Eur. Phys. J. C manuscript No.
(will be inserted by the editor)

HERAFitter

Open Source QCD Fit Project

Version 0.91 (svn 1536 - post EB)

S. Alekhin[16,17], O. Behnke[1], P. Belov[1,12], M. Botje[18], D. Britzger[1], S. Camarda[1], A.M. Cooper-Sarkar[2], K. Daum[29,30], C. Diaconu[3], J. Feltesse[19], A. Gizhko[1], A. Glazov[1], A. Guffanti[20], M. Guzzi[1], F. Hautmann[13,14,15], A. Jung[31], H. Jung[1,32], V. Kolesnikov[4], H. Kowalski[1], O. Kuprash[1], A. Kusina[21], S. Levonian[1], K. Lipka[1], B. Lobodzinski[28], K. Lohwasser[16], A. Luszczak[5], B. Malaescu[24], R. McNulty[27], V. Myronenko[1], S. Naumann-Emme[1], K. Nowak[1], F. Olness[21], E. Perez[23], H. Pirumov[1], R. Plačakytė[1], K. Rabbertz[6], V. Radescu[1], R. Sadykov[4], G. Salam[25,26], A. Sapronov[4], A. Schöning[10], T. Schörner-Sadenius[1], S. Shushkevich[1], W. Slominski[7], H. Spiesberger[22], P. Starovoitov[1], M. Sutton[8], J. Tomaszewska[9], O. Turkot[1], A. Vargas[1], G. Watt[11], K. Wichmann[1]

[1]Deutsches Elektronen-Synchrotron (DESY), Hamburg, Germany
[2] Department of Physics, University of Oxford, Oxford, United Kingdom
[3] CPPM, IN2P3-CNRS, Univ. Mediterranee, Marseille, France
[4] Joint Institute for Nuclear Research (JINR), Joliot-Curie 6, 141980, Dubna, Moscow Region, Russia
[5] T. Kosciuszko Cracow University of Technology
[6] Institut für Experimentelle Kernphysik, Karlsruhe, Germany
[7] Jagiellonian University, Institute of Physics, Reymonta 4, PL-30-059 Cracow, Poland
[8] University of Sussex, Department of Physics and Astronomy, Sussex House, Brighton BN1 9RH, United Kingdom
[9] Warsaw University of Technology, Faculty of Physics, Koszykowa 75, 00-662 Warsaw, Poland
[10] Physikalisches Institut, Universität Heidelberg, Heidelberg, Germany
[11] Institute for Particle Physics Phenomenology, Durham University, Durham, DH1 3LE, United Kingdom
[12] Current address: Department of Physics, St. Petersburg State University, Ulyanovskaya 1, 198504 St. Petersburg, Russia
[13] Dept. of Physics and Astronomy, University of Sussex, Brighton BN1 9QH, United Kingdom
[14] Rutherford Appleton Laboratory, Chilton OX11 0QX, United Kingdom
[15] Dept. of Theoretical Physics, University of Oxford, Oxford OX1 3NP, United Kingdom
[16] Deutsches Elektronen-Synchrotron (DESY), Platanenallee 6, D15738 Zeuthen, Germany
[17] Institute for High Energy Physics,142281 Protvino, Moscow region, Russia
[18] Nikhef, Science Park, Amsterdam, the Netherlands
[19] CEA, DSM/Irfu, CE-Saclay, Gif-sur-Yvette, France
[20] Niels Bohr Institute, University of Copenhagen, Denmark
[21] Southern Methodist University, Dallas, Texas
[22] PRISMA Cluster of Excellence, Institut für Physik (WA THEP), Johannes-Gutenberg-Universität, D-55099 Mainz, Germany
[23] CERN, European Organization for Nuclear Research, Geneva, Switzerland
[24] Laboratoire de Physique Nucléaire et de Hautes Energies, UPMC and Université, Paris-Diderot and CNRS/IN2P3, Paris, France
[25] CERN, PH-TH, CH-1211 Geneva 23, Switzerland
[26] LPTHE; CNRS UMR 7589; UPMC Univ. Paris 6; Paris 75252, France
[27] University College Dublin, Dublin 4, Ireland
[28] Max Planck Institut Für Physik, Werner Heisenberg Institut, Föhringer Ring 6, Muńchen
[29] Fachbereich C, Universität Wuppertal, Wuppertal, Germany
[30] Rechenzentrum, Universität Wuppertal, Wuppertal, Germany
[31] FERMILAB, Batavia, IL, 60510, USA
[32] Elementaire Deeltjes Fysica, Universiteit Antwerpen, B 2020 Antwerpen, Belgium

Received: date / Accepted: date

Abstract `HERAFitter` [1] is an open-source package which provides a framework for the determination of the parton distribution functions (PDFs) of the proton and for multifold analyses in Quantum Chromodynamics (QCD).

Measurements of lepton-proton deep inelastic scattering and of proton-proton (proton-antiproton) collisions at hadron colliders are included in the `HERAFitter` package, and are used to probe and constrain the partonic content of the proton.

The partonic distributions are determined by using the factorisation properties of the hadronic cross sections in which short-distance perturbatively calculable partonic scattering cross sections and long-distance contributions that are the non-perturbative universal PDFs, are factorised.

The HERAFitter platform provides a broad choice of options for the treatment of the experimental uncertainties and a common environment where a large number of theoretical calculations and methodological options are used to perform detailed QCD analyses. The general structure of HERAFitter together with available methods are described in this paper.

Keywords PDFs · QCD · Fit · proton structure

Contents

1 Introduction

The constant inflow of new experimental measurements with unprecedented accuracy from hadron colliders is a remarkable challenge for the high energy physics community to provide higher-order theory predictions and to develop efficient tools and methods for data analysis. The recent discovery of the Higgs boson [2, 3] and the extensive searches for signals of new physics in LHC proton-proton collisions demand high-precision computations to test the validity of the Standard Model (SM) and factorisation in Quantum Chromodynamics (QCD). According to the collinear factorisation in perturbative QCD (pQCD) hadronic inclusive cross sections are written as

$$\sigma(\alpha_s(\mu_R),\mu_R,\mu_F) = \sum_{a,b} \int_0^1 dx_1\, dx_2 f_a(x_1,\mu_F) f_b(x_2,\mu_F) \times \hat{\sigma}^{ab}(x_1,x_2;\alpha_s(\mu_R),\mu_R,\mu_F) \quad (1)$$

where the cross section σ for any hard-scattering inclusive process is expressed as a convolution of Parton Distribution Functions (PDFs) f_a and f_b with the partonic cross section $\hat{\sigma}^{ab}$. At Leading-Order (LO), the PDFs represent the probability of finding a specific parton a (b) in the first (second) proton carrying a fraction x_1 (x_2) of its momentum. Indices a and b in the Eq. 1 indicate the various kinds of partons, i.e. gluons, quarks and antiquarks of different flavours, that are considered as the constituents of the proton. The PDFs depend on factorisation scale,μ_F, while the partonic cross sections depend on the strong coupling α_s, and the factorisation and renormalisation scales, μ_F and μ_R. The partonic cross sections $\hat{\sigma}^{ab}$ are calculated in pQCD whereas PDFs are constrained by global fits to variety of the hard-process experimental data employing universality of PDFs within a particular factorisation scheme [4, 5].

Measurements of the inclusive Neutral Current (NC) and Charged Current (CC) Deep Inelastic Scattering (DIS) at the ep collider HERA provide crucial information for determining the PDFs. The gluon density in small and medium x can be accurately determined solely from the HERA data. Many processes in pp and $p\bar{p}$ collisions at LHC and Tevatron, respectively, probe PDFs in the kinematic ranges, complementary to the DIS measurements. Therefore inclusion of the LHC and Tevatron data in the QCD analysis of the proton structure provide additional constraints on the PDFs, improving either their precision, or providing valuable information on the correlations of PDFs with the fundamental QCD parameters like the strong coupling or the quark masses. In this context, the processes of interest at hadron colliders are Drell-Yan (DY) production, W-boson asymmetries, associated production of W or Z bosons and heavy quarks, top quark, jet and prompt photon production.

This paper describes the open-source QCD fit platform HERAFitter which includes the set of tools essential for a comprehensive global QCD analysis of pp, $p\bar{p}$ and ep scattering processes of the experimental measurement. It is developed for determination of PDFs and extraction of fundamental QCD parameters such as the heavy quark masses and the strong coupling constant. This platform also provides the basis for comparisons of different theoretical approaches and can be used for direct tests of the impact of new experimental data on the SM parameters in the QCD analyses.

This paper is organised as follows. The structure and overview of `HERAFitter` are presented in section 2. Section 3 discusses the various processes and corresponding theoretical calculations performed in the collinear factorisation using the DGLAP [6–10] formalism, available in `HERAFitter`. Section 4 presents various fast techniques employed by the theory calculations used in `HERAFitter`. Section 5 elucidates the methodology of determining PDFs through fits based on various χ^2 definitions used in the minimisation procedure. Alternative approaches to the DGLAP formalism are presented in section 6. Specific applications of the package are given in section 8 and the summary is presented in section 9.

Data	Process	Reaction	Theory calculations, schemes
HERA Fixed Target	DIS NC	$ep \to eX$	TR′, ACOT ZM (`QCDNUM`) FFN (`OPENQCDRAD`, `QCDNUM`), TMD (uPDFevolv)
HERA	DIS CC	$ep \to \nu_e X$	ACOT, ZM (`QCDNUM`) FFN (`OPENQCDRAD`)
	DIS jets	$ep \to e$ jets$X$	`NLOJet++` (`fastNLO`)
	DIS heavy quarks	$ep \to ec\bar{c}X$, $ep \to eb\bar{b}X$	ZM (`QCDNUM`), TR′, ACOT, FFN (`OPENQCDRAD`, `QCDNUM`)
Tevatron LHC	Drell-Yan	$pp(\bar{p}) \to l\bar{l}X$, $pp(\bar{p}) \to l\nu X$	`MCFM` (`APPLGRID`)
	top pair	$pp(\bar{p}) \to t\bar{t}X$	`MCFM` (`APPLGRID`), `HATHOR`
	single top	$pp(\bar{p}) \to tl\nu X$, $pp(\bar{p}) \to tX$, $pp(\bar{p}) \to tWX$	`MCFM` (`APPLGRID`)
	jets	$pp(\bar{p}) \to$ jets$X$	`NLOJet++` (`APPLGRID`), `NLOJet++` (`fastNLO`)
LHC	DY+heavy quarks	$pp \to VhX$	`MCFM` (`APPLGRID`)

Table 1 The list of processes implemented in the `HERAFitter` package. The references for the individual calculations and their implementations are given in the text.

2 The HERAFitter Structure

`HERAFitter` is a flexible open-source platform for the QCD analyses of different experimental measurements, providing a versatile environment for benchmarking studies. It is widely used within the LHC experiments [11–16].

