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

HERAFitter

Open Source QCD Fit Project

First Author[a,1], **Second Author**[??,2,3]

[1]First address
[2]Second address
[3]*Present Address:* if needed

Received: date / Accepted: date

Abstract The HERAFitter project is presented which provides a framework for QCD analyses related to the proton structure in the context of multi-processes and multi-experiments. Utilising QCD factorisation, HERAFitter allows determining the proton parton distribution functions (PDF) from various hard scattering measurements or obtaining benchmark predictions for such processes using different PDFs. The main processes and data sets that are currently included are Deep-Inelastic-Scattering (DIS) in ep collisions at HERA and Drell Yan (DY), jet and top quark production in pp ($p\bar{p}$) collisions at the LHC (Tevatron). A large number of theoretical and methodological options is available in HERAFitter via interfaces to different modules.

Keywords First keyword · Second keyword · More

1 Introduction

In the era of the Higgs discovery and scrupulous searches for signals of new physics at the LHC it is crucial to have accurate Standard Model (SM) predictions for many hard scattering processes such as the Higgs production at the LHC. A most common approach to calculate the SM cross sections for such reactions is to use perturbative QCD collinear factorisation:

$$\sigma^{pp\to H+X}(\alpha_s,\mu_r,\mu_f) = \sum_{a,b}\int_0^1 dx_1\int_0^1 dx_2\, f_a(x_1,\alpha_s,\mu_F) f_b(x_2,\alpha_s,\mu_F)\times\hat{\sigma}^{ab\to H+X}(x_1,x_2;\alpha_s,\mu_R,\mu_F). \quad (1)$$

Here the desired cross section $\sigma^{pp\to H+X}$ for inclusive Higgs production is expressed as a convolution of Parton Distribution Functions (PDF) f_a and f_b with the partonic cross section $\hat{\sigma}^{ab\to H+X}$. The PDFs describe the probability of finding a specific parton a (b) in the first (second) proton carrying the fraction x_1 (x_2) of its momentum. The sum in Eq. 1 over a and b is over all different kind of partons, i.e. gluons and the various quarks and antiquarks flavours, that are considered as the constituents of the proton. Both the PDFs and the partonic cross section depend on the strong coupling constant α_s and the factorisation and renormalisation scales μ_F and μ_R. The partonic cross sections are calculable in pQCD while the PDFs cannot be determined solely with pQCD but are assumed to be universal (that is process independent). This make it possible to use different scattering reactions to constrain the PDFs; in particular one can use specific reaction data for determining the PDFs and then take these PDFs for predicting other processes via Eq. 1.

Key information on the PDFs is provided by the Deep Inelastic Scattering (DIS) data from the ep collider HERA. For instance, the gluon density relevant for calculating the dominant gluon gluon fusion contribution to Higgs production at the LHC can be accurately determined from the HERA data alone. Specific data from the Tevatron $p\bar{p}$ and the LHC pp collider can help to further constrain the PDFs. The most sensitive processes are Drell Yan production, W and Z asymmetries, top quark production, jet production and others.

HERAFitter is a QCD analysis tool based on QCD factorisation that aims at determining precise PDFs by integrating/using all the PDF sensitive information from HERA, Tevatron and the LHC. The processes that are currently included are listed in Tab. 1.

The basic functionality of HERAFitter is shown in Fig. 1 and consists of four parts: **needs to update figure!**

- Input data: All relevant cross section data from the various reactions are stored internally in HERAFitter with the full information on their uncorrelated and correlated uncertainties.
- Theory: Predictions are obtained using the factorisation approach (Eq. 1). The PDFs are parametrised at a start-

[a]e-mail: fauthor@example.com

Process Process	Diagram	Theory calculation	Available data
DIS NC			HERA
DIS CC			HERA
DIS jets			HERA
DIS heavy quark prod			HERA
$pp(\bar{p})$ Drell Yan		Applgrid Sapronov	Tevatron, LHC
$pp(\bar{p})$ W charge asym		Applgrid	Tevatron, LHC
$pp(\bar{p})$ top			Tevatron, LHC
$pp(\bar{p})$ jets		Applgrid FastNLO	Tevatron, LHC
$pp(\bar{p})$ DY+heavy quark		Applgrid	LHC

Table 1 The list of processes available in `HERAFitter` .

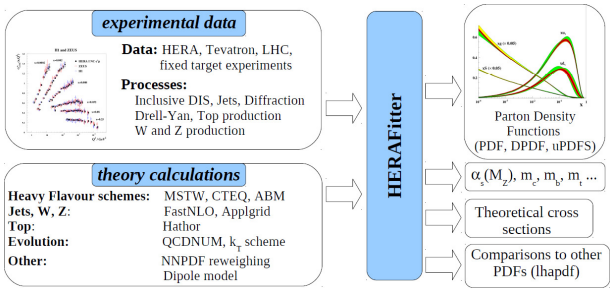

Fig. 1 Schematic structure of the `HERAFitter` program.

ing scale Q_0 by a certain functional form. They depend on a set of fit parameters **p**. The PDFs are evolved from Q_0 to the factorisation scale μ_F using DGLAP evolution as implemented in QCDNUM [1] and then multiplied (Eq. 1) with the hard parton cross sections calculated by a specific theory program (also listed in Tab. 1).

– Minimization: A least square fit is performed, constructing a χ^2 from the input data and the theory prediction. The χ^2 is minimized iteratively with respect to the PDF parameters using the MINUIT program.
– Output: The fitted parameters **p** and their estimated covariances from MINUIT are provided. The resulting PDFs can be graphically discplayed at arbitrary scales with their one sigma uncertainties bands. To demonstrate the fit consistency, plots are provided in which the input data are compared to the fitted theory predictions.

