 Reweighting

Reweighting and unweigthing were performed by making 4 different polarised jet predictions. The combinations are: two different center of mass energies (200 and $510GeV$) with the two jet types: inclusive jet and dijet. The ensemble of data was then used to generate a single χ^2 per replica, as correlations between datasets can then be accounted for. The error per replica distribution turned out to be rather smooth and approximately centered at 1 as can be seen in the figure below. With an effective amount of replicas of $N_{\text{eff}} = 173$, the entropy is around $H \approx 1$, which is low enough to retain quite a lot of information.



(a) χ^2 as distribution over the replicas before reweighting, binned in 100 bins.



(b) Entropy as a function of the new replica number after unweighting. The red line corresponds to the effective number of replicas, $N_{\text{eff}} = 173$, as provided in equation 71.

Figure 10: Summary of the unweighting run.

Experiment	Target	N_{dat}	x	Q^2[GeV^2]	Observable	Ref.
EMC	p	10	.015 - .466	3.5 - 29.5	g_1	[32]
SMC	d	13	.002 - .48	.50 - 54.80	g_1	[33]
	p	13	.002 - .48	.50 - 54.80	g_1	[33]
E142	n	8	.035 - .466	1.1 - 5.5	g_1	[34]
E143	d	28	.031 - .749	1.27 - 9.52	g_1	[35]
	p	28	.035 - .466	1.27 - 9.52	g_1	[35]
E154	n	11	.017 - .024	1.2 - 15.0	g_1	[36]
E155	p	24	.015 - .750	1.22 - 34.72	g_1/F_1	[37]
	n	24	.015 - .750	1.22 - 34.72	g_1/F_1	[37]
JLAB E06 014	n	6	.277 - .548	3.078 - 3.078	g_1/F_1	[38]
JLAB E97 103	n	5	.160 - .200	.57 - 1.34	g_1	[39]
JLAB E99 117	n	3	.33 - .60	2.71 - 4.83	g_1/F_1	[40]
JLAB EG1 DVCS	d	44	.158 - .574	1.078 - 4.666	g_1/F_1	[41]
	p	47	.154 - .578	1.064 - 4.115	g_1/F_1	[41]
COMPASS-D	d	15	.0046 - .567	1.1 - 60.8	g_1	[42]
COMPASS-P	p	17	.0036 - .57	1.1 -67.4	g_1	[43]
HERMES97	n	9	.033 - .464	1.22 - 5.25	g_1	[44]
HERMES	p	15	.0264 - .7248	1.12 - 12.21	g_1	[45]
	d	15	.0264 - .7248	1.12 - 12.21	g_1	[45]
Total DIS		335				

Table 1: The available polarised DIS data. For each measurement, we report the name of the experiment, the nature of the polarised target, the number of data points, the kinematic coverage in the (x, Q^2) plane, the measured observable and the corresponding reference.

Experiment	Type	N_{dat}	p_T/m_{jj} [GeV]	$\sqrt{s}$ [GeV]	Observable	Ref.
PHENIX	1JET	8	2.4 - 10.8	200	A_{LL}	[46]
STAR 2005	1JET	10	2.4 - 10.8	200	A_{LL}	[47]
STAR 2006	1JET	9	8.5 - 34.7	200	A_{LL}	[47]
STAR 2009	1JET	22	5.5 - 32.4	200	A_{LL}	[48]
	2JET	33	17.06 - 67.88	200	A_{LL}	[49]
STAR 2012	1JET	14	6.758 - 55.29	510	A_{LL}	[50]
	2JET	42	19.61 - 110.81	510	A_{LL}	[50]
STAR 2013	1JET	14	8.65 - 63.32	510	A_{LL}	[51]
	2JET	49	14.2 - 133.2	510	A_{LL}	[51]
STAR 2015	1JET	22	5.8 - 33.5	200	A_{LL}	[52]
	2JET	14	19.76 - 71.27	200	A_{LL}	[52]
Total JET		237				

Table 2: The available polarised Jet data. For each measurement, we report the name of the experiment, the number of data points, the kinematic coverage in either p_T for 1JET and m_{jj} for 2JET, the center of mass energy, the measured observable and the corresponding reference.