The functionality of `HERAFitter` is schematically illustrated in Fig. 1 and it can be divided into four main blocks:

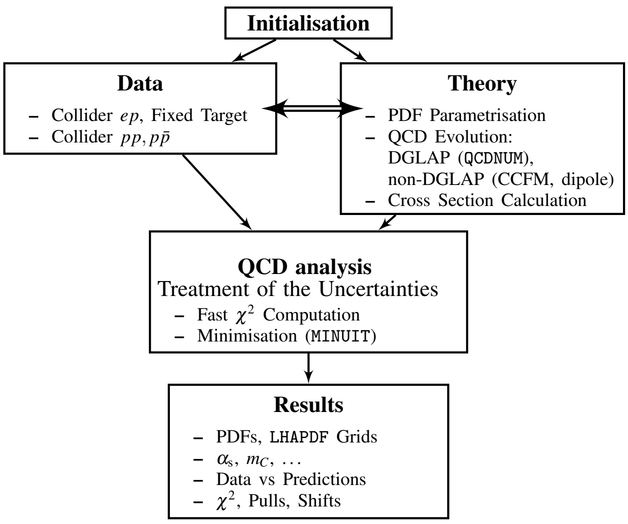

Fig. 1 Schematic structure of the `HERAFitter` program.

Data: Different available measurements from various processes are implemented in the `HERAFitter` package including the full information on their uncorrelated and correlated uncertainties. HERA data are sensitive to light quark and gluon densities mostly through scaling violations, covering low and medium x ranges. These data are the basis of any proton PDF extraction, and are used by all global PDF groups [17–21]. However, improvements in precision of PDFs require additional constraints on the gluon and quark distributions at high-x, better understanding of heavy quark distributions and decomposition of the light-quark sea. For these purposes, the measurements of the fixed-target experiments, Tevatron and LHC are of particular importance. The processes that are currently available in `HERAFitter` framework are listed in Tab. 1.

Theory: Predictions for cross section of different processes are obtained using the factorisation approach (Eq. 1). The PDFs are parametrised at a starting input scale Q_0^2 by a chosen functional form with a set of free parameters $\mathbf{p}$. These PDFs are evolved to the scale of the measurement Q^2, $Q^2 > Q_0^2$. The evolution follows either DGLAP [6–10] (as implemented in `QCDNUM` [22]), CCFM [23–26] (as implemented in `uPDFevolv` [27]). The prediction of a particular process cross section is obtained by a convolution of the evolved PDFs and the partonic cross section, calculated at a certain order in QCD with a appropriate theory calculation (as listed in Tab. 1). Alternatively, predictions using dipole models [28–30] can be also obtained.

QCD analysis: The PDFs are determined by the least square fit, minimising the χ^2 function, formed using the input data and theory predictions, with the `MINUIT` [31] program. Various choices of accounting for the experimental uncertainties are employed in `HERAFitter`, either using a nuisance parameter method for the correlated systematic uncertainties, or a covariance matrix method as described in section 5.2). In addition, `HERAFitter` allows to study different statistics assumptions for the distributions of the systematic uncertainties, like Gauss, LogNormal [32] (see section 5.3).

Results: The resulting PDFs are provided in a format ready to be used by the `LHAPDF` library [33, 34] (or by `TMDlib` [35]). `HERAFitter` drawing tools can be used to display the PDFs with their uncertainties at a chosen scale. As an example, a first set of PDFs extracted using `HERAFitter` from HERA I data, HERAPDF1.0 [36], is shown in Fig. 2.

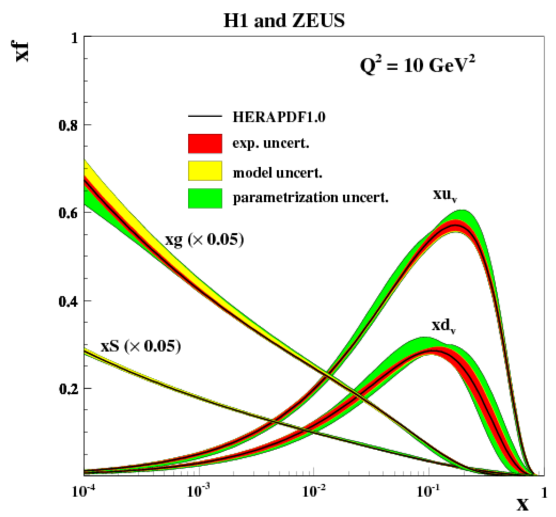

Fig. 2 Distributions of valence (xu_v, xd_v), sea (xS) and the gluon (g) densities in HERAPDF1.0 [36]. The gluon and the sea distributions are scaled down by a factor of 20. The experimental, model and parametrisation uncertainties are shown as colored bands.

3 Theoretical formalism using DGLAP evolution

In this section the theoretical formalism for various processes available in `HERAFitter` is described.

A direct consequence of factorisation is that scale dependence or "evolution" of PDFs can be predicted by the renormalisation group equations (RGE's). By imposing that physical observables are independent on $\mu_{\rm F}$, it leads to a representation of parton evolution in terms of integro-differential equations known as DGLAP [6–10] (Dokshitzer-Gribov-Lipatov-Altarelli-Parisi) equations:

$$\frac{d\, f_i(x,\mu_{\rm R},\mu_{\rm F})}{d\log\mu_{\rm F}^2} = \sum_{j=q\bar{q},g}\int_x^1 \frac{dy}{y} P_{ij}\left(\frac{x}{y};\alpha_s,\mu_{\rm R},\mu_{\rm F}\right) f_j(y,\mu_{\rm R},\mu_{\rm F}), \quad (2)$$

where the functions P_{ij} are the evolution kernels or splitting functions, which represent the probability of finding parton i in parton j, and have perturbative expansion in α_s. The analytic structure of P_{ij} is known at 3-loop in perturbation theory and can be found in the literature.

Therefore, once PDFs are determined by a direct comparison with the experiments at the initial scale Q_0, their evolution at the scale μ is entirely determined by DGLAP equations. Alternative approaches to DGLAP evolution, valid in different kinematic regimes, are also implemented in `HERAFitter` and will be discussed in the next sections.

3.1 Deep Inelastic Scattering and Proton Structure

DIS data provide the backbone of any PDF fit. The formalism that relates the DIS measurements to pQCD and the PDFs has been described in detail in many extensive reviews (see e.g. [37]) and it is only briefly summarised here. DIS is the process where a lepton scatters off the constituents of the proton by a virtual exchange of a NC or CC vector boson and, as a result, a scattered lepton and a multi-hadronic final state are produced. The common DIS kinematic variables are the absolute squared four-momentum of the exchange boson, Q^2, the Bjorken x, and the inelasticity y, related by $y = Q^2/sx$, where s is the squared centre-of-mass (c.o.m.) energy.

The NC cross section can be expressed in terms of generalised structure functions:

$$\frac{d^2\sigma_{NC}^{e^\pm p}}{dxdQ^2} = \frac{2\pi\alpha^2}{xQ^4}\cdot\sigma_{r,NC}^{e^\pm p}, \quad (3)$$

$$\sigma_{r,NC}^{e^\pm p} = Y_+\tilde{F}_2^\pm \mp Y_- x\tilde{F}_3^\pm - y^2\tilde{F}_L^\pm, \quad (4)$$

where the electromagnetic coupling constant α, the photon propagator and a helicity factor are absorbed in the definition of reduced cross section σ_r, and $Y_\pm = 1\pm(1-y)^2$ (additional terms of $O(1/Q^2)$ are numerically small at the HERA kinematics and are neglected). The generalised structure functions $\tilde{F}_{2,3}$ can be written as linear combinations of the proton structure functions $F_2^\gamma, F_{2,3}^{\gamma Z}$ and $F_{2,3}^Z$ associated to pure photon exchange terms, photon-Z interference terms and pure Z exchange terms, respectively. The structure function $\tilde{F}_2$ is the dominant contribution to the cross section, $x\tilde{F}_3$ becomes important at high Q^2 and $\tilde{F}_L$ is sizable only at high y.

The inclusive CC ep cross section, analogous to the NC case, can be expressed in terms of another set of structure functions and in LO in α_s, the e^+p and e^-p cross sections are sensitive to different combinations of the quark flavour densities:

$$\frac{d^2\sigma_{CC}^{e^{\pm}p}}{dxdQ^2} = \frac{1 \pm P}{2}\frac{G_F^2}{2\pi x}\left[\frac{M_W^2}{M_W^2+Q^2}\right]\cdot \sigma_{r,CC}^{e^{\pm}p} \quad (5)$$

$$\sigma_{r,CC}^{e^{\pm}p} = Y_+\tilde{W}_2^{\pm} \mp Y_- x\tilde{W}_3^{\pm} - y^2\tilde{W}_L^{\pm}, \quad (6)$$

where P represents the lepton beam polarisation and $\tilde{W}_2$, $\tilde{W}_3$,$\tilde{W}_L$ are structure functions analogous to the above NC case. The QCD predictions for the DIS structure functions are obtained by convoluting the PDFs with the respective coefficient functions.

The DIS measurements span a large range of Q^2 from few GeV^2 to about 10^5 GeV^2, crossing heavy-quark mass thresholds, thus the treatment of heavy quarks (charm and beauty) and of their masses becomes important. There are different approaches to the treatment of heavy quark production that should be equivalent if calculations are carried out to all orders in α_s. Several variants of these schemes are implemented in `HERAFitter` and they are briefly discussed below.

Zero-Mass Variable Flavour Number (ZM-VFN)[38]: In this scheme, the heavy quark densities appear in the proton at Q^2 values above $\sim m_h^2$ (heavy quark mass) and the heavy quarks are treated as massless in both the initial and final states of the hard scattering process. The lowest order process is the scattering of lepton off the heavy quark via boson exchange. This scheme is expected to be reliable only in the region with $Q^2 \gg m_h^2$. In `HERAFitter` this scheme is available for the DIS structure function calculation via the interface to the `QCDNUM` [22] package and it benefits from the fast `QCDNUM` convolution engine.

Fixed Flavour Number (FFN)[39–41]: In this scheme only the gluon and the light quarks are considered as partons within the proton and massive quarks are produced perturbatively in the final state. The lowest order process is the heavy quark-antiquark pair production in the boson-gluon fusion. In `HERAFitter` this scheme can be accessed via the `QCDNUM` implementation or through the interface to the open-source code `OPENQCDRAD` [42], as implemented by the ABM group. Through `QCDNUM`, the calculation of the heavy quark contributions to DIS structure functions are available at Next-to-Leading-Order (NLO), at $O(\alpha_s^2)$, and only electromagnetic exchange contributions are taken into account. Through the ABM implementation the heavy quark contributions to CC structure functions are available and, for the NC case, the QCD corrections to the coefficient functions at Next-to-Next-to Leading Order (NNLO) are provided at the best currently known approximation [43]. The ABM implementation also includes the definition in $\overline{\text{MS}}$ scheme with the running heavy-quark mass [44]. The scheme has the advantage of reducing the sensitivity of the DIS cross sections to higher order corrections, and improving the theoretical precision of the mass definition.