The outline of this paper is as follows. Section 2 discusses the various processes and corresponding theoretical calculations that are available in `HERAFitter` . Section 3 elucidates the methodology of determining PDF based on various approaches **(what do you mean here by approaches?)** and χ^2 definitions used in the minimisation. Specific applications of the package are given in section 4. **add something more here?.**

2 Theoretical Input

The proton PDFs are classically extracted from QCD fits by a measure of the agreement between experimental data and corresponding theory models. During the fit procedure in the `HERAFitter` framework, the PDFs are parametrised at a starting scale Q_0^2 chosen to be below the charm mass threshold and then evolved using coupled, integro-differential Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) [2, 3, 4, 5, 6] evolution equations as implemented in the QCDNUM [1] program in the $\overline{\text{MS}}$ scheme. The evolution can be performed in the LO, NLO or NNLO accuracy [7, 8].

In following, the theoretical models for various processes available in `HERAFitter` are described.

2.1 Deep Inelastic Scattering Formalism and Schemes

Deep inelastic scattering (DIS) is the lepton scattering on the constituents of the proton by a virtual exchange of a neutral (NC) or charged (CC) boson and, as a result, a scattered lepton and a multihadronic final state are produced. The DIS kinematic variables are the negative squared four-momentum of the exchange boson, Q^2, the scaling variable x, which can be related in the parton model to the fraction of momentum carried by the struck quark, and the inelasticity parameter y, which is the fraction of the energy transferred to the hadronic vertex.
The NC (and similarly CC) cross section can be expressed in terms of structure functions:

$$\frac{d^2\sigma_{NC}^{e^\pm p}}{dxdQ^2} = \frac{2\pi\alpha^2}{xQ^4}\left[Y_+\tilde{F}_2^\pm \mp Y_- x\tilde{F}_3^\pm - y^2\tilde{F}_L^\pm\right],$$

where $Y_\pm = 1 \pm (1-y)^2$ with y being the inelasticity. The structure function $\tilde{F}_2$ is the dominant contribution to the cross section, $x\tilde{F}_3$ is important at high Q^2 and $\tilde{F}_L$ is sizable only at high y. In the framework of perturbative QCD the structure functions are directly related to the parton distribution functions, i.e. in leading order (LO) F_2 is the momentum sum of quark and anti-quark distributions, $F_2 \approx x\sum e_q^2(q+\overline{q})$, and xF_3 is related to their difference, $xF_3 \approx x\sum 2e_q a_q(q-\overline{q})$. At higher orders, terms related to the gluon density distribution ($\alpha_s g$) appear.
In analogy to neutral currents, the inclusive CC ep cross section can be expressed in terms of structure functions and in LO the e^+p and e^-p cross sections are sensitive to different quark densities:

$$e^+: \ \tilde{\sigma}_{CC}^{e^+p} = x[\overline{u}+\overline{c}] + (1-y)^2 x[d+s]$$
$$e^-: \ \tilde{\sigma}_{CC}^{e^-p} = x[u+c] + (1-y)^2 x[\overline{d}+\overline{s}].$$

The QCD predictions for the DIS structure functions are obtained by convoluting the PDFs with the coefficient functions calculated using various schemes, i.e. the general mass

Variable-Flavour number (GM-VFN) [9] schemes or the Fixed-Flavour number (FFN) [10, 11, 12], The following VFN schemes with various treatments for the heavy quark thresholds are considered in HERAFitter : The Thorne Roberts (TR) scheme with its variants at NLO and NNLO [13, 14] as provided by the MSTW group, the ACOT scheme with its variants at LO and NLO as provided by the CTEQ group. In addition, the zero-mass variable flavour number scheme (ZM-VFNS) where heavy quark densities are included in the proton for $Q^2 >> m_h^2$ but are treated as massless in both the initial and final states can be used in HERAFitter . The FFN scheme is available via the QCDNUM implementation and via the OPENQCDRAD [15] interface. Each of these schemes is briefly discussed below.

2.1.1 GM-VFN Thorne-Roberts scheme

The Thorne-Roberts (TR) scheme (referred as RT scheme in the HERAFitter) smoothly connect the two regions: scales below ($Q^2 < m_h^2$) and scales much above the heavy quark scale threshold ($Q^2 >> m_h^2$). Therea retwo different variants of the TR schemes: TR standard (as used in MSTW PDF sets [14, 16]) and TR optimal [17], with a smoother transition across the heavy quark mass scales and both of them are accessible within the HERAFitter package. The calculations are available to NLO and NNLO. In addition, a fast version of the scheme is available (i.e. RT FAST) by using the $k-factor$ technique where $k-factors$ are defined as the ratio between massless (accessed by QCDNUM) and massive scheme. The $k-factors$ are only calculated for the PDF parameters at the first fir iteration and are not anymore updated through the minimisation procedure, hence, the recommended is the full TR scheme.

2.1.2 GM-VFN ACOT scheme

The Aivazis-Collins-Olness-Tung scheme belongs to the group of VFN factorisation schemes that use the renormalization method of Collins-Wilczek-Zee (CWZ) [18]. This scheme involves a mixture of the $\overline{\text{MS}}$ scheme for light partons (and for heavy partons when the factorisation scale is larger than the heavy quark mass) and the zero-momentum subtraction renormalisation scheme for graphs with heavy quark lines (if the factorisation scale is smaller than the mass of the heavy quark threshold). The DGLAP kernels and PDF evolution are pure $\overline{\text{MS}}$, therefore, the ACOT scheme is considered to be a minimal extension of the $\overline{\text{MS}}$ scheme.