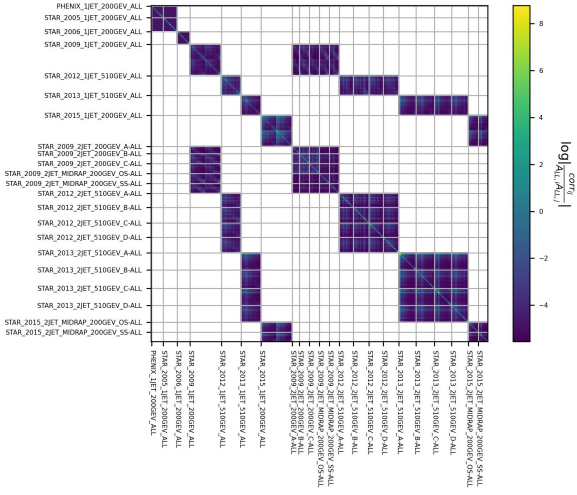


Figure 11: Correlation between jet and dijet datasets mentioned in table 2 based on the relative covariance.

Experiment	N_{dat}	η_e	$\sqrt{s}$ [GeV]	E_T [GeV]	Observable	Ref.
STAR	12	-1.25 - 1.25	510	25-50	A_L	[53]
Total W	12					

Table 3: The available charged current W production data. For each measurement, we report the name of the experiment, the number of data points, the kinematic coverage in η_e, $\sqrt{s}$ and E_T, the measured observable and the corresponding reference.

Experiment	Target	N_{dat}	x	Q^2[GeV^2]	Observable	Ref.
EIC	p	31	.00024-.1036	2.55-52.39	g_1^c/F_1^c	[54]
ATHENA	p	512	.00013-.815	1.29-5150	A_{LL}^p	[55]
Total EIC		543				

Table 4: The available polarised EIC projection data. For each measurement, we report the name of the experiment, the nature of the polarised target, the number of data points, the kinematic coverage in the (x, Q^2) plane, the measured observable and the corresponding reference.

4.2 Polarised jets

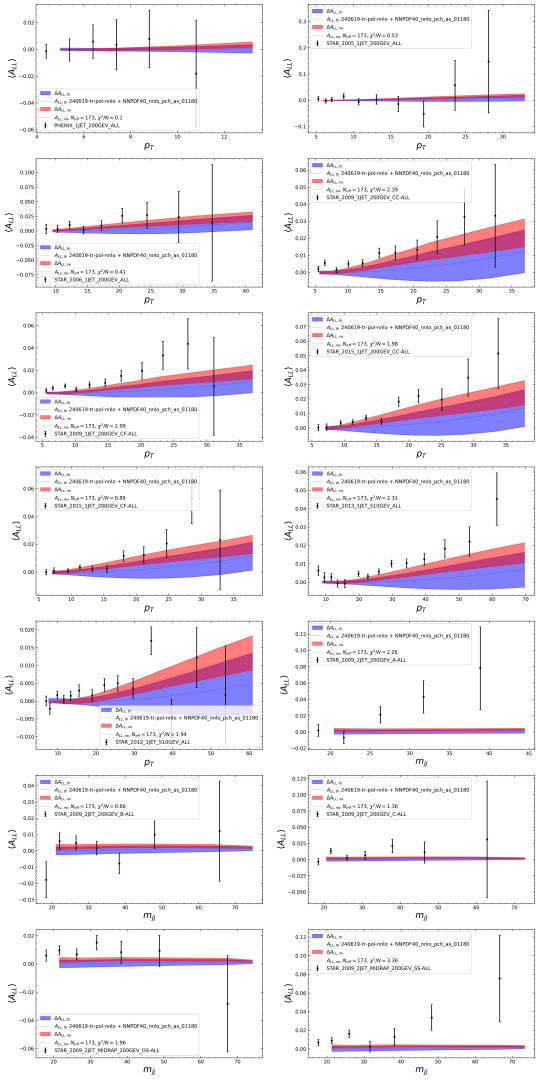


Figure 12: Comparison between the longitudinal double spin asymmetry before (blue) and after unweighting (red), for various experiments listed in table 2. at NNLO pt. 1.