General-Mass Variable Flavour Number (GM-VFN)[45]: It this scheme, heavy quark production is treated for $Q^2 \leq m_h^2$ in the FFN scheme and for $Q^2 \gg m_h^2$ in a masless scheme. The recent series of PDF groups that use this scheme are MSTW, CT(CTEQ), NNPDF, and HERAPDF. `HERAFitter` implements different variants of the GM-VFN scheme and they are presented below:

- *GM-VFN Thorne-Roberts scheme:* The Thorne-Roberts (TR) scheme [46] was designed to provide a smooth transition from the massive FFN scheme at low scales $Q^2 < m_h^2$ to the massless ZM-VFNS scheme at high scales $Q^2 \gg m_h^2$. However, the original version was technically difficult to implement beyond NLO, and was updated to the TR' scheme [47]. There are two different variants of the TR' schemes: TR' standard (as used in MSTW PDF sets [17, 47]) and TR' optimal [48], with a smoother transition across the heavy quark threshold region. Both variants are accessible within the `HERAFitter` package at LO, NLO and NNLO.
- *GM-VFN ACOT scheme:* The Aivazis-Collins-Olness-Tung (ACOT) scheme belongs to the group of VFN factorisation schemes that use the renormalisation method of Collins-Wilczek-Zee (CWZ) [49]. This scheme unifies the low scale $Q^2 < m_h^2$ and high scale $Q^2 > m_h^2$ regions with a smooth interpolation across the full energy range. Within the `ACOT`package, different variants of the ACOT scheme are available: ACOT-Full [50], S-ACOT-χ [51, 52], ACOT-ZM [50], $\overline{\text{MS}}$ at LO and NLO. For the longitudinal structure function higher order calculations are also available. A compasion of PDFs extracted from the QCD fits to the HERA data with the TR' and ACOT-Full schemes is illustrated in Fig. 3.

3.2 Electroweak Corrections to DIS

Calculations of higher-order electroweak corrections to DIS scattering at HERA are available in `HERAFitter` in the on-shell scheme. In this scheme the gauge bosons masses M_W and M_Z are treated symmetrically as basic parameters together with the top, Higgs and fermion masses. These electroweak corrections are based on the `EPRC`package [53]. The code provides the running of electromagnetic coupling α using the most recent parametrisation of the hadronic contribution to Δ_α [54], as well as an older version from Burkhard [55].

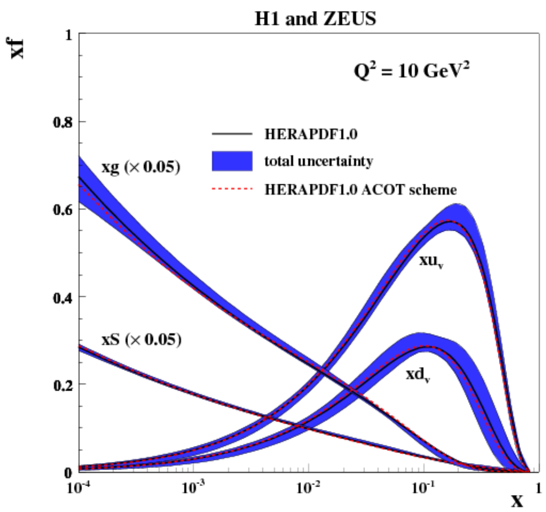

Fig. 3 Overview showing the u- and d-valence, the total sea (scaled), and gluon (scaled) PDFs of the NLO HERAPDF1.0 set [36] with their total uncertainty at the scale of $Q^2 = 10$ GeV^2 obtained using the TR' scheme and compared to the PDFs obtained with the ACOT scheme using the k-factor technique (red).

3.3 Diffractive PDFs

Similarly to standard PDFs, diffractive parton distributions (DPDFs) can be determined from QCD fits to diffractive cross sections. About 10% of deep inelastic interactions at HERA are diffractive, i.e. leading to events in which the interacting proton stays intact ($ep \to eXp$). In the diffractive process the proton is well separated from the rest of the hadronic final state by a large rapidity gap. This is interpreted as the dissociation of the virtual photon into hadronic system X with the invariant mass much smaller than the photon-proton c.o.m. energy $W = ys - Q^2 + m_p^2(1-y)$, where m_p is proton's mass, and the same net quantum numbers as the exchanged photon. For such a processes, the diffractive DIS is mediated by the exchange of a hard Pomeron or a secondary Reggeon with the vacuum quantum numbers. The factorisable pomeron picture has proved remarkably successful in the description of most of these data.

The kinematic variables squared four-momentum transfer t (the undetected momentum transfer to the proton system) and the mass M_X of the diffractively produced final state appear for the diffrative process in addition to the usual DIS variables x, Q^2. In practice, the variable M_X is often replaced by dimensionless quantity $\beta = \frac{Q^2}{M_X^2+Q^2-t}$. In models based on a factorisable pomeron, β may be viewed at LO as the fraction of the pomeron longitudinal momentum which is carried by the struck parton, $x = \beta x_{IP}$.
For the inclusive case, the diffractive cross-section reads as:

$$\frac{d\sigma}{d\beta\, dQ^2 dx_{IP}\, dt} = \frac{2\pi\alpha^2}{\beta Q^4}\left(1+(1-y)^2\right)\overline{\sigma}^{D(4)}(\beta, Q^2, x_{IP}, t) \qquad (7)$$

with the "reduced cross-section":

$$\overline{\sigma}^{D(4)} = F_2^{D(4)} - \frac{y^2}{1+(1-y)^2} F_L^{D(4)}. \qquad (8)$$

Substituting $x = x_{IP}\beta$ we can relate Eq. 7 to the standard DIS formula. In this way, the diffractive structure functions can be expressed as convolutions of the calculable coefficient functions with the diffractive quark and gluon distribution functions, which in general depend on x_{IP}, Q^2, β, t.

The diffractive PDFs in `HERAFitter` [56, 57] are implemented as a sum of two factorised contributions:

$$\Phi_{IP}(x_{IP}, t)\, f_a^{Pom}(\beta, Q^2) + \Phi_{IR}(x_{IP}, t)\, f_a^{IR}(\beta, Q^2), \qquad (9)$$

where $\Phi(x_{IP}, t)$ are the Regge type fluxes. The Reggeon PDFs, f_a^{IR} are taken as those of the pion, while the Pomeron ones, f_a^{IP}, are obtained from a fit to the data.

3.4 Drell-Yan Processes in pp or $p\bar{p}$ Collisions

Drell-Yan process provides further valuable information about PDFs. In pp and $p\bar{p}$ scattering, the Z/γ^* and W production probe bi-linear combinations of quarks. Complementary information on the different quark densities can be obtained from the W asymmetry (d, u and their ratio), the ratio of the W and Z cross sections (sensitive to the flavour composition of the quark sea, in particular to the s density), and associated W and Z production with heavy quarks (sensitive to s- and c-quark densities). Measurements at large boson $p_T \gtrsim M_{W,Z}$ are potentially sensitive to the gluon density [58].

The LO DY for NC triple differential cross section in invariant mass M, boson rapidity y and lepton scattering angle $\cos\theta$ in the parton c.o.m. frame can be written as [59, 60]:

$$\frac{d^3\sigma}{dMdy\, d\cos\theta} = \frac{\pi\alpha^2}{3MS}\sum_q P_q\left[f_q(x_1, Q^2) f_{\bar{q}}(x_2, Q^2) + (q \leftrightarrow \bar{q})\right], \qquad (10)$$

where S is the squared c.o.m. beam energy, $x_{1,2} = \frac{M}{\sqrt{S}}\exp(\pm y)$, $f_q(x_1, Q^2)$ are the quark distribution functions, and P_q is a partonic cross section.

The LO expression for CC scattering has a form:

$$\frac{d^3\sigma}{dMdy\, d\cos\theta} = \frac{\pi\alpha^2}{48S\sin^4\theta_W}\frac{M^3(1-\cos\theta)^2}{(M^2 - M_W^2) + \Gamma_W^2 M_W^2} \sum_{q_1, q_2} V_{q_1 q_2}^2 f_{q_1}(x_1, Q^2) f_{q_2}(x_2, Q^2), \qquad (11)$$

where $V_{q_1 q_2}$ is the Cabibbo-Kobayashi-Maskawa (CKM) quark mixing matrix and M_W and Γ_W are the W boson mass and decay width, respectively.

The simple form of these expressions allows the calculation of integrated cross sections without the use of Monte

Carlo (MC) techniques which often introduce statistical fluctuations. In both NC and CC expressions the PDFs depend only on boson rapidity y and invariant mass M, while the integral in $\cos\theta$ can be solved analytically including the case of realistic kinematic cuts.

Currently, the predictions for W and Z/γ^* production are available to NNLO and W, Z in association with heavy flavour quarks to NLO. There are several possibilities for obtaining the theoretical predictions for DY production in HERAFitter.

The NLO and NNLO calculations are computing power and time consuming and k-factor or fast grid techniques must be employed (see section 4 for details), interfaced to programs such as MCFM [61–63], available for NLO calculations, or FEWZ [64] and DYNNLO [65] for NLO and NNLO.

3.5 Jet Production in ep and pp or $p\bar{p}$ Collisions

Cross section for production of the high-transverse-momentum hadronic jets is sensitive to the high-x gluon PDF (see e.g. [17]) therefore this process can be used to improve determination of the gluon PDF, which is particularly important for the Higgs production and searches for new physics. Jet production cross sections are currently only known to NLO, although calculations for higher-order contributions to jet production in proton-proton collisions are now quite advanced [66–68]. Within HERAFitter, the NLOJet++ program [69, 70] may be used for the calculations of jet production. Similarly to the DY case, the calculation is very demanding in terms of computing power. Therefore fast grid techniques are used to facilitate the QCD analyses including jet cross section measurements. in ep, pp and $p\bar{p}$ collisions (for details see section 4).

3.6 Top-quark Production in pp or $p\bar{p}$ Collisions

Top-quark pairs ($t\bar{t}$) are produced at hadron colliders dominantly via gg fusion (at the LHC) and $q\bar{q}$ annihilation (at the Tevatron). Measurements of the $t\bar{t}$ cross sections provide additional constraints in particular on the gluon density at medium to high values of x, on α_s and on the top-quark mass, m_t [71]. Precise predictions for the total $t\bar{t}$ cross section are available to full NNLO [72]. They can be computed within HERAFitter via an interface to the program HATHOR [73]. Differential $t\bar{t}$ cross section predictions can be used with MCFM [63, 74–77] at NLO accuracy interfaced to HERAFitter with fast grid techniques.

Single top quarks are produced via electroweak interactions and single-top cross sections can be used, for example, to probe the ratio of the u and d densities in the proton as well as the b-quark PDF. Predictions for single-top production are available only at NLO accuracy using MCFM.