Within the ACOT package, different variants of the ACOT scheme are available: ACOT-Full, S-ACOT-χ, ACOT-ZM, $\overline{\text{MS}}$ at LO and NLO. For the longitudinal structure function higher order calculations are also available. The ACOT-Full implementation fully takes into account the quark masses and it reduces to ZM $\overline{\text{MS}}$ scheme in the limit of masses going to zero, but it has the disadvantage of being quite slow. Therefore the $k-factor$ technique has been adopted within the HERAFitter framework. The $k-factors$ can be defined in two different ways: as the ratio between same order calculations but massless vs massive (i.e. NLO (ZM-VFNS)/NLO (ACOT), or as the ratio between LO (massless)/NLO (massive). Both options are available in the HERAFitter package and they give similar results. For convergence of the k-factors usually $2-3$ repetitions of the fit are needed.

2.1.3 Fixed-Flavour Number Scheme

In the FFN 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. In HERAFitter this scheme can be accessed via the ABM [15] or QCDNUM implementation. The ABM implementation also includes the running mass definition of the heavy quark mass in the FFN scheme [19] which is realised via the interface to the open-source code OPENQCDRAD [15]. This 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. In QCDNUM, the calculation of the heavy quark contributions to DIS structure functions are available at NLO and only electromagnetic exchange contributions are taken into account. In the ABM implementation, the QCD corrections to the massive Wilson coefficients up to the currently best known approximate NNLO for the neutral-current (NC) heavy-quark production [20] and up to NLO for the charged-current (CC) case are available.

2.2 DIPOLE models

The dipole picture provides an alternative approach to the virtual photon-proton scattering at low x because it allows 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 [21]. The dipoles can be viewed 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 virtual photon fluctuates into a quark-antiquark pair and subsequently interacts with the target, and the dynamics of the interaction are embedded in the dipole scattering amplitude.

Several dipole models have been developed to describe various DIS reactions. They vary due to different assumption made about the behavior of the dipole-proton cross sections. In the HERAFitter three representative models are implemented: the Golec-Biernat-Wüsthoff (GBW) [22] dipole saturation model, the colour glass condensate approach to the high parton density regime Iancu-Itakura-Munier (IIM)

model [23] and a modified GBW model which takes into account the effects of DGLAP evolution Bartels-Golec-Kowalski (BGK) [24].

2.2.1 GBW model

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

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

where r corresponds to the transverse separation between the quark and the antiquark, and R_0^2 is an x dependent scale parameter which has corresponds to a saturation radius, $R_0^2(x) = (x/x_0)^{\lambda}$. The free fitted parameters are the cross-section normalisation σ_0 as well as x_0 and λ.

2.2.2 IIM model

The IIM model assumes an improved expression for the dipole cross section which is based on the Balitsky-Kovchegov equation [25]. The explicit formula for $\sigma_{\rm dip}$ can be found in [23]. The free fitted parameters are an alternative scale parameter $\tilde{R}$, x_0 and λ.

2.2.3 BGK model

The BGK model modifies the GBW model by taking into account the DGLAP evolution of the gluon density. The dipole cross section is given by

$$\sigma_{\rm dip}(x,r^2) = \sigma_0 \left(1 - \exp\left[-\frac{\pi^2 r^2 \alpha_s(\mu^2) x g(x,\mu^2)}{3\sigma_0}\right]\right).$$

The factorization scale μ^2 has the form $\mu^2 = C_{bgk}/r^2 + \mu_0^2$. In this model the gluon density, which is parametrized at some starting scale Q_0^2 by

$$x g(x,Q_0^2) = A_g x^{-\lambda_g}(1-x)^{C_g}.$$

is evolved to larger Q^2's using LO and NLO DGLAP evolution. The free fitted parameters for this model are σ_0, μ_0^2 and three parameters for the gluon density: A_g, λ_g, C_g. The parameter C_{bgk} is kept fixed: $C_{bgk} = 4.0$.

2.2.4 BGK model with valence quarks

The dipole models are valid in the low-x region only, where the valence quark contribution is small, of the order of 5%. The new HERA F_2 data have a precision which is better than 2 %. Therefore, in the `HERAFitter` the contribution of the valence quarks is taken from the PDF fits and added to the original BGK model, this is uniquely possible within the `HERAFitter` framework.

2.3 Transverse Momentum Dependent (unintegrated PDF) with CCFM

In this subsection another alternative approach to collinear DGLAP evolution is presented. In high energy factorization [26] generally the measured cross section is written as a convolution of the partonic cross section $\hat{\sigma}(k_t)$, which depends on the transverse momentum k_t of the incoming parton, with the k_t-dependent parton distribution function $\tilde{\mathcal{A}}(x,k_t,p)$ (transverse momentum dependent (TMD) or unintegrated uPDF):

$$\sigma = \int \frac{dz}{z} d^2k_t \hat{\sigma}(\frac{x}{z},k_t)\tilde{\mathcal{A}}(x,k_t,p) \quad (3)$$