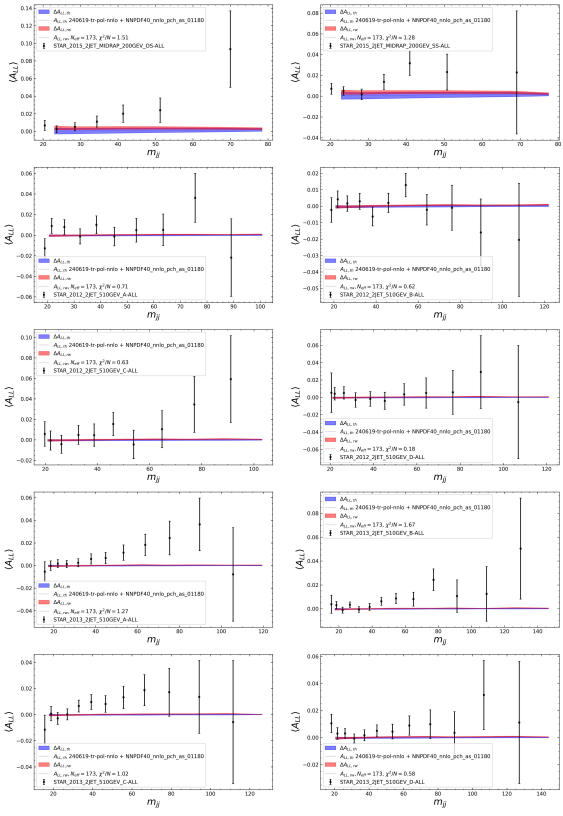


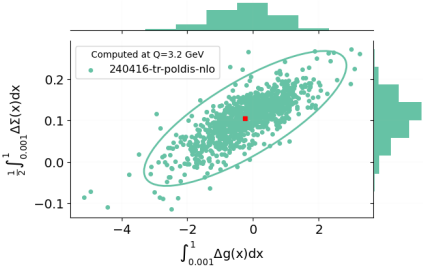
Figure 13: Comparison between the longitudinal double spin asymmetry before (blue) and after unweighting (red), for various experiments listed in table 2. at NNLO pt. 2.

4.3 pPDFs

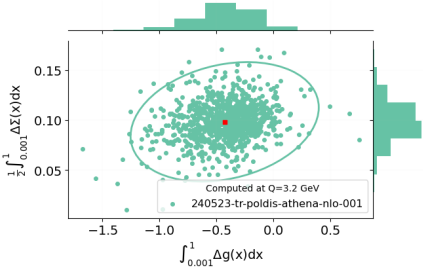
Now, we can compare the pPDFS at next-to-next leading order. Next to including the (di)jet data by reweighting, a method was found to generate dedicated FK-tables. Hence, we also compare both cases (unweighted Jet v.s. Jet). In the figures in 15, we see that the results are quite similar to NNPDFpol1.0. The gluon distributions, however, are mostly different. Moreover, the V and V_3 at low x have a higher uncertainty then for NNPDFpol1.0.

4.4 EIC Projection

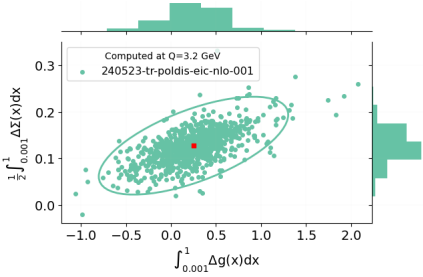
After the inclusion of respectively the EIC and Athena data, we can observe what the restricting effect is on the singlet and gluon spin. In figure 14, we can see the result.



(a) The singlet and spin contributions.

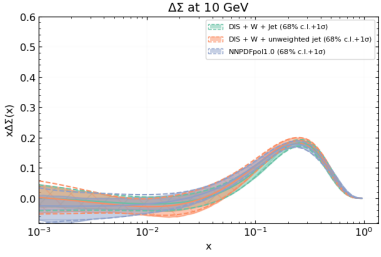


(b) The singlet and spin contributions after inclusion of Athena projections.

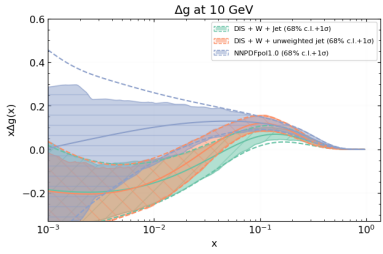


(c) The singlet and spin contributions after inclusion of EIC projections.