4 Computational Techniques

Precise measurements require theoretical predictions with equally good accuracy in order to maximise their impact in PDF fits. Perturbative calculations, however, get more and more involved with order due to an increasing number of Feynman diagrams. Nowadays even the most advanced perturbative techniques in combination with modern computing hardware do not lead to sufficiently small turn-around times. The direct inclusion of computationally demanding higher-order calculations into iterative fits therefore is not possible. Relying on the fact that a full repetition of the perturbative calculation for arbitrary changes in input parameters is not necessary at each iteration step, two methods have been developed to resolve this problem: the techniques of k-factors and *fast grids*. Both are available in HERAFitter and described as follows.

4.1 k-factor Technique

The k-factors are defined as the ratio of the prediction of a higher-order (slow) pQCD calculation to a lower-order (fast) calculation. Because the k-factors depend on the phase space probed by the measurement, they have to be stored in a table including dependence on the relevant kinematic variables. Before the start of a fitting procedure, the table of k-factors has to be computed once for a given PDF with the time consuming higher-order code. In subsequent iteration steps the theory prediction is derived from the fast lower-order calculation multiplied by the pre-tabulated k-factors.

This procedure, however, neglects the fact that the k-factors can be PDF dependent, as a consequence, they have to be re-evaluated for the newly determined PDF at the end of the fit for the consistency check. Usually, the fit is repeated until input and output k-factors have converged. In summary, this technique avoids iteration of the higher-order calculation at each step, but still requires a couple of repetitions depending on the analysis.

An implementation of k-factor technique in HERAFitter is used for the fast approximation of the time-consuming GM-VFN schemes for heavy quarks in DIS. 'FAST" heavy-flavour schemes are implemented with k-factors defined as the ratio of calculations at the same perturbative order but for massive vs. massless quarks, e.g. NLO (massive)/NLO (massless). These k-factors are calculated only for the starting PDF and hence, the "FAST" heavy flavour schemes should only be used for quick checks, i.e. full heavy flavour schemes are normally recommended. For the ACOT case, due to long computation time, the k-factors are used in the default settings in HERAFitter.

4.2 Fast Grid Techniques

Fast grid techniques exploit the factorisable nature of the cross sections and the fact that iterative PDF fitting procedures do not impose completely arbitrary change in the shape of the parameterised functions that represent each PDF. Instead, it can be assumed that a generic PDF can be approximated by a set of interpolating functions with a sufficient number of support points. The accuracy of this approximation, can be checked and optimised in various ways with the simplest one being an increase in the number of support points. Having ensured that the approximation bias is negligibly small compared to the experimental and theoretical accuracy for all practical purposes, this method can be used to perform the time consuming higher-order calculations (Eq. 1) only once for the set of interpolating functions. Further iteration of a cross section evaluation for a particular PDF set is fast and implies only sums over the set of interpolators multiplied by factors depending on the PDF. The approach applies equally for the cross sections of processes involving one or two hadrons in the initial state as well as to their renormalisation and factorisation scale variation.

This technique was pioneered in the `fastNLO` project [78] to facilitate the inclusion of notoriously time consuming jet cross sections at NLO into PDF fits. The `APPLGRID` [79] project developed an alternative method and, in addition to jets, extended its applicability to other scattering processes, such as DY, heavy quark pair production is association with boson production, etc. While differing in their interpolation and optimisation strategies, both packages construct tables with grids for each bin of an observable in two steps: In the first step, the accessible phase space in the parton momentum fractions x and the renormalisation and factorisation scales μ_{R} and μ_{F} is explored in order to optimize the table size. The second step consists of the actual grid filling for the requested observables. Higher-order cross sections can then be restored very efficiently from the pre-produced grids while varying externally provided PDF sets, μ_{R} and μ_{F}, or the strong coupling $\alpha_s(\mu_R)$. The approach can in principle be extended to arbitrary processes, but requires to establish an interface between the higher-order theory programs and the fast interpolation frameworks. Work in that direction is ongoing for both packages and described in more details in the following:

- The `fastNLO` project [78] has been interfaced to the `NLOJet++` program [69] for the calculation of jet production in DIS [80] as well as 2- and 3-jet production in hadron-hadron collisions at NLO [70, 81]. To demonstrate the applicability to higher-orders, threshold corrections at 2-loop order, which approximate the NNLO for the inclusive jet cross section, have been included into the framework as well [82] following Ref. [83].
  The latest version of `fastNLO` convolution program [84] allows for a creation of tables where renormalisation and factorisation scales can be varied as a function of two pre-defined observables, e.g. jet transverse momentum $p_\perp$ and Q for DIS. The `fastNLO` code is available online and the jet cross-section grids computed for kinematics of various experiments can be downloaded as well [85]. Dedicated `fastNLO` libraries and tables with theory predictions for comparison to particular cross section measurements are included into the `HERAFitter` package. For the `HERAFitter` implementation, the evaluation of the strong coupling constant is taken consistently with the PDF evolution from the `QCDNUM` code.
- In the `APPLGRID` package [79, 86], in addition to the jet cross sections from `NLOJet++` in $pp(\bar{p})$ and DIS processes, the calculations of DY production are alsoimplemented. The look-up tables (grids) can be generated with the customised versions of the `MCFM` parton level DY generator [61–63]. The variation of the renormalisation and factorisation scales is possible a posteriori, when calculating theory predictions with the `APPLGRID` tables, and independent variation of the strong coupling constant is also allowed. For NNLO predictions in `HERAFitter`, the k-factors technique can be also applied within the `APPLGRID` framework.
  The `HERAFitter` interface to `APPLGRID` was in particular used by the ATLAS collaboration to extract the strange quark density of the proton from W and Z cross sections [11]. An illustration of ATLAS PDFs extracted employing these techniques is displayed in Fig. 4 together with the comparison to global PDF sets CT10 [18] and NNPDF2.1 [19].

5 Fit Methodology

Performing a QCD analysis one usually needs to check stability of the results w.r.t. different assumptions, e.g. the functional parametrisation form, the heavy quarks mass values, alternative theoretical calculations, method of minimisation, interpretation of uncertainties, etc. It is also desirable to be able to discriminate or quantify the effect of the chosen ansatz, ideally within a common framework, and `HERAFitter` is optimally designed for such tests. The methodology employed by `HERAFitter` relies on a flexible and modular framework that allows for independent integration of the state-of-the-art techniques, either related to the inclusion of a new theoretical calculation, or of new approaches to treat uncertainties.

In this section we describe the available options in `HERAFitter`. In addition, as an alternative approach to a complete QCD fit, the Bayesian reweighting method, which is also available in `HERAFitter`, is described.

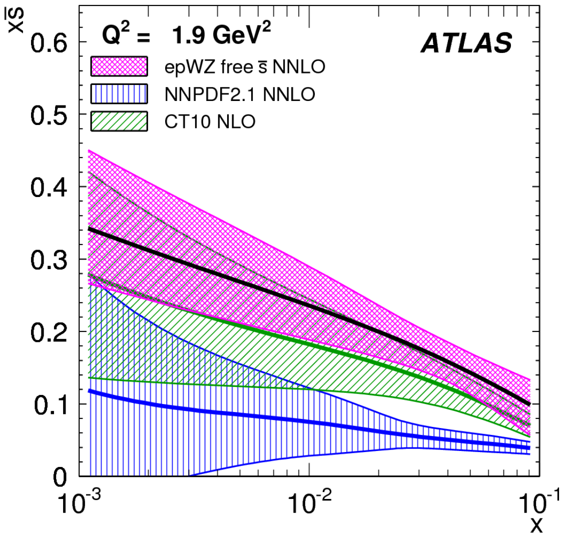

Fig. 4 The strange anti-quark density versus x for the ATLAS epWZ free sbar NNLO fit [11] (magenta band) compared to predictions from NNPDF2.1 (blue hatched) and CT10 (green hatched) at $Q^2 = 1.9$ GeV^2. The ATLAS fit was performed using a k-factor approach for NNLO corrections.

5.1 Functional Forms for PDF Parametrisation

The PDFs are parametrised using several predefined functional form and different flavour decomposition. In `HERAFitter`, various functional forms to parametrise PDFs can be used:

Standard Polynomials: The standard polynomial form is most commonly used by the PDF groups. A polynomial functional form is used to parametrise the x-dependence of the PDFs, where index j denotes each parametrised PDF:

$$xf_j(x) = A_j x^{B_j} (1-x)^{C_j} P_i(x), \tag{12}$$

The parametrised PDFs are the valence distributions xu_v and xd_v, the gluon distribution xg, and the u-type and d-type sea as constrained by HERA data alone, $x\bar{U}$, $x\bar{D}$, where $x\bar{U} = x\bar{u}$, $x\bar{D} = x\bar{d} + x\bar{s}$ at the starting scale. The form of polynomials $P_i(x)$ depdend on the style, defined as a steering parameter. The form $(1+\varepsilon_j\sqrt{x}+D_j x+E_j x^2)$ is used for the HERAPDF [36] style with additional constraints relating to the flavour decomposition of the light sea. For the CTEQ style, $P_i(x)$ takes the form $e^{a_3 x}(1+e^{a_4}x+e^{a_5}x^2)$ and, in contrast to polynomial form, is positive by constraction. QCD number and momentum sum rules are used to determine the normalisations A for the valence and gluon distributions, and the sum rule integrals are solved analytically.

Bi-Log-Normal Distributions: The parametrisation is motivated by multi-particle statistics and holds the following functional form:

$$xf_j(x) = a_j x^{p_j - b_j \log(x)} (1-x)^{q_j - d_j \log(1-x)}. \tag{13}$$

This function can be regarded as a generalisation of the standard polynomial form described above, however, numerical integration of Eq. 13 is required in order to satisfy the QCD sum rules.

Chebyshev Polynomials: A flexible parameterization employed for the gluon and sea distributions and based on the Chebyshev polynomials. For better modeling the low-x asymptotic of those PDFs, the polynomial of the argument $\log(x)$ are considered. Furthermore, the PDFs are multiplied by the factor of $(1-x)$ to ensure that they vanish as $x \to 1$. The resulting parametric form reads

$$xg(x) = A_g (1-x) \sum_{i=0}^{N_g-1} A_{g_i} T_i \left(-\frac{2\log x - \log x_{\min}}{\log x_{\min}} \right), \tag{14}$$

$$xS(x) = (1-x) \sum_{i=0}^{N_S-1} A_{S_i} T_i \left(-\frac{2\log x - \log x_{\min}}{\log x_{\min}} \right), \tag{15}$$

where T_i are the first-type Chebyshev polynomials of the order i. The normalisation factor A_g is defined from the momentum sum rule which can be evaluated analytically. The values of $N_{g,S}$ up to 15 are allowed, however, already starting from $N_{g,S} \geq 5$ the fit quality is already similar to the standard-polynomial parametrisation with a similar number of parameters.