The evolution of $\tilde{\mathcal{A}}(x,k_t,p)$ can proceed via the BFKL, DGLAP or via the CCFM evolution equations. In `HERAFitter` an extension of the CCFM [27, 28, 29, 30] evolution has been implemented. Since the evolution cannot be easily obtained in a closed form, first a kernel $\tilde{\mathcal{A}}(x'',k_t,p)$ is determined from the MC solution of the CCFM evolution equation, and is then folded with the non-perturbative starting distribution $\mathcal{A}_0(x)$ [31]:

$$\begin{aligned} x\mathcal{A}(x,k_t,p) &= x\int dx' \int dx'' \mathcal{A}_0(x)\tilde{\mathcal{A}}\left(x'',k_t,p\right)\delta(x'\cdot x''-x) \\ &= \int dx' \int dx'' \mathcal{A}_0(x)\tilde{\mathcal{A}}\left(x'',k_t,p\right)\frac{x}{x'}\delta(x''-\frac{x}{x'}) \\ &= \int dx' \mathcal{A}_0(x')\cdot\frac{x}{x'}\tilde{\mathcal{A}}\left(\frac{x}{x'},k_t,p\right) \end{aligned} \quad (4)$$

The kernel $\tilde{\mathcal{A}}$ includes all the dynamics of the evolution, Sudakov form factors and splitting functions and is determined in a grid of $50\otimes 50\otimes 50$ bins in x,k_t,p.

The calculation of the cross section according to Eq.(3) involves a multidimensional Monte Carlo integration which is time consuming and suffers from numerical fluctuations, and therefore cannot be used directly in a fit procedure involving the calculation of numerical derivatives in the search for a minimum. Instead the following procedure is applied:

$$\sigma_r(x,Q^2) = \int_x^1 dx_g \mathcal{A}(x_g,k_t,p)\hat{\sigma}(x,x_g,Q^2) \quad (5)$$

$$= \int_x^1 dx' \mathcal{A}_0(x')\cdot\tilde{\sigma}(x/x',Q^2) \quad (6)$$

The kernel $\tilde{\mathcal{A}}$ has to be provided separately and is not calculable within the program. A starting distribution $\mathcal{A}_0$, at the starting scale Q_0, of the following form is used:

$$x\mathcal{A}_0(x,k_t) = Nx^{-B_g}\cdot(1-x)^{C_g}(1-D_g x) \quad (7)$$

with free parameters N, B_g, C_g, D_g.

The calculation of the ep cross section follows eq.(3), with the off-shell matrix element including quarks masses taken from [26] in its implementation in `CASCADE` [32]. In addition to the boson gluon fusion process, valence quark initiated $\gamma q \to q$ processes are also included, with the valence quarks taken from [33].

2.4 Diffractive PDFs

It was observed at HERA that about 10% of deep inelastic interactions are diffractive leading to events in which the interacting proton stays intact ($ep \to eXp$). In the diffractive process the proton appears well separated from the rest of the hadronic final state by a large rapidity gap, in all other respects the events look similar to normal deep inelastic events. This process is usually interpreted as the diffractive dissociation of the exchanged virtual photon to produce a hadronic system X with mass much smaller than W and the same net quantum numbers as the exchanged photon. For this, the proton vertex factorisation approach is assumed such that the diffractive DIS is mediated by the exchange of hard Pomeron and a secondary Reggeon. The factorisable pomeron picture has proved remarkably successful for the description of most of these data.

In addition to x, Q^2 and the squared four-momentum transfer t (the undetected momentum transfer to the proton system), the mass M_X of the diffractively produced final state provides a further degree of freedom. In practice, the variable M_X is often replaced by β,

$$\beta = \frac{Q^2}{M_X^2 + Q^2 - t}. \tag{8}$$

In models based on a factorisable pomeron, β may be viewed 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 can be expressed as:

$$\frac{d\sigma}{d\beta\, dQ^2 dx_{IP}\, dt} = \frac{2\pi\alpha^2}{\beta Q^4}\left(1+(1-y)^2\right)\overline{\sigma}^{D(4)}(\beta, Q^2, x_{IP}, t) \tag{9}$$

where the "reduced cross-section" , $\overline{\sigma}$, is defined as

$$\overline{\sigma}^{D(4)} = F_2^{D(4)} - \frac{y^2}{1+(1-y)^2} F_L^{D(4)} = F_T^{D(4)} + \frac{2(1-y)}{1+(1-y)^2} F_L^{D(4)} \tag{10}$$

The dimension of $F_k^{D(4)}(\beta, Q^2, x_{IP}, t)$ is GeV^{-2} and thus quantities integrated over t.

$$F_k^{D(3)}(\beta, Q^2, x_{IP}) \equiv \int_{t_{\min}}^{t_{\max}} dt F_k^{D(4)}(\beta, Q^2, x_{IP}, t) \tag{11}$$

are dimensionless. The maximum kinematically allowed value of t is given by

$$t_{\mathrm{MAX}} = -\frac{x_{IP}^2 m_p^2 + p_\perp^2}{1 - x_{IP}} \approx -\frac{x_{IP}^2}{1 - x_{IP}} m_p^2 \tag{12}$$

where m_p is the proton mass. As $x = x_{IP}\beta$ we can normalize to the standard DIS formula

$$\frac{d\sigma}{d\beta\, dQ^2 dx_{IP}\, dt} = \frac{2\pi\alpha^2}{xQ^4}\left(1+(1-y)^2\right) x_{IP}\overline{\sigma}^{D(4)}(\beta, Q^2, x_{IP}, t) \tag{13}$$

which upon integration over t reads

$$\frac{d\sigma}{d\beta\, dQ^2\, dx_{IP}} = \frac{2\pi\alpha^2}{xQ^4}\left(1+(1-y)^2\right) x_{IP}\overline{\sigma}^{D(3)}(\beta, Q^2, x_{IP}). \tag{14}$$

The diffractive structure functions can be expressed as convolutions of the calculable coefficient functions with diffractive quark and gluon distribution functions, which in general depend on all of x_{IP}, Q^2, β, t.