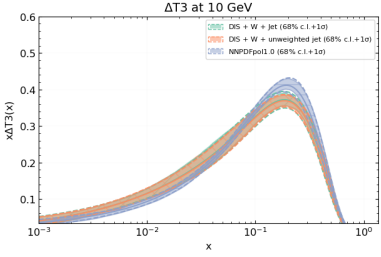
Figure 14: Comparison between NNPDFpol2.0 (DIS only) and Athena and EIC projection data at NLO for the singlet and gluon contributions at $3.2GeV$.



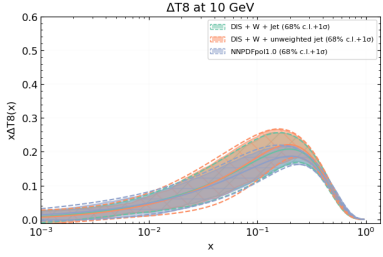
(a) The singlet spin contribution at NNLO.



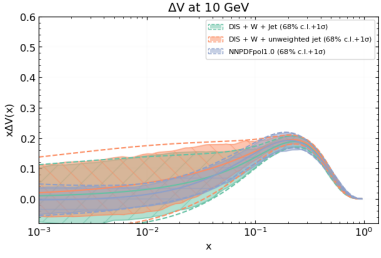
(b) The gluon spin contribution at NNLO.



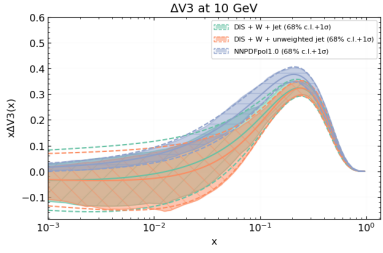
(c) The ΔT_3 contribution at NNLO.



(d) The ΔT_8 contribution at NNLO.



(e) The ΔV contribution at NNLO.



(f) The ΔV_3 contribution at NNLO.

Figure 15: Comparison between NNPDFpol1.0 and NNPDFpol2.0 (unweighted and non-reweighted) at NNLO for the singlet and gluon contributions at $1 GeV$.

5 Conclusion

We have derived the framework to compute parton distribution functions from first principles up to twist-2 and motivated the physical background. Moreover, with the adjusted NNPDF framework, we have been able to create pPDFs at NNLO. In constraining the space of parton distribution functions using the positivity bound of 59, the question arises why 1 standard deviation is the optimal choice. The answer to this is that it probably is not. Next to this, we could also choose to add the standard deviation of the polarised distributions themselves. However, this would require us to iteratively calculate the full set of polarised replicas and hope for convergence. A possibility of making the bound more 'fair' for the former issue, would be to implement a cost function dependent on the quotient with the standard deviation instead.

Then, using the principle of unweighting, these results were refined to match proton-proton collisions better. The predictions done by the polarised jets code seem to be suitable to be reweighted by the inclusive jet data, however, the dijet predictions follow the trend in the data less well. This can have various reasons. One of the major issues with the unweighting was the code that produces the polarised dijet predictions. Various cuts, such as η and p_T, have until now not been imposed on the dijet, even though the data does provide values for them. This definitely should be attempted in the future. There also seems to be another reason for the flatlining behaviour of the fits.

Then, we can turn to comparing pPDFs with NNPDFpol1.0. Both the singlet and non-singlet distributions seem to be mostly unaffected. This hints at the conclusion that massive effects do not play a significant role at the energy regimes in question. The unweighted and fitted jet seem to largely agree in most regimes, however, the fitted jet definitely seems smoother (due to a higher replica number). The fact that the gluon PDF of NNPDFpol2.0 (figure 14b) mostly differs from NNPDFpol1.0 in the NNLO case can be explained by the additional corrections that are imposed in our novel framework. The overall negative sign, however, is not immediately expected, but at the same time not immediately punished by the reweighting procedure. The question arises whether this could be a genuine result. The answer to this question, is most probably given in a recent study [56], highlighting that gluon-gluon fusion

$$gg \to H \tag{84}$$

is dependent on the gluon pPDF in such a way that $\Delta g < 0$ leads to a negative cross-section, which is unphysical. Thus, with the current data, it is fair to say that most probably we have an erroneous sign. There might be multiple solutions to mitigate this problem, some of which are looking at gluon-gluon fusion itself or waiting for the results of the EIC.