The low-x uncertainties in the PDFs determined from the HERA data using different parameterizatons were studied in Ref. [87]. Figure 5 shows the comparison of the gluon density obtained with the parameterization Eq. 14,15 to the standard-polynomial one, for $N_{g,S} = 9$.

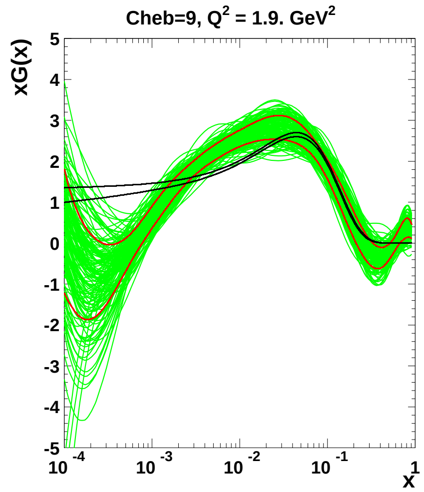

Fig. 5 The gluon density is shown at the starting scale. The black lines correspond to the uncertainty band of the gluon distribution using a standard parameterisation and it is compared to the case of the Chebyshev parameterisation [87]. The uncertainty band for the latter case is estimated using the Monte Carlo technique with the green lines denoting fits to data replica. aRed lines indicate the standard deviation about the mean value of these replicas.

External PDFs: HERAFitter provides the possibility to access external PDF sets, which can be used to compute theoretical predictions for the various processes of interest as implemented in HERAFitter. This is possible via an interface to LHAPDF [33, 34] providing access to the global PDF sets. HERAFitter also allows to evolve PDFs from LHAPDF with QCDNUM using the corresponding grids as a starting scale. Figure 6 illustrates the comparison of the PDFs accessed from LHAPDF as produced with the drawing tools available in HERAFitter.

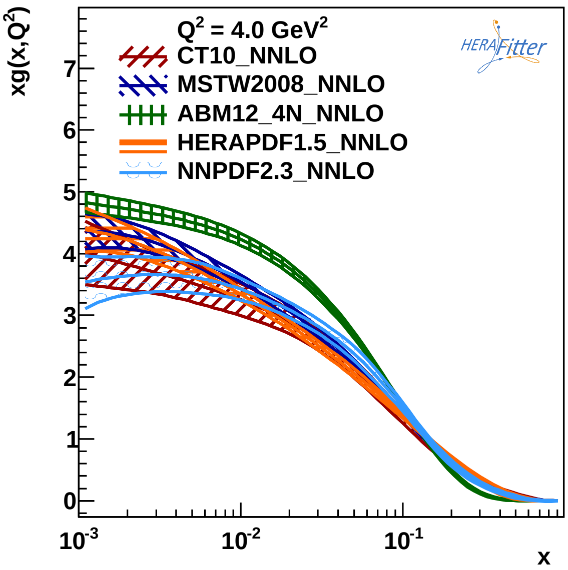

Fig. 6 Gluon density as extracted by various PDF groups at the scale of $Q^2 = 4$ GeV^2, plotted using the drawing tools from HERAFitter.

5.2 Representation of χ^2

The PDF parameters are determined in HERAFitter by minimisation of the χ^2 function taking into account correlated and uncorrelated measurement uncertainties. There are various forms of χ^2 differing by method used to include the experimental uncertainties, e.g. using covariance matrix or providing nuisance parameters to encode dependence of each systematic source for each measurement data point, different scaling options, etc. The options available in HERAFitter are following.

Covariance Matrix Representation: For a data point μ_i with a corresponding theory prediction m_i, the χ^2 function can be expressed in the following form:

$$\chi^2(m) = \sum_{i,k} (m_i - \mu_i) C^{-1}_{ik} (m_k - \mu_k), \tag{16}$$

where the experimental uncertainties are given in a form of covariance matrix $C_{i,k}$ for measurements in bins i an k. The covariance matrix C_{ik} is given by the sum of statistical, uncorrelated and correlated systematic contributions:

$$C_{ik} = C^{stat}_{ik} + C^{uncor}_{ik} + C^{sys}_{ik}. \tag{17}$$

With this representation the effect of a certain systematic source of the uncertainty cannot be distinguished from others.

Nuisance Parameters Representation: For the case when systematic uncertainties are separated by sources the χ^2 form is expressed as

$$\chi^2(m,b) = \sum_i \frac{\left[\mu_i - m_i\left(1 - \sum_j \gamma^i_j b_j\right)\right]^2}{\delta^2_{i,\mathrm{unc}} m_i^2 + \delta^2_{i,\mathrm{stat}}\, \mu_i m_i \left(1 - \sum_j \gamma^i_j b_j\right)} + \sum_j b_j^2, \tag{18}$$

where, $\delta_{i,\mathrm{stat}}$ and $\delta_{i,\mathrm{unc}}$ are relative statistical and uncorrelated systematic uncertainties of the measurement i. Further, γ^i_j quantifies the sensitivity of the measurement to the correlated systematic source j. The function χ^2 depends in addition on the set of systematic nuisance parameters b_j. This definition of the χ^2 function assumes that systematic uncertainties are proportional to the central prediction values (multiplicative errors), whereas the statistical uncertainties scale with the square root of the expected number of events.

During the χ^2 minimisation, the nuisance parameters b_j and the PDFs are determined.

Mixed Form Representation: In some cases, the statistical and systematic uncertainties are provided in different forms. For example, the correlated experimental systematic uncertainties are available as nuisance parameters but the bin-to-bin statistical correlations are given in a form of covariance matrix. HERAFitter offers possibilities to include also the mixed form of treating statistical, uncorrelated and correlated systematic uncertainties.

Any source of the measurement systematic uncertainty can be treated as additive or multiplicative. The statistical uncertainties can be included as additive or Poisson. Minimisation with respect to nuisance parameters is performed analytically, however for more detailed studies of correlations individual nuisance parameters can be included in the MINUIT minimisation.

5.3 Treatment of the Experimental Uncertainties

Three distinct methods for propagating experimental uncertainties to PDFs are implemented in HERAFitter and reviewed here: the Hessian, Offset, and Monte Carlo method.

Hessian (Eigenvector) method: The PDF uncertainties reflecting the uncertainties in experimental data are esitimated by examining the shape of χ^2 in the neighborhood of the minimum [88]. Following approach of Ref. [88], the Hessian matrix is defined by the second derivatives of χ^2 on the fitted PDF parameters. The matrix is diagonalized and the Hessian eigenvectors are computed. Due to orthogonality, these vectors correspond to statistically independent sources of the uncertainties in the PDFs obtained.

Offset method: The Offset method [89] uses also the χ^2 function for the central fit for which only uncorrelated uncertainties are taken into account. The goodness of the fit can no longer be judged from the χ^2 since correlated uncertainties are ignored. The correlated uncertainties are propagated into the PDF uncertainties performing the variants of fit with the experimental data varied by $\pm 1\sigma$ from the central value for each systematic source. Since the resulting deviation of the PDF parameters from the ones obtained in the central fit are statistically independent, they are combined in quadrature to arive to the total PDF systematic uncertainty.
In most cases, the uncertainties estimated by the offset method are larger than those from the Hessian method.

Monte Carlo method: The Monte-Carlo technique [90, 91] can be used to determine PDF uncertainties. The uncertainties are estimated using the pseudo-data replicas (typically > 100) randomly generated from the measurement central values and their systematic and statistical uncertainties taking into account all point-to-point correlations. The QCD fit is performed for each replica and the PDF central values with their experimental uncertainties are estimated using distribution of the PDF parameters over these fits, i.e. the mean values and standard deviations over the replicas.
The MC method was checked against the standard error estimation of the PDF uncertainties obtained by the Hessian method. A good agreement was found between the methods once the Gaussian distribution of statistic and systematic uncertainties is assumed in the MC approach [32]. This comparison is illustrated in Fig. 7. Similar findings were reported by the MSTW global analysis [92].
Since the MC method requires large number of replicas, the eigenvector representation is often more practical to represent PDF uncertainties. As it was illustrated by [93], it is possible to transform MC to eigenvector representation. Tools to perform this transformation are provided with `HERAFitter` and were recently employed to obtain correlated sets of PDFs at different perturbative order [94].

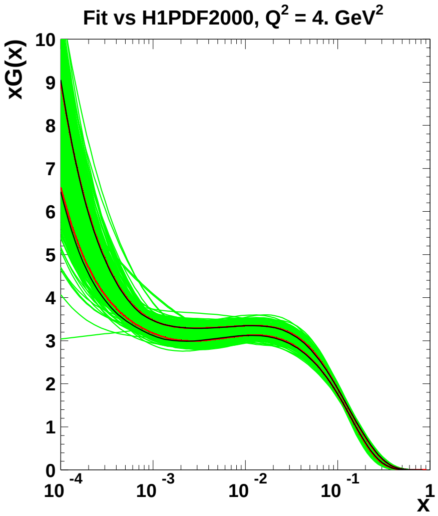

Fig. 7 Comparison between the standard error calculations as employed by the Hessian approach (black lines) and the MC approach (with more than 100 replicas) assuming Gaussian distribution for uncertainty distributions, shown here for each replica (green lines) together with the evaluated standard deviation (red lines) [32]. The black lines in the figure are mostly covered by the red lines.

The nuisance parameter representation of χ^2 in Eq. 18 is derived assuming symmetric experimental errors, however, the published systematic uncertainties are rather often asymmetric. `HERAFitter` provides the possibility to use asymmetric systematic uncertainties. The implementation relies on the assumption that asymmetric uncertainties can be described by a parabolic function and the nuisance parameter in Eq. 18 is modified as follows

$$\gamma^i_j \to \omega^i_j b_j + \gamma^i_j, \tag{19}$$

where the coefficients ω^i_j, γ^i_j are defined by the up and down values of the systematic uncertainties, $S^{\pm}_{ij}$,

$$\omega^i_j = \frac{1}{2}\left(S^+_{ij} + S^-_{ij}\right), \qquad \gamma^i_j = \frac{1}{2}\left(S^+_{ij} - S^-_{ij}\right). \tag{20}$$

5.4 Treatment of the Theoretical Input Parameters

The results of a QCD fit depend not only on the input data but also on the input parameters used in the theoretical calculations. Nowadays, the PDF groups address the impact of the choices of theoretical parameters by providing alternative PDFs with different choices of the mass of the charm quarks m_c, mass of the bottom quarks m_b and the value of $\alpha_s(M_Z)$. Another important issue is the choice of the functional form for the PDFs at the starting scale and the value of the starting scale itself. `HERAFitter` provides possibility of different user choices of various input parameters of the theory.