Regge factorization For a better description of data, a contribution from a secondary Reggeon, *IR*, is included, hence

$$F_k^{D(4)}(\beta, Q^2, x_{IP}, t) = \sum_{\mathscr{X}=IP,IR} \phi_{\mathscr{X}}(x_{IP}, t)\, F_k^{\mathscr{X}}(\beta, Q^2) \tag{15}$$

or

$$F_k^{D(3)}(\beta, Q^2, x_{IP}) = \sum_{\mathscr{X}=IP,IR} \Phi_{\mathscr{X}}(x_{IP})\, F_k^{\mathscr{X}}(\beta, Q^2) \tag{16}$$

where

$$\Phi_{\mathscr{X}}(x_{IP}) = \int_{t_{\min}}^{t_{\max}} dt\, \phi_{\mathscr{X}}(x_{IP}, t). \tag{17}$$

The fluxes are parametrized as

$$\phi_{\mathscr{X}}(x_{IP}, t) = \frac{A_{\mathscr{X}}\, e^{b_{\mathscr{X}} t}}{x_{IP}^{2\alpha_{\mathscr{X}}(t)-1}} \tag{18a}$$

where

$$\alpha_{\mathscr{X}}(t) = \alpha_{\mathscr{X}}(0) + \alpha'_{\mathscr{X}} t. \tag{18b}$$

The function $F_k^{IR}(\beta, Q^2)$ is taken to be that of the pion.

2.5 Electroweak corrections for ep scattering

In order to compare the experimental data with theoretical predictions, QED corrections are necessary. In the `HERAFitter` the electroweak corrections for the DIS process are based on the EPRC package [34].

The calculations of higher-order electroweak corrections to DIS scattering at HERA are performed in the on-shell scheme where the gauge bosons masses M_W abd M_Z are treated symmetrically as basic parameters together with the top and Higgs masses, besides the fine structure constant α and other fermion masses.

The code provides the running of α using the most recent parametrisation of the hadronic contribution to Δ_α [35], as well as an older one from Burkhard [36].

2.6 Drell Yan processes

The calculations of the Drell Yan processes are known for many observables up to NNLO order. For example, there are packages such as FEWZ [37] and DYNNLO [38] for NNLO or MCFM [39] for NLO calculations. However, due to the complicated nature of these calculation involving an increased number of diagrams with each additional order, these calculations are too slow to be used iteratively in a fit. There are various methods to overcome this shortage: using the $k-factor$ approximation from lower to higher order, or using the so-called grid technique (storing the matrix elements on grids such that the cross sections maybe calculated later by convoluting these grids with the input PDFs) when available.

`HERAFitter` provides two implementations for pp Drell Yan processes. The first implementation uses calculations at LO which can be extended to NLO using k-factors, the second uses the APPLGRID interface.

The leading order Drell-Yan [40, 41] triple differential cross section in invariant mass M, boson rapidity y and CMS lepton scattering angle $\cos\theta$, for the neutral current, can be written as

$$\frac{\mathrm{d}^3\sigma}{\mathrm{d}M\mathrm{d}y\mathrm{d}\cos\theta} = \frac{\pi\alpha^2}{3MS}\sum_q P_q\left[F_q(x_1,Q^2)F_{\bar{q}}(x_2,Q^2)+(q\leftrightarrow\bar{q})\right], \tag{19}$$

where S is the squared CMS beam energy, $x_{1,2}=\frac{M}{\sqrt{S}}\exp(\pm y)$, $F_q(x_1,Q^2)$ is the parton number density, and

$$\begin{aligned} P_q &= e_l^2 e_q^2(1+\cos^2\theta) \\ &+ e_l e_q \frac{2M^2(M^2-M_Z^2)}{\sin^2\theta_W\cos^2\theta_W\left[(M^2-M_Z^2)^2+\Gamma_Z^2M_Z^2\right]} \\ &\left[aA_q(1+\cos^2\theta)+2bB_q\cos\theta\right] \\ &+ \frac{M^4}{\sin^4\theta_W\cos^4\theta_W\left[(M^2-M_Z^2)^2+\Gamma_Z^2M_Z^2\right]} \\ &\left[(a^2+b^2)(A_q^2+B_q^2)(1+\cos^2\theta)+8abA_qB_q\cos\theta\right]. \end{aligned} \tag{20}$$

Here θ_W is the Weinberg angle, M_Z and Γ_Z are Z boson mass and width, and

$$\begin{aligned} a &= -\frac{1}{4}+\sin^2\theta_W, \\ b &= -\frac{1}{4}, \\ A_q &= \frac{1}{2}I_q^3 - e_q\sin^2\theta_W, \\ B_q &= \frac{1}{2}I_q^3, \\ I_u^3 &= -I_d^3 = \frac{1}{2}, \\ e_l &= -1, e_u = \frac{2}{3}, e_d = -\frac{1}{3} \end{aligned} \tag{21}$$

give the electro-weak couplings.
The expression for charged current scattering has a simpler form.

$$\frac{\mathrm{d}^3\sigma}{\mathrm{d}M\mathrm{d}y\mathrm{d}\cos\theta} = \frac{\pi\alpha^2}{48S\sin^4\theta_W}\frac{M^3(1-\cos\theta)^2}{(M^2-M_W^2)+\Gamma_W^2M_W^2} \sum_{q_1,q_2} V_{q_1q_2}^2 F_{q_1}(x_1,Q^2)F_{q_2}(x_2,Q^2), \tag{22}$$

where $V_{q_1q_2}$ is the CKM quark mixing matrix and M_W and Γ_W are W boson mass and decay width.