Moreover, we considered projections for the novel EIC experiments. The initial reduction that we observe is not as large as seen in the proposal 3. However, relatively speaking, the Athena data does shrink the error by quite a significant amount, still preferring negative gluon spin. On the other hand we have less reduction of the error bands for the EIC data, but with some preferred shift in the gluon pPDF. We have some remarks here as well. The NNPDFpol2.0 neural network was trained on primarily available DIS data and checked using jet data, both in a specified range (see 1, 2). When choosing to use the fitted model to make predictions about the EIC (which has data points outside of the training bounds 4), we essentially believe that our model has

some emergent extrapolation power. Even though limiting behaviour for small-x has previously been studied and imposed in our framework, according to the author, all the produced projections should be taken with a grain of salt and should merely be considered as a rough sketch of how our own models possibly will change when new data is available.

Summarising, we have made some steps towards NNPDFpol2.0 and we remain hopeful that by including the dijet data in a correct manner and propagating the new fits, an optimal set of pPDFs and projections can be found.

References

[1] Michiel Botje. Lecture notes Particle Physics II QCD in Eight Lectures. 2013. URL: http://www.nikhef.nl/user/h24/qcdcourse.

[2] Walter Greiner, Stefan Schramm, and Eckart Stein. Quantum chromodynamics. Springer Berlin Heidelberg, 2007, pp. 1–553. ISBN: 3540485341. DOI: 10.1007/978-3-540-48535-3/COVER.

[3] The spin of a proton - Physics World. URL: https://physicsworld.com/a/the-spin-of-a-proton/.

[4] Masashi Wakamatsu. "Is gauge-invariant complete decomposition of the nucleon spin possible ?" In: International Journal of Modern Physics A 29 (9 Feb. 2014). DOI: 10.1142/S0217751X14300129. URL: http://arxiv.org/abs/1402.4193%20http://dx.doi.org/10.1142/S0217751X14300129.

[5] Masashi Wakamatsu. "Unravelling the physical meaning of the Jaffe-Manohar decomposition of the nucleon spin". In: Physical Review D 94 (5 July 2016). DOI: 10.1103/PhysRevD.94.056004. URL: http://arxiv.org/abs/1607.04018%20http://dx.doi.org/10.1103/PhysRevD.94.056004.

[6] Paul Anderson et al. "Complete Formalism of Cross Sections and Asymmetries for Longitudinally and Transversely Polarized Leptons and Hadrons in Deep Inelastic Scattering". In: 30th International Workshop on Deep-Inelastic Scattering and Related Subjects. May 2023. arXiv: 2306.00097 [hep-ph].

[7] Richard D. Ball et al. "The path to proton structure at 1% accuracy". In: Eur. Phys. J. C 82.5 (2022), p. 428. DOI: 10.1140/epjc/s10052-022-10328-7. arXiv: 2109.02653 [hep-ph].

[8] Richard D. Ball et al. "An open-source machine learning framework for global analyses of parton distributions". In: Eur. Phys. J. C 81.10 (2021), p. 958. DOI: 10.1140/epjc/s10052-021-09747-9. arXiv: 2109.02671 [hep-ph].

[9] Emanuele Roberto Nocera. "Unbiased spin-dependent Parton Distribution Functions". In: (Mar. 2014). URL: https://arxiv.org/abs/1403.0440v1.

[10] A. Accardi et al. "Electron Ion Collider: The Next QCD Frontier: Understanding the glue that binds us all". In: Eur. Phys. J. A 52.9 (2016). Ed. by A. Deshpande, Z. E. Meziani, and J. W. Qiu, p. 268. DOI: 10.1140/epja/i2016-16268-9. arXiv: 1212.1701 [nucl-ex].

[11] Frederik Van der Veken. "Wilson lines : applications in QCD". PhD thesis. Antwerp U., 2014.

[12] Aneesh V. Manohar. "Polarized parton distribution functions". In: Phys. Rev. Lett. 66 (3 Jan. 1991), pp. 289–292. DOI: 10.1103/PhysRevLett.66.289. URL: https://link.aps.org/doi/10.1103/PhysRevLett.66.289.