5.5 Bayesian Reweighting Techniques

As an alternative to performing a full QCD fit, `HERAFitter` allows to assess the impact of including new data in an existing fit using the Bayesian Reweighting technique. The method provides a fast estimate of the impact of new data on PDFs. Bayesian Reweighting was first proposed for the PDF sets delivered in form of MC replicas ensembles by [90] and further developed by the NNPDF Collaboration [95, 96]. More recently, a method to preform Bayesian Reweighting studies starting from PDF fits where uncertainties are provided in form of parameter eigenvectors has been also developed [92]. The latter is based on generating replica set by introducing Gaussian fluctuations on the central PDF set with a variance determined by the PDF uncertainty given by the eigenvectors. Both reweighting methods are implemented in `HERAFitter`.

The Bayesian Reweighting technique relies on the fact that the MC replicas of a PDF sets (i.e. NNPDF) give a representation of the probability distribution in the space of PDFs. In particular, the PDFs are represented as ensembles of $N_{\rm rep}$ equiprobable (*i.e.* having all weight equal to unity) replicas. The central value for a given observable, $\mathscr{O}(\mathrm{PDF})$, is computed as the average of the predictions obtained from the ensemble as

$$\langle \mathscr{O}(\mathrm{PDF}) \rangle = \frac{1}{N_{\rm rep}} \sum_{k=1}^{N_{\rm rep}} \mathscr{O}(\mathrm{PDF}_k), \tag{21}$$

and the uncertainty as the standard deviation of the sample.

Upon inclusion of new data the prior probability distribution, given by the prior PDF set, is updated according to Bayes Theorem and the weight of each replica, w_k, is updated according to

$$w_k = \frac{(\chi^2_k)^{\frac{1}{2}(N_{\rm data}-1)} e^{-\frac{1}{2}\chi^2_k}}{\frac{1}{N_{\rm rep}} \sum_{k=1}^{N_{\rm rep}} (\chi^2_k)^{\frac{1}{2}(N_{\rm data}-1)} e^{-\frac{1}{2}\chi^2_k}}, \tag{22}$$

where $N_{\rm data}$ is the number of new data points, k denotes the specific replica for which the weight is calculated and χ^2_k is the chi-square of the new data obtained using the k-th PDF replica. Given a PDF set and a corresponding set of weights, which describes the impact on the same set of the inclusion of new data, the prediction for a given observable can be computed as the *weighted* average,

$$\langle \mathscr{O}(\mathrm{PDF}) \rangle = \frac{1}{N_{\rm rep}} \sum_{k=1}^{N_{\rm rep}} w_k \mathscr{O}(\mathrm{PDF}_k). \tag{23}$$

To simplify the use of reweighted set, an unweighted set (*i.e.* a set of equiprobable replicas which incorporates the information of the original weights) is generated using the method described in [95].

The number of effective replicas of a reweighted sets, that is the size of an equiprobable replicas set containing the same amount of information as the reweighted set in question, is measured by the Shannon Entropy

$$N_{\rm eff} \equiv \exp\left\{ \frac{1}{N_{\rm rep}} \sum_{k=1}^{N} \mathrm{rep} w_k \ln(N_{\rm rep}/w_k) \right\}. \tag{24}$$

On the one hand there is no reason in generating a final unweighted set that has a number of replicas (significantly) larger than $N_{\rm eff}$ as no extra information is gained. On the other hand it is advisable to start from a prior PDF set which has as many replicas as possible in order to have a more accurate posterior set at the end of the reweighting procedure.

6 Alternatives to DGLAP Formalism

The QCD calculations based on the DGLAP [6–10] evolution equations are very successful in describing all relevant hard scattering data in the perturbative region $Q^2 \gtrsim 1\,\mathrm{GeV}^2$. At small-$x$ and small-Q^2 the DGLAP dynamics may be modified by non-perturbative QCD effects like saturation-based dipole models and other higher twist effects. Different approaches that are alternatives to the DGLAP formalism can be used to analyse DIS data in `HERAFitter`. These include several different dipole models and the use of transverse momentum dependent, or unintegrated PDFs (uPDFs).

6.1 Dipole Models

The dipole picture provides an alternative approach to the proton-virtual photon scattering at low x providing the description of both inclusive and diffractive processes. In this approach, the virtual photon fluctuates into a $q\bar{q}$ (or $q\bar{q}g$) dipole which interacts with the proton [97]. The dipoles can be considered as quasi-stable quantum mechanical states, which have very long life time $\propto 1/m_p x$ and a size which is not changed by scattering. The dynamics of the interaction are embedded in the dipole scattering amplitude.

Several dipole models which assume different behavior of the dipole-proton cross sections are implemented in `HERAFitter`: the Golec-Biernat-Wüsthoff (GBW) dipole saturation model [28], the colour glass condensate approach to the high parton density regime called the Iancu-Itakura-Munier (IIM) dipole model [29] and a modified GBW model which takes into account the effects of DGLAP evolution called the Bartels-Golec-Kowalski (BGK) dipole model [30].

GBW model: In the GBW model the dipole-proton cross section $\sigma_{\rm dip}$ is given by

$$\sigma_{\rm dip}(x, r^2) = \sigma_0 \left(1 - \exp\left[-\frac{r^2}{4R_0^2(x)}\right]\right), \tag{25}$$

where r corresponds to the transverse separation between the quark and the antiquark, and R_0^2 is an x-dependent scale parameter which represents the spacing of the gluons in the proton. $R_0^2(x) = (x/x_0)^\lambda / 1\ \mathrm{GeV}^{-2}$ is called the saturation radius. The cross-section normalisation σ_0, x_0, and λ are parameters of the model commonly fitted to the DIS data. This model gives exact Bjorken scaling when the dipole size r is small.

IIM model: The IIM model assumes an improved expression for the dipole cross section which is based on the Balitsky-Kovchegov equation [98]. The explicit formula for σ_{dip} can be found in [29]. The alternative scale parameter $\tilde{R}$, x_0 and λ are fitted parameters of the model.

BGK model: The BGK model is a modification of the GBW model assuming that the spacing R_0 is inverse of the gluon density and taking into account the DGLAP evolution of the latter. The gluon density parametrised at some starting scale Q_0^2 by Eq. 12 is evolved to larger scales using DGLAP evolution.

BGK model with valence quarks: The dipole models are valid in the low-x region only, where the valence quark contribution to the total proton momentum is 5% to 15% for x from 0.0001 to 0.01 [99]. The new HERA F_2 measurements have a precision which is better than 2%. Therefore, in `HERAFitter` the contribution of the valence quarks can be taken into account in the original BGK model [100].

6.2 Transverse Momentum Dependent PDFs

QCD calculations of multiple-scale processes and complex final-states require in general transverse-momentum dependent (TMD) [101], or unintegrated, parton distribution and parton decay functions [102–110]. The TMD factorisation has been proven recently [101] for inclusive DIS. For particular hadron-hadron scattering processes, like heavy flavor, vector boson and Higgs production, TMD factorisation has also been proven in the high-energy (small-x) limit [111–113]

In the framework of high-energy factorisation [111, 114, 115] the DIS cross section can be written as a convolution in both longitudinal and transverse momenta of the TMD parton density function $\mathscr{A}(x,k_t,\mu)$ with the off-shell partonic matrix elements, as follows

$$\sigma_j(x,Q^2) = \int_x^1 dz \int d^2k_t\ \hat{\sigma}_j(x,Q^2,z,k_t)\ \mathscr{A}(z,k_t,\mu) \qquad (26)$$

with the DIS cross sections σ_j, $(j = 2,L)$ related to the structure functions F_2 and F_L. The hard-scattering kernels $\hat{\sigma}_j$ of Eq. 26, are k_t-dependent and the evolution of the transverse-momentum dependent gluon density $\mathscr{A}$ is obtained by combining the resummation of small-x logarithmic contributions [116–118] with medium-x and large-x contributions to parton splitting [6, 9, 10] according to the CCFM evolution equation [25, 119, 120].

The factorisation formula (26) allows resummation of logarithmically enhanced small-x contributions to all orders in perturbation theory, both in the hard scattering coefficients and in the parton evolution, fully taking into account the dependence on the factorisation scale μ and on the factorisation scheme [121, 122].

The cross section σ_j, $(j = 2,L)$ is calculated in a FFN scheme, where only the boson-gluon fusion process ($\gamma^* g^* \to q\bar{q}$) is included. The masses of the quarks are explicitly included as parameters of the model. In addition to $\gamma^* g^* \to q\bar{q}$, the contribution from valence quarks is included via $\gamma^* q \to q$ by using a CCFM evolution of valence quarks [123, 124].

CCFM Grid Techniques: The CCFM evolution cannot be written easily in an analytic closed form. For this reason a Monte Carlo method is employed, which is however time-consuming, and cannot be used in a straightforward manner in a fit program.

Following the convolution method introduced in [124, 125], the kernel $\tilde{\mathscr{A}}(x'',k_t,p)$ is determined from the Monte Carlo solution of the CCFM evolution equation, and then folded with the non-perturbative starting distribution $\mathscr{A}_0(x)$.

$$\begin{aligned} x\mathscr{A}(x,k_t,p) &= x\int dx' \int dx'' \mathscr{A}_0(x')\tilde{\mathscr{A}}(x'',k_t,p)\,\delta(x'x''-x) \\ &= \int dx' \mathscr{A}_0(x') \cdot \frac{x}{x'}\, \tilde{\mathscr{A}}\left(\frac{x}{x'},k_t,p\right), \qquad (27) \end{aligned}$$

where k_t denotes the transverse momentum of the propagator gluon and p is the evolution variable.

The kernel $\tilde{\mathscr{A}}$ incorporates all of the dynamics of the evolution. It is defined on a grid of $50 \otimes 50 \otimes 50$ bins in x, k_t, p. The binning in the grid is logarithmic, except for the longitudinal variable x where 40 bins in logarithmic spacing below 0.1, and 10 bins in linear spacing above 0.1 are used.

Calculation of the cross section according to Eq. 26 involves a multidimensional Monte Carlo integration which is time consuming and suffers from numerical fluctuations. This cannot be employed directly in a fit procedure involving the calculation of numerical derivatives in the search for the minimum. Instead the following equation is applied:

$$\begin{aligned} \sigma(x,Q^2) &= \int_x^1 dx_g \mathscr{A}(x_g,k_t,p)\hat{\sigma}(x,x_g,Q^2) \\ &= \int_x^1 dx' \mathscr{A}_0(x') \cdot \tilde{\sigma}(x/x',Q^2) \qquad (28) \end{aligned}$$

Here, first $\tilde{\sigma}(x',Q^2)$ is calculated numerically with a Monte Carlo integration on a grid in x for the values of Q^2 used in the fit. Then the last step in Eq. 28 is performed with a fast numerical gauss integration, which can be used in standard fit procedures.

Functional Forms for TMD parameterisation: For the starting distribution $\mathscr{A}_0$, at the starting scale Q_0, the following form is used:

$$x\mathscr{A}_0(x,k_t) = Nx^{-B} \cdot (1-x)^C \left(1 - Dx + E\sqrt{x}\right) \exp[-k_t^2/\sigma^2]\ , \qquad (29)$$

with $\sigma^2 = Q_0^2/2$ and the free parameters N, B, C, D, E. Valence quarks are treated using the method of [123] as described in [124] with a starting distribution taken from any collinear PDF and imposing the flavor sum rule at every scale p.
The TMD parton densities can be plotted either with HERAFitter provided tools or with TMDplotter[35].