The simple form of these expressions allows the calculation of integrated cross sections without utilization of Monte-Carlo techniques. This is particularly useful for PDF fitting purposes because statistical fluctuations are avoided in this case. In both neutral and charged current expressions the parton distribution functions factorise as functions dependent only on boson rapidity y and invariant mass M. The integral in $\cos\theta$ can be computed analytically and integrations in y and M can be performed with the Simpson method. The $\cos\theta$ parts are kept in the equation explicitly because their integration is asymmetric for data in lepton η bins and also because of the need to apply lepton $p_\perp$ cuts.

The fact that PDF functions factorise, allows high speed calculations when performing parameter fits over lepton rapidity data. In this case the factorised part of the expression which is independent of PDFs can be calculated only once for all minimisation iterations. The leading order code in `HERAFitter` package implements this optimisation and uses fast convolution routines provided by QCDNUM. Currently the full width LO calculations are optimised for lepton pseudorapidity and boson rapidity distributions with the possibility to apply lepton $p_\perp$ cuts. This flexibility allows the calculations to be performed within the phase space corresponding to the available measurement.

The calculated leading order cross sections are multiplied by $k-factors$ to obtain predictions at NLO or NNLO precision.

Alternatively, one can obtain the NLO predictions directly by using APPLGRID or FASTNLO techniques, which rely on the factorisation theorem by decoupling the hard scattering coefficients from PDFs. The hard scattering coefficients are calculated once and stored into a grid for a given kinematic bin, speeding up the convolution process with the PDFs and thus allowing to for fast QCD fits.

2.7 Cross Sections for $t\bar{t}$ production in pp or $p\bar{p}$ collisions

Top-quark pairs ($t\bar{t}$) are mainly produced at hadron colliders via gg fusion and $q\bar{q}$ annihilation. There are also qq' and qg production modes. The program HATHOR [42] allows the calculation of the expected total $t\bar{t}$ cross section at $p\bar{p}$ and pp colliders up to approximate NNLO accuracy. Version 1.3 of

HATHOR includes the exact NNLO for $q\bar{q} \to t\bar{t}$ [43] as well as a new high-energy constraint on the approximate NNLO calculation obtained from soft-gluon resummation [44]. The default choice for renormalization and factorization scale in $t\bar{t}$ production is the top-quark mass, m_t. The pole mass scheme is typically employed for m_t but HATHOR also supports calculations in the $\overline{\text{MS}}$ scheme.

2.8 Jet production

This sections presents various fast calculational techniques for jet production based on the factorization formalism.

Similarly to DY case, the calculation of higher order jet cross sections is very demanding in terms of computing power. Therefore. in order to enable the inclusion of jet-cross section measurements in PDF and α_s fits, the perturbative coefficients have to be pre-computed in a PDF and α_s independent way. For this purpose, two similar tools are interfaced to the `HERAFitter` .

2.8.1 FastNLO

The fastNLO project [45, 46, 47] enables the inclusion of jet data in PDF and α_s fits. This tool uses multi-dimensional interpolation techniques to convert the convolutions of perturbative coefficients with parton distribution functions and the strong coupling into simple products. The perturbative coefficients are calculated by the `NLOJET++` program [48] where calculations for jet-production in DIS [49] as well as in hadron-hadron collisions [50, 51] with threshold-corrections of $\mathcal{O}$(NNLO) for inclusive jet cross sections [52] are available.

The fastNLO libraries are included in the `HERAFitter` package and in order to include a new measurement into the PDF fit, only the fastNLO tables have to be specified. These tables include all necessary information about the perturbative coefficients and the calculated process for all bins of a certain dataset. Tables for almost all published jet measurements are available through the project website `http://fastnlo.hepforge.org`. The fastNLO tables are conventionally calculated for multiple factors of the factorization scale, and the renormalization scale factor can be chosen freely. Some of the fastNLO tables already allow for the free choice [47] of the renormalization and the factorization scale as a function of two pre-defined observables. The evaluation of the strong coupling constant, which enters the cross section calculation, is taken consistently from the QCDNUM evolution code.

2.8.2 APPLGRID

The APPLGRID [53] package allows the fast computation of NLO cross sections for particular processes for arbitrary sets of proton parton distribution functions. The package implements calculations of Drell Yan cross sections of electroweak boson (Z, W) production as well as jet production in proton-(anti)proton collisions and DIS processes.

The approach is based on storing the perturbative coefficients of NLO QCD calculations of final-state observables measured in hadron colliders in look-up tables. The PDFs and the strong couplings are included during the final calculations, e.g. during PDF fitting. The method allows variation of factorization and renormalization scales in calculations.

The look-up tables (grids) can be generated with modified versions of MCFM [54, 55] or NLOjet++ [51] software as distributed with the full version of APPLGRID package.

APPLGRID supports an interface to the MCFM parton level generators, hence model input parameters such as electroweak parameters are in fact pre-set following the MCFM input steering card, while binning and definitions of the observables for which the differential cross sections are needed are set in the APPLGRID code. The grid parameters, Q^2 binning and interpolation orders are also defined in the code.

APPLGRID constructs the grid tables in two steps: *(i)* exploration of the phase space in order to optimize the memory storage and *(ii)* actual grid construction in the phase space corresponding to the requested observables.

Afterwards the NLO cross sections are restored from the grids using externally provided PDFs, α_S, factorization and renormalization scales. QCD NNLO k-factors can be applied if requested.

3 Methodology

The methodology employed by HERAFitter provides, on one hand, faithful conditions adopted by different theoretical groups in extracting global PDF sets, and, on the other hand, presents a flexible framework to implement new theoretical developments for direct comparison.