[13] Xiangdong Ji, Yang Xu, and Yong Zhao. "Gluon Spin, Canonical Momentum, and Gauge Symmetry". In: Journal of High Energy Physics 2012 (8 May 2012). DOI: 10.1007/JHEP08(2012)082. URL: http://arxiv.org/abs/1205.0156%20http://dx.doi.org/10.1007/JHEP08(2012)082.

[14] M. Anselmino, P. Gambino, and J. Kalinowski. "Polarized deep inelastic scattering at high energies and parity violating structure functions". In: Zeitschrift für Physik C Particles and Fields 64 (2 Jan. 1994), pp. 267–273. DOI: 10.1007/BF01557397. URL: http://arxiv.org/abs/hep-ph/9401264%20http://dx.doi.org/10.1007/BF01557397.

[15] Thomas Becher, Alessandro Broggio, and Andrea Ferroglia. "Introduction to Soft-Collinear Effective Theory". In: 896 (Oct. 2014). DOI: 10.1007/978-3-319-14848-9. URL: http://arxiv.org/abs/1410.1892%20http://dx.doi.org/10.1007/978-3-319-14848-9.

[16] Iain W Stewart. Lectures on the Soft-Collinear Effective Theory. 2013.

[17] John C Collins et al. "Factorization of Hard Processes in QCD". In: (Sept. 2004), pp. 1–91. DOI: 10.1142/9789814503266_0001. URL: https://arxiv.org/abs/hep-ph/0409313v1.

[18] Perturbative PDF evolution – NNPDF documentation. URL: https://docs.nnpdf.science/theory/PTevol.html.

[19] Richard D. Ball et al. "Intrinsic charm in a matched general-mass scheme". In: Physics Letters, Section B: Nuclear, Elementary Particle and High-Energy Physics 754 (Sept. 2015), pp. 49–58. DOI: 10.1016/j.physletb.2015.12.077. URL: http://arxiv.org/abs/1510.00009%20http://dx.doi.org/10.1016/j.physletb.2015.12.077.

[20] Stefano Forte et al. "Heavy quarks in deep-inelastic scattering". In: Nuclear Physics B 834 (1-2 July 2010), pp. 116–162. ISSN: 05503213. DOI: 10.1016/j.nuclphysb.2010.03.014.

[21] Amanda Cooper-Sarkar. "PDF Fits at HERA". In: (Dec. 2011). URL: https://arxiv.org/abs/1112.2107v1.

[22] S. Alekhin, J. Blümlein, and S. Moch. "The ABM parton distributions tuned to LHC data". In: Phys. Rev. D 89 (5 Mar. 2014), p. 054028. DOI: 10.1103/PhysRevD.89.054028. URL: https://link.aps.org/doi/10.1103/PhysRevD.89.054028.

[23] Sayipjamal Dulat et al. "New parton distribution functions from a global analysis of quantum chromodynamics". In: Phys. Rev. D 93 (3 Feb. 2016), p. 033006. DOI: 10.1103/PhysRevD.93.033006. URL: https://link.aps.org/doi/10.1103/PhysRevD.93.033006.

[24] Emanuele R. Nocera et al. "Unbiased Polarised Parton Distribution Functions and their Uncertainties". In: (June 2012), p. 5. URL: https://arxiv.org/abs/1206.0201v2.

[25] Daniel de Florian et al. "Global Analysis of Helicity Parton Densities and Their Uncertainties". In: Physical Review Letters 101 (7 Apr. 2008). DOI: 10.1103/PhysRevLett.101.072001. URL: http://arxiv.org/abs/0804.0422%20http://dx.doi.org/10.1103/PhysRevLett.101.072001.

[26] NNPDF Collaboration. Methodology overview. URL: https://docs.nnpdf.science/n3fit/methodology.html.

[27] D. de Florian et al. "Next-to-leading order jet cross sections in polarized hadronic collisions". In: Nuclear Physics B 539 (3 Aug. 1998), pp. 455–476. DOI: 10.1016/S0550-3213(98)00673-7. URL: http://arxiv.org/abs/hep-ph/9808262%20http://dx.doi.org/10.1016/S0550-3213(98)00673-7.

[28] NNPDF Collaboration et al. "Reweighting and Unweighting of Parton Distributions and the LHC W lepton asymmetry data". In: (Aug. 2011). URL: http://arxiv.org/abs/1108.1758.