7 HERAFitter Code Organisation

HERAFitter is an open source code and it can be downloaded from the dedicated webpage [1] together with its supporting documentation and fast grid theory files (described in section 4) associated with the properly formatted data files available in HERAFitter. The source code contains all the relevant information to perform QCD fits with HERA DIS data as a default set [1]. The performance time depends on the fitting options and varies from 10 minutes (using 'FAST' techniques as described in section 4) to several hours when full uncertainties are estimated. The HERAFitter code is a combination of C++ and Fortran 77libraries with minimal dependencies, i.e. for the default fitting options no external dependences are required except QCDNUM evolution program [22] and CERN libraries. The ROOT libaries are only required for the drawing tools and when invoking APPLGRID . Drawing tool inbuilt in HERAFitter provides a qualitative and quantitative assessment of the results. Fig. 8 shows an illustration of a comparison between the inclusive NC data from the HERA I with the predictions based on HERAPDF1.0 PDFs. The consistency of the measurements and the theory is expressed by pulls, defined as a difference between data and theory divided by the uncorrelated error of the data. In each kinematic bin of the measurement, pulls are provided in units of standard deviation (sigma). As an additional consistency check between data and the theory predictions, pull information, defined as the difference between data and prediction divided by the uncorrelated uncertainty of the data, is displayed in units of sigma shifts for each given data bin.

There are also cache options, fast evolution kernels, and usage of the OpenMP (Open Multi-Processing) interface which allows parallel applications of the GM-VFNS theory predictions in DIS. In addition, the HERAFitter references and GNU public licence are provided together with the main source code.

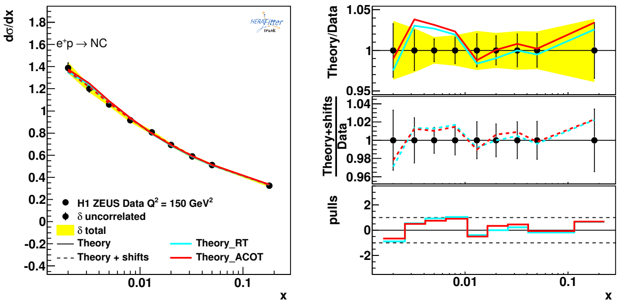

Fig. 8 An illustration of the consistency of HERA measurements [36] and the theory predictions, obtained in HERAFitter with the default drawing tool.

8 Applications of HERAFitter

The HERAFitter program was used in a number of experimental and theoretical analyses. This list includes several LHC analyses of SM processes, namely inclusive Drell-Yan and $W$and Z production [11, 13, 14], inclusive jet production [12]. The results of QCD analyses using HERAFitter were also published by HERA experiments in the inclusive [15] and the heavy flavour production measurements [16, 126]. Following theory and phenomenology studies were performed with HERAFitter: a determination of the transverse momentum dependent gluon density using precision HERA data [127], an analysis of HERA data within a dipole model [100], the study of the low-x uncertainties in PDFs determined from the HERA data using different parameterisatons [87] and the impact of QED radiative corrections on PDFs [128]. A recent study based on a set of PDFs determined with the HERAFitter and addressing the correlated uncertainties between orders was published in [94].

The HERAFitter framework has been used to produce PDF grids from the QCD analyses performed at HERA [36, 129] and at the LHC [130], using measurements from ATLAS [11, 12], which can be used to study predictions for SM or beyond SM processes. Moreover, HERAFitter provides a possibility to perform impact studies for possible future colliders as demonstrated by the QCD studies at the LHeC [131].

[1]Default settings in HERAFitter are tuned to reproduce the central HERAPDF1.0 set.

9 Summary

The HERAFitter project is a unique platform for QCD analyses to study the structure of the proton. The project successfully encapsulates a wide variety of QCD tools to facilitate analyses of the experimental data and theoretical calculations. HERAFitter is the first open source platform which can be used for benchmarking studies. It allows for direct comparisons of various theoretical approaches under the same settings, and a variety of different methodologies in treating the experimental and model uncertainties. The growth of HERAFitter is driven by the QCD advances in theoretical calculations and in precision of experimental data.

Acknowledgements HERAFitter developers team acknowledges the kind hospitality of DESY and funding by the Helmholtz Alliance "Physics at the Terascale" of the Helmholtz Association. We are grateful to the DESY IT department for their support of the HERAFitter developers. Additional support was received from the BMBF-JINR cooperation program, the Heisenberg-Landau program, the RFBR grant 12-02-91526-CERN a, the Polish NSC project DEC-2011/03/B/ST2/00220 and a dedicated funding of the Initiative and Networking Fond of Helmholtz Association SO-072. We also acknowledge Nathan Hartland with Luigi Del Debbio for contributing to the implementation of the Bayesian Reweighting technique and would like to thank R. Thorne for fruitful discussions.

References

1. *HERAFitter*, https://www.herafitter.org.
2. G. Aad *et al.* [ATLAS Collaboration], Phys.Lett. **B716**, 1 (2012), [arXiv:1207.7214].
3. S. Chatrchyan *et al.* [CMS Collaboration], Phys.Lett. **B716**, 30 (2012), [arXiv:1207.7235].
4. E. Perez and E. Rizvi, Rep.Prog.Phys. **76**, 046201 (2013), [arXiv:1208.1178].
5. S. Forte and G. Watt, Ann.Rev.Nucl.Part.Sci. **63**, 291 (2013), [arXiv:1301.6754].
6. V. N. Gribov and L. N. Lipatov, Sov. J. Nucl. Phys. **15**, 438 (1972).
7. V. N. Gribov and L. N. Lipatov, Sov. J. Nucl. Phys. **15**, 675 (1972).
8. L. N. Lipatov, Sov. J. Nucl. Phys. **20**, 94 (1975).
9. Y. L. Dokshitzer, Sov. Phys. JETP **46**, 641 (1977).
10. G. Altarelli and G. Parisi, Nucl. Phys. B **126**, 298 (1977).
11. G. Aad *et al.* [ATLAS Collaboration], Phys. Rev. Lett. **109**, 012001 (2012), [arXiv:1203.4051].
12. G. Aad *et al.* [ATLAS Collaboration], Eur.Phys.J. **73**, 2509 (2013), [arXiv:1304:4739].
13. G. Aad *et al.* [ATLAS Collaboration], Phys. Lett. **B725**, 223 (2013), [arXiv::1305.4192].
14. S. Chatrchyan *et al.* [CMS Collaboration], submitted to Phys. Rev. **D** (2014), [arXiv:1312.6283].
15. F. Aaron *et al.* [H1 Collaboration], JHEP **1209**, 061 (2012), [arXiv:1206.7007].
16. H. Abramowicz *et al.* [H1 and ZEUS Collaborations], Eur. Phys. J. **C73**, 2311 (2013), [arXiv:1211.1182].
17. A. Martin, W. Stirling, R. Thorne, and G. Watt, Eur. Phys. J. C **63**, 189 (2009), [arXiv:0901.0002], URL `http://mstwpdf.hepforge.org/`.
18. J. Gao, M. Guzzi, J. Huston, H.-L. Lai, Z. Li, *et al.*, Phys.Rev. **D89**, 033009 (2014), [arXiv:1302.6246], URL `http://hep.pa.msu.edu/cteq/public/`.
19. R. D. Ball *et al.*, Nucl.Phys. **B867**, 244 (2013), [arXiv:1207.1303], URL `https://nnpdf.hepforge.org/`.
20. S. Alekhin, J. Bluemlein, and S. Moch, Phys.Rev. **D89**, 054028 (2014), [arXiv:1310.3059].
21. P. Jimenez-Delgado and E. Reya, Phys.Rev. **D80**, 114011 (2009), [arXiv:0909.1711], URL `http://www.het.physik.tu-dortmund.de/pdfserver/index.html`.
22. M. Botje (2010), http://www.nikef.nl/h24/qcdnum/index.html, [arXiv:1005.1481].
23. M. Ciafaloni, Nucl. Phys. B **296**, 49 (1988).
24. S. Catani, F. Fiorani, and G. Marchesini, Phys. Lett. B **234**, 339 (1990).
25. S. Catani, F. Fiorani, and G. Marchesini, Nucl. Phys. B **336**, 18 (1990).
26. G. Marchesini, Nucl. Phys. B **445**, 49 (1995).
27. H. Jung *et al.*, *The CCFM uPDF evolution* (2014), DESY-14-060.
28. K. Golec-Biernat and M. Wüsthoff, Phys. Rev. D **59**, 014017 (1999), [hep-ph/9807513].
29. E. Iancu, K. Itakura, and S. Munier, Phys. Lett. **B590**, 199 (2004), [hep-ph/0310338].
30. J. Bartels, K. Golec-Biernat, and H. Kowalski, Phys. Rev. D **66**, 014001 (2002), [hep-ph/0203258].
31. F. James and M. Roos, Comput. Phys. Commun. **10**, 343 (1975).
32. M. Dittmar, S. Forte, A. Glazov, S. Moch, G. Altarelli, *et al.* (2009), [arXiv:0901.2504].
33. M. Whalley, D. Bourilkov, and R. Group (2005), [hep-ph/0508110].
34. *LHAPDF*, URL `http://lhapdf.hepforge.org`.
35. H. Jung *et al.*, *TMDlib and TMDplotter: library and plotting tools for Transverse Momentum Dependent parton distributions* (2014), DESY-14-059.
36. F. Aaron *et al.* [H1 and ZEUS Collaborations], JHEP **1001**, 109 (2010), [arXiv:0911.0884].
37. R. Devenish and A. Cooper-Sarkar (2011), *Deep Inelastic Scattering*, ISBN: 0199602255,9780199602254.
38. J. C. Collins and W.-K. Tung, Nucl. Phys. B **278**, 934 (1986).
39. E. Laenen *et al.*, Phys. Lett. **B291**, 325 (1992).
40. E. Laenen *et al.*, Nucl. Phys. **B392**, 162, 229 (1993).
41. S. Riemersma, J. Smith, and van Neerven. W.L., Phys. Lett. **B347**, 143 (1995), [hep-ph/9411431].
42. S. Alekhin, J. Blümlein, and S. Moch, *OPENQCDRAD*, http://www-zeuthen.desy.de/∼alekhin/OPENQCDRAD.
43. H. Kawamura, N. Lo Presti, S. Moch, and A. Vogt, Nucl.Phys. **B864**, 399 (2012).
44. S. Alekhin and S. Moch, Phys. Lett. **B699**, 345 (2011), [arXiv:1011.5790].
45. R. Demina, S. Keller, M. Kramer, S. Kretzer, R. Martin, *et al.* (1999), [hep-ph/0005112].
46. R. S. Thorne and R. G. Roberts, Phys. Rev. D **57**, 6871 (1998), [hep-ph/9709442].