The QCD fit formalism within HERAFitter gives access to different functional forms used to parametrise PDFs at the starting scale, different definitions for χ^2 to account for systematic uncertainties in extraction of theory parameters, different treatement of experimental uncertainties, and alternative theoretical models.

The speedy performance of such a complex framework is achieved by optimising the time of calculations using and incorporating innovative techniques such as cache option, fast evolution kernels, grid techniques making the platform a practical engine for iterative usage.

As an alternative to a complete QCD fit, reweighing method to estimate impact of new data is provided. This has been already advocated by the NNPDF collaboration [**?**] and HERAFitter can provide direct access to this method. The method has been extended to work not only on the replica method,

but also on the eigenvectors (as introduced by MSTW group [**?**]).

3.1 Chisquare representation

The PDF parameters are extracted through the χ^2 minimization method by employing MINUIT package linked to HERAFitter. There are various forms to represent the χ^2 form, i.e. covariance matrix or decomposed into nuisance parameters. In addition, there are various methods in dealing with the correlated systematic (or statistical) uncertainties. HERAFitter is prviding both options.

3.1.1 Covariance Matrix Representation

In the case of off-diagonal statistical uncertainties, the χ^2 function is

$$\chi^2_{\rm exp}(m,b) = \sum_{ij}\left(m^i - \sum_l \Gamma^i_l(m^i)b_l - \mu^i\right) C^{-1}_{{\rm stat.}\ ij}(m^i,m^j) \left(m^j - \sum_l \Gamma^j_l(m^j)b_l - \mu^j\right) + \sum_l b_l^2. \tag{23}$$

Here the scaling properties of the correlated systematic uncertainties Γ^i_j and of the covariance matrix $C_{{\rm stat.}\ ij}$ are expressed as a dependence on m_i and the dependence of $\delta_{\rm stat}$ on b_j is ignored.

3.1.2 Nuisance Parameters Representation

$$\chi^2_{\rm exp}(m,b) = \sum_i \frac{\left[m^i - \sum_j \gamma^i_j m^i b_j - \mu^i\right]^2}{\delta^2_{i,{\rm stat}}\mu^i\left(m^i - \sum_j \gamma^i_j m^i b_j\right) + \left(\delta_{i,{\rm uncor}}\, m^i\right)^2} + \sum_j b_j^2. \tag{24}$$

Here μ^i is the measured central value at a point i with relative statistical $\delta_{i,stat}$ and relative uncorrelated systematic uncertainty $\delta_{i,unc}$. Further, γ^i_j quantifies the sensitivity of the measurement μ^i at the point i to the systematic source j. The function $\chi^2_{\rm exp}$ depends on the set of underlying physical quantities m^i (denoted as the vector m) and the set of systematic uncertainties b_j (b). This definition of the χ^2 function takes into account that systematic uncertainties are proportional to the central values (multiplicative errors), whereas the statistical errors scale with the square roots of the expected number of events.

3.2 Treatment of Experimental Uncertainties

HERAFitter provides three methods in assessing the experimental uncertainties on PDFs: Hessian method, Offset Method, Monte Carlo method.

3.2.1 Hessian method

3.2.2 Offset method

3.2.3 Monte Carlo method

3.3 Treatment of Theoretical Input Parameters

The results of a QCD fit depends not only on the input data but also on the input theoretical ansatz, which is also uncertain. Nowadays, the modern PDFs try to address the impact of this ansatz on the resulting PDFs by assessing an uncertainty on the choice of the initial parameter, such as mass of charm m_c, mass of the bottom quarks m_b. Another important input is the choice of the functional form for the PDFs at the starting scale. For this, HERAFitter provides a series of choices ranging from simple functional forms to more complex forms such as Chebyshev Polynomials with larger flexibility. Larger flexibility usually requires some regularisation methods in order for the results to be physical.

3.4 Performance Optimisation

The above mentioned features make HERAFitter a powerful project that encapsulates state of the art developments from struggles on reaching atmost experimental precision to the state of the art theory developments.

4 Application of `HERAFitter`

The application of `HERAFitter` consists of:

- Description of product of the `HERAFitter` in terms of PDFGrids: HERA, LHC, LHeC.
- Application of `HERAFitter` in the QCD interpretation of data:
  - at HERA
  - at LHC (WZ, DY, jets..)
  - possibility to make impact studies for future colliders.