[29] Claude Duhr, Falko Dulat, and Bernhard Mistlberger. "Charged Current Drell-Yan Production at N3LO". In: Journal of High Energy Physics 2020 (11 July 2020). ISSN: 10298479. DOI: 10.1007/jhep11(2020)143. URL: https://arxiv.org/abs/2007.13313v2.

[30] Justin R. Stevens. "Longitudinal Single-Spin Asymmetry and Cross Section for $W^{\pm}$ Boson production in Polarized Proton-Proton Collisions at $\sqrt{s}$ = 500 GeV". PhD thesis. Indiana U., 2012.

[31] "Heavy Quarks in Polarised Deep-Inelastic Scattering at the Electron-Ion Collider". In: (Jan. 2024). URL: https://arxiv.org/abs/2401.10127v2.

[32] J. Ashman et al. "An Investigation of the Spin Structure of the Proton in Deep Inelastic Scattering of Polarized Muons on Polarized Protons". In: Nucl. Phys. B 328 (1989). Ed. by V. W. Hughes and C. Cavata, p. 1. DOI: 10.1016/0550-3213(89)90089-8.

[33] B. Adeva et al. "Spin asymmetries A(1) and structure functions g1 of the proton and the deuteron from polarized high-energy muon scattering". In: Phys. Rev. D 58 (1998), p. 112001. DOI: 10.1103/PhysRevD.58.112001.

[34] P. L. Anthony et al. "Deep inelastic scattering of polarized electrons by polarized He-3 and the study of the neutron spin structure". In: Phys. Rev. D 54 (1996), pp. 6620–6650. DOI: 10.1103/PhysRevD.54.6620. arXiv: hep-ex/9610007.

[35] K. Abe et al. "Measurements of the proton and deuteron spin structure functions g(1) and g(2)". In: Phys. Rev. D 58 (1998), p. 112003. DOI: 10.1103/PhysRevD.58.112003. arXiv: hep-ph/9802357.

[36] K. Abe et al. "Precision determination of the neutron spin structure function g1(n)". In: Phys. Rev. Lett. 79 (1997). Ed. by B. Frois, V. W. Hughes, and N. De Groot, pp. 26–30. DOI: 10.1103/PhysRevLett.79.26. arXiv: hep-ex/9705012.

[37] P. L Anthony et al. "Measurements of the Q**2 dependence of the proton and neutron spin structure functions g(1)**p and g(1)**n". In: Phys. Lett. B 493 (2000), pp. 19-28. DOI: 10.1016/S0370-2693(00)01014-5. arXiv: hep-ph/0007248.

[38] D. Flay et al. "Measurements of d_2^n and A_1^n: Probing the neutron spin structure". In: Phys. Rev. D 94.5 (2016), p. 052003. DOI: 10.1103/PhysRevD.94.052003. arXiv: 1603.03612 [nucl-ex].

[39] K. M. Kramer. "The Search for Higher Twist Effects in the Spin-Structure Functions of the Neutron". In: AIP Conf. Proc. 675.1 (2003). Ed. by Y. I. Makdisi, A. U. Luccio, and W. W. MacKay, pp. 615-619. DOI: 10.1063/1.1607208.

[40] X. Zheng et al. "Precision measurement of the neutron spin asymmetries and spin-dependent structure functions in the valence quark region". In: Phys. Rev. C 70 (2004), p. 065207. DOI: 10.1103/PhysRevC.70.065207. arXiv: nucl-ex/0405006.

[41] Y. Prok et al. "Precision measurements of g_1 of the proton and the deuteron with 6 GeV electrons". In: Phys. Rev. C 90.2 (2014), p. 025212. DOI: 10.1103/PhysRevC.90.025212. arXiv: 1404.6231 [nucl-ex].

[42] C. Adolph et al. "Final COMPASS results on the deuteron spin-dependent structure function g_1^d and the Bjorken sum rule". In: Phys. Lett. B 769 (2017), pp. 34-41. DOI: 10.1016/j.physletb.2017.03.018. arXiv: 1612.00620 [hep-ex].

[43] C. Adolph et al. "The spin structure function g_1^p of the proton and a test of the Bjorken sum rule". In: Phys. Lett. B 753 (2016), pp. 18-28. DOI: 10.1016/j.physletb.2015.11.064. arXiv: 1503.08935 [hep-ex].