47. R. S. Thorne, Phys. Rev. **D73**, 054019 (2006), [hep-ph/0601245].
48. R. S. Thorne, Phys. Rev. D **86**, 074017 (2012), [arXiv:1201.6180].
49. J. C. Collins, Phys.Rev. **D58**, 094002 (1998), [hep-ph/9806259].
50. M. Aivazis, J. C. Collins, F. I. Olness, and W.-K. Tung, Phys.Rev. **D50**, 3102 (1994), [hep-ph/9312319].
51. M. Kramer, F. I. Olness, and D. E. Soper, Phys. Rev. **D62**, 096007 (2000), [hep-ph/0003035].
52. S. Kretzer, H. Lai, F. Olness, and W. Tung, Phys.Rev. **D69**, 114005 (2004), [hep-ph/0307022].
53. H. Spiesberger, Private communication.
54. F. Jegerlehner, Proceedings, LC10 Workshop **DESY 11-117** (2011).
55. H. Burkhard, F. Jegerlehner, G. Penso, and C. Verzegnassi, in CERN Yellow Report on "Polarization at LEP" 1988.
56. A. Aktas *et al.* [H1 Collaboration], Eur.Phys.J. **C48**, 715 (2006), [hep-ex/0606004].
57. S. Chekanov *et al.* [ZEUS Collaboration], Nucl. Phys. **B831**, 1 (2010), [hep-ex/09114119].
58. S. A. Malik and G. Watt, JHEP **1402**, 025 (2014), [arXiv:1304.2424].
59. S. D. Drell and T.-M. Yan, Phys. Rev. Lett. **25**, 316 (1970).
60. M. Yamada and M. Hayashi, Nuovo Cim. **A70**, 273 (1982).
61. J. M. Campbell and R. K. Ellis, Phys. Rev. **D60**, 113006 (1999), [arXiv:9905386].
62. J. M. Campbell and R. K. Ellis, Phys. Rev. **D62**, 114012 (2000), [arXiv:0006304].
63. J. M. Campbell and R. K. Ellis, Nucl. Phys. Proc. Suppl. **205-206**, 10 (2010), [arXiv:1007.3492].
64. Y. Li and F. Petriello, Phys.Rev. **D86**, 094034 (2012), [arXiv:1208.5967].
65. G. Bozzi, J. Rojo, and A. Vicini, Phys.Rev. **D83**, 113008 (2011), [arXiv:1104.2056].
66. A. Gehrmann-De Ridder, T. Gehrmann, E. Glover, and J. Pires, Phys. Rev. Lett. **110**, 162003 (2013), [arXiv:1301.7310].
67. E. Glover and J. Pires, JHEP **1006**, 096 (2010), [arXiv:1003.2824].
68. J. Currie, A. Gehrmann-De Ridder, E. Glover, and J. Pires, JHEP **1401**, 110 (2014), [arXiv:1310.3993].
69. Z. Nagy and Z. Trocsanyi, Phys.Rev. **D59**, 014020 (1999), [hep-ph/9806317].
70. Z. Nagy, Phys.Rev.Lett. **88**, 122003 (2002), [hep-ph/0110315].
71. S. Chatrchyan *et al.* [CMS Collaboration], Phys.Lett. **B728**, 496 (2014), [arXiv:1307.1907].
72. M. Czakon, P. Fiedler, and A. Mitov, Phys. Rev. Lett. **110**, 252004 (2013), [arXiv:1303.6254].
73. M. Aliev, H. Lacker, U. Langenfeld, S. Moch, P. Uwer, *et al.*, Comput.Phys.Commun. **182**, 1034 (2011), [arXiv:1007.1327].
74. J. M. Campbell, R. Frederix, F. Maltoni, and F. Tramontano, Phys.Rev.Lett. **102**, 182003 (2009), [arXiv:0903.0005].
75. J. M. Campbell and F. Tramontano, Nucl.Phys. **B726**, 109 (2005), [hep-ph/0506289].
76. J. M. Campbell, R. K. Ellis, and F. Tramontano, Phys.Rev. **D70**, 094012 (2004), [hep-ph/0408158].
77. J. M. Campbell and R. K. Ellis (2012), report FERMILAB-PUB-12-078-T, [arXiv:1204.1513].
78. T. Kluge, K. Rabbertz, and M. Wobisch, pp. 483–486 (2006), [hep-ph/0609285].
79. T. Carli *et al.*, Eur. Phys. J. **C66**, 503 (2010), [arXiv:0911.2985].
80. Z. Nagy and Z. Trocsanyi, Phys.Rev.Lett. **87**, 082001 (2001), [hep-ph/0104315].
81. Z. Nagy, Phys.Rev. **D68**, 094002 (2003), [hep-ph/0307268].
82. M. Wobisch, D. Britzger, T. Kluge, K. Rabbertz, and F. Stober [fastNLO Collaboration] (2011), [arXiv:1109.1310].
83. N. Kidonakis and J. Owens, Phys.Rev. **D63**, 054019 (2001), [hep-ph/0007268].
84. D. Britzger, K. Rabbertz, F. Stober, and M. Wobisch [fastNLO Collaboration] (2012), [arXiv:1208.3641].
85. http://fastnlo.hepforge.org, URL `http://fastnlo.hepforge.org`.
86. http://applgrid.hepforge.org, URL `http://applgrid.hepforge.org`.
87. A. Glazov, S. Moch, and V. Radescu, Phys. Lett. B **695**, 238 (2011), [arXiv:1009.6170].
88. J. Pumplin, D. Stump, R. Brock, D. Casey, J. Huston, *et al.*, Phys.Rev. **D65**, 014013 (2001), [hep-ph/0101032].
89. M. Botje, J.Phys. **G28**, 779 (2002), [hep-ph/0110123].
90. W. T. Giele and S. Keller, Phys.Rev. **D58**, 094023 (1998), [hep-ph/9803393].
91. W. T. Giele, S. Keller, and D. Kosower (2001), [hep-ph/0104052].
92. G. Watt and R. Thorne, JHEP **1208**, 052 (2012), [arXiv:1205.4024].
93. J. Gao and P. Nadolsky, JHEP **1407**, 035 (2014), [arXiv:1401.0013].
94. HERAFitter Developers Team and M. Lisovyi (2014), [arXiv:1404.4234].
95. R. D. Ball, V. Bertone, F. Cerutti, L. Del Debbio, S. Forte, *et al.*, Nucl.Phys. **B855**, 608 (2012), [arXiv:1108.1758].
96. R. D. Ball *et al.* [NNPDF Collaboration], Nucl.Phys. **B849**, 112 (2011), [arXiv:1012.0836].

97. N. N. Nikolaev and B. Zakharov, Z.Phys. **C49**, 607 (1991).
98. I. Balitsky, Nucl. Phys. B **463**, 99 (1996), [hep-ph/9509348].
99. F. Aaron *et al.* [H1 Collaboration], Eur.Phys.J. **C71**, 1579 (2011), [arXiv:1012.4355].
100. A. Luszczak and H. Kowalski (2013), [arXiv:1312.4060].
101. J. Collins, *Foundations of perturbative QCD*, vol. 32 (Cambridge monographs on particle physics, nuclear physics and cosmology., 2011).
102. S. M. Aybat and T. C. Rogers, Phys.Rev. **D83**, 114042 (2011), [arXiv:1101.5057].
103. M. Buffing, P. Mulders, and A. Mukherjee, Int.J.Mod.Phys.Conf.Ser. **25**, 1460003 (2014), [arXiv:1309.2472].
104. M. Buffing, A. Mukherjee, and P. Mulders, Phys.Rev. **D88**, 054027 (2013), [arXiv:1306.5897].
105. M. Buffing, A. Mukherjee, and P. Mulders, Phys.Rev. **D86**, 074030 (2012), [arXiv:1207.3221].
106. P. Mulders, Pramana **72**, 83 (2009), [arXiv:0806.1134].
107. S. Jadach and M. Skrzypek, Acta Phys.Polon. **B40**, 2071 (2009), [arXiv:0905.1399].
108. F. Hautmann, Acta Phys.Polon. **B40**, 2139 (2009).
109. F. Hautmann, M. Hentschinski, and H. Jung (2012), [arXiv:1205.6358].
110. F. Hautmann and H. Jung, Nucl.Phys.Proc.Suppl. **184**, 64 (2008), [arXiv:0712.0568].
111. S. Catani, M. Ciafaloni, and F. Hautmann, Phys. Lett. B **242**, 97 (1990).
112. J. C. Collins and R. K. Ellis, Nucl. Phys. B **360**, 3 (1991).
113. F. Hautmann, H. Jung, and V. Pandis, AIP Conf.Proc. **1350**, 263 (2011), [arXiv:1011.6157].
114. S. Catani, M. Ciafaloni, and F. Hautmann, Nucl. Phys. B **366**, 135 (1991).
115. S. Catani, M. Ciafaloni, and F. Hautmann, Phys. Lett. B **307**, 147 (1993).
116. L. Lipatov, Phys.Rept. **286**, 131 (1997), [hep-ph/9610276].
117. V. S. Fadin, E. Kuraev, and L. Lipatov, Phys.Lett. **B60**, 50 (1975).
118. I. I. Balitsky and L. N. Lipatov, Sov. J. Nucl. Phys. **28**, 822 (1978).
119. M. Ciafaloni, Nucl. Phys. **B296**, 49 (1988).
120. G. Marchesini, Nucl. Phys. B **445**, 49 (1995), [hep-ph/9412327].
121. S. Catani and F. Hautmann, Nucl. Phys. B **427**, 475 (1994), [hep-ph/9405388].
122. S. Catani and F. Hautmann, Phys.Lett. **B315**, 157 (1993).
123. M. Deak, F. Hautmann, H. Jung, and K. Kutak, *Forward-Central Jet Correlations at the Large Hadron Collider* (2010), [arXiv:1012.6037].
124. F. Hautmann and H. Jung, Nuclear Physics B **883**, 1 (2014), [arXiv:1312.7875].
125. H. Jung and F. Hautmann (2012), [arXiv:1206.1796].
126. H. Abramowicz *et al.* [ZEUS Collaboration] (2014), [1405.6915].
127. F. Hautmann and H. Jung (2013), [arXiv:1312.7875].
128. R. Sadykov (2014), [arXiv:1401.1133].
129. *HERAPDF1.5LO, NLO and NNLO* (H1prelim-13-141 and ZEUS-prel-13-003, H1prelim-10-142 and ZEUS-prel-10-018, H1prelim-11-042 and ZEUS-prel-11-002), available via: http://lhapdf.hepforge.org/pdfsets.
130. *ATLAS NNLO epWZ12*, available via: http://lhapdf.hepforge.org/pdfsets.
131. J. L. Abelleira Fernandez *et al.* [LHeC Study Group], Journal of Phys. **G**, 075001 (2012), [arXiv:1206.2913].