5 Summary

The new `HERAFitter` package has been presented

References

1. M. Botje (2010), http://www.nikef.nl/h24/qcdnum/index.html, [arXiv:1005.1481].
2. V. N. Gribov and L. N. Lipatov, Sov. J. Nucl. Phys. **15**, 438 (1972).
3. V. N. Gribov and L. N. Lipatov, Sov. J. Nucl. Phys. **15**, 675 (1972).
4. L. N. Lipatov, Sov. J. Nucl. Phys. **20**, 94 (1975).
5. Y. L. Dokshitzer, Sov. Phys. JETP **46**, 641 (1977).
6. G. Altarelli and G. Parisi, Nucl. Phys. B **126**, 298 (1977).
7. G. Curci, W. Furmanski, and R. Petronzio, Nucl.Phys. **B175**, 27 (1980).
8. W. Furmanski and R. Petronzio, Phys.Lett. **B97**, 437 (1980).
9. R. Demina, S. Keller, M. Kramer, S. Kretzer, R. Martin, *et al.* (1999), [hep-ph/0005112].
10. E. L. *et al.*, Phys. Lett. **B291**, 325 (1992).
11. E. L. *et al.*, Nucl. Phys. **B392**, 162, 229 (1993).
12. S. Riemersma, J. Smith, and van Neerven. W.L., Phys. Lett. **B347**, 143 (1995), [hep-ph/9411431].
13. R. S. Thorne and R. G. Roberts, Phys. Rev. D **57**, 6871 (1998), [hep-ph/9709442].
14. R. S. Thorne, Phys. Rev. **D73**, 054019 (2006), [hep-ph/0601245].
15. S. Alekhin, *OPENQCDRAD*, a program description and the code are available via: http://www-zeuthen.desy.de/∼alekhin/OPENQCDRAD.
16. A. D. Martin, Eur. Phys. J. C **63**, 189 (2009).
17. R. S. Thorne (2012), [arXiv:1201.6180].
18. J. C. Collins, Phys.Rev. **D58**, 094002 (1998), [hep-ph/9806259].
19. S. Alekhin and S. Moch, Phys. Lett. **B699**, 345 (2011), [arXiv:1011.5790].
20. K. H., N. Lo Presti, S. Moch, and A. Vogt, Nucl.Phys. **B864**, 399 (2012).
21. N. N. Nikolaev and B. Zakharov, Z.Phys. **C49**, 607 (1991).
22. K. Golec-Biernat and M. Wüsthoff, Phys. Rev. D **59**, 014017 (1999), [hep-ph/9807513].
23. E. Iancu, K. Itakura, and S. Munier, Phys. Lett. **B590**, 199 (2004), [hep-ph/0310338].
24. J. Bartels, K. Golec-Biernat, and H. Kowalski, Phys. Rev. D **66**, 014001 (2002), [hep-ph/0203258].
25. I. Balitsky, Nucl. Phys. B **463**, 99 (1996), [hep-ph/9509348].
26. S. Catani, M. Ciafaloni, and F. Hautmann, Nucl. Phys. B **366**, 135 (1991).
27. M. Ciafaloni, Nucl. Phys. B **296**, 49 (1988).
28. S. Catani, F. Fiorani, and G. Marchesini, Phys. Lett. B **234**, 339 (1990).
29. S. Catani, F. Fiorani, and G. Marchesini, Nucl. Phys. B **336**, 18 (1990).
30. G. Marchesini, Nucl. Phys. B **445**, 49 (1995).
31. H. Jung and F. Hautmann (2012), [arXiv:1206.1796].
32. H. Jung, S. Baranov, M. Deak, A. Grebenyuk, F. Hautmann, *et al.*, Eur.Phys.J. **C70**, 1237 (2010), [arXiv:1008.0152].
33. M. Deak, F. Hautmann, H. Jung, and K. Kutak, *Forward-Central Jet Correlations at the Large Hadron Collider* (2010), [arXiv:1012.6037].
34. H. Spiesberger, Private communication.
35. Jegerlehner, Proceedings, LC10 Workshop **DESY 11-117** (2011).
36. H. Burkhard, F. Jegerlehner, G. Penso, and C. Verzegnassi, in CERN Yellow Report on "Polarization at LEP" 1988.
37. Y. Li and F. Petriello, Phys.Rev. **D86**, 094034 (2012), [arXiv:1208.5967].
38. G. Bozzi, J. Rojo, and A. Vicini, Phys.Rev. **D83**, 113008 (2011), [arXiv:1104.2056].
39. A. Falkowski, M. L. Mangano, A. Martin, G. Perez, and J. Winter (2012), [arXiv:1212.4003].
40. S. D. Drell and T.-M. Yan, Phys. Rev. Lett. **25**, 316 (1970).
41. M. Yamada and M. Hayashi, Nuovo Cim. **A70**, 273 (1982).
42. M. Aliev, H. Lacker, U. Langenfeld, S. Moch, P. Uwer, *et al.*, Comput.Phys.Commun. **182**, 1034 (2011), [arXiv:1007.1327].
43. P. Bärnreuther, M. Czakon, and A. Mitov (2012), [arXiv:1204.5201].
44. S. Moch, P. Uwer, and A. Vogt, Phys.Lett. **B714**, 48 (2012), [hep-ph/1203.6282].
45. T. Kluge, K. Rabbertz, and M. Wobisch, pp. 483–486 (2006), [hep-ph/0609285].
46. M. Wobisch, D. Britzger, T. Kluge, K. Rabbertz, and F. Stober [fastNLO Collaboration] (2011), [arXiv:1109.1310].
47. D. Britzger, K. Rabbertz, F. Stober, and M. Wobisch [fastNLO Collaboration] (2012), [arXiv:1208.3641].
48. Z. Nagy and Z. Trocsanyi, Phys.Rev. **D59**, 014020 (1999), [hep-ph/9806317].
49. Z. Nagy and Z. Trocsanyi, Phys.Rev.Lett. **87**, 082001 (2001), [hep-ph/0104315].
50. Z. Nagy, Phys.Rev. **D68**, 094002 (2003), [hep-ph/0307268].
51. Z. Nagy, Phys.Rev.Lett. **88**, 122003 (2002), [hep-ph/0110315].
52. N. Kidonakis and J. Owens, Phys.Rev. **D63**, 054019 (2001), [hep-ph/0007268].
53. T. Carli *et al.*, Eur. Phys. J. **C66**, 503 (2010), [arXiv:0911.2985].

54. J. M. Campbell and R. K. Ellis, Phys. Rev. **D60**, 113006 (1999), [arXiv:9905386].
55. J. M. Campbell and R. K. Ellis, Nucl. Phys. Proc. Suppl. **205-206**, 10 (2010), [arXiv:1007.3492].