[44] K. Ackerstaff et al. "Measurement of the neutron spin structure function g1(n) with a polarized He-3 internal target". In: Phys. Lett. B 404 (1997). Ed. by B. Frois, V. W. Hughes, and N. De Groot, pp. 383-389. DOI: 10.1016/S0370-2693(97)00611-4. arXiv: hep-ex/9703005.

[45] A. Airapetian et al. "Precise determination of the spin structure function g(1) of the proton, deuteron and neutron". In: Phys. Rev. D 75 (2007), p. 012007. DOI: 10.1103/PhysRevD.75.012007. arXiv: hep-ex/0609039.

[46] A. Adare et al. "Event Structure and Double Helicity Asymmetry in Jet Production from Polarized $p + p$ Collisions at $\sqrt{s} = 200$-GeV". In: Phys. Rev. D 84 (2011), p. 012006. DOI: 10.1103/PhysRevD.84.012006. arXiv: 1009.4921 [hep-ex].

[47] L. Adamczyk et al. "Longitudinal and transverse spin asymmetries for inclusive jet production at mid-rapidity in polarized $p + p$ collisions at $\sqrt{s} = 200$ GeV". In: Phys. Rev. D 86 (2012), p. 032006. DOI: 10.1103/PhysRevD.86.032006. arXiv: 1205.2735 [nucl-ex].

[48] L. Adamczyk et al. "Precision Measurement of the Longitudinal Double-spin Asymmetry for Inclusive Jet Production in Polarized Proton Collisions at $\sqrt{s} = 200$ GeV". In: Phys. Rev. Lett. 115.9 (2015), p. 092002. DOI: 10.1103/PhysRevLett.115.092002. arXiv: 1405.5134 [hep-ex].

[49] L. Adamczyk et al. "Measurement of the cross section and longitudinal double-spin asymmetry for di-jet production in polarized pp collisions at $\sqrt{s}$ = 200 GeV". In: Phys. Rev. D 95.7 (2017), p. 071103. DOI: 10.1103/PhysRevD.95.071103. arXiv: 1610.06616 [hep-ex].

[50] J. Adam et al. "Longitudinal double-spin asymmetry for inclusive jet and dijet production in pp collisions at $\sqrt{s} = 510$ GeV". In: Phys. Rev. D 100.5 (2019), p. 052005. DOI: 10.1103/PhysRevD.100.052005. arXiv: 1906.02740 [hep-ex].

[51] M. S. Abdallah et al. "Longitudinal double-spin asymmetry for inclusive jet and dijet production in polarized proton collisions at $\sqrt{s} = 510$ GeV". In: Phys. Rev. D 105.9 (2022), p. 092011. DOI: 10.1103/PhysRevD.105.092011. arXiv: 2110.11020 [hep-ex].

[52] M. S. Abdallah et al. "Longitudinal double-spin asymmetry for inclusive jet and dijet production in polarized proton collisions at $\sqrt{s} = 200$ GeV". In: Phys. Rev. D 103.9 (2021), p. L091103. DOI: 10.1103/PhysRevD.103.L091103. arXiv: 2103.05571 [hep-ex].

[53] STAR Collaboration et al. "Measurement of the longitudinal spin asymmetries for weak boson production in proton-proton collisions at $\sqrt{\{s\}}$ = 510 GeV". In: Physical Review D 99 (5 Dec. 2018), p. 52. DOI: 10.1103/PhysRevD.99.051102. URL: http://arxiv.org/abs/1812.04817%20http://dx.doi.org/10.1103/PhysRevD.99.051102.

[54] Daniele Paolo Anderle et al. "Probing gluon helicity with heavy flavor at the Electron-Ion Collider". In: Phys. Rev. D 104.11 (2021), p. 114039. DOI: 10.1103/PhysRevD.104.114039. arXiv: 2110.04489 [hep-ex].

[55] J. Adam et al. "ATHENA detector proposal – a totally hermetic electron nucleus apparatus proposed for IP6 at the Electron-Ion Collider". In: JINST 17.10 (2022), P10019. DOI: 10.1088/1748-0221/17/10/P10019. arXiv: 2210.09048 [physics.ins-det].

[56] Daniel de Florian, Stefano Forte, and Werner Vogelsang. "Higgs production at RHIC and the positivity of the gluon helicity distribution". In: (Jan. 2024). URL: http://arxiv.org/abs/2401.10814.