HERAFitter - an open source QCD fit framework

Andrey Sapronov
on behalf of the `HERAFitter` team

Joint Institute for Nuclear Research, 141980 Russia, Dubna, Moscow Region, Joliot-Curie 6.

E-mail: `Andrey.Sapronov@cern.ch`

Abstract. The open source project `HERAFitter` was established to increase applicability of the QCD analysis in the hadron collider experiments. The framework may be used to extract parton density functions from a variety of experimental measurements and to assess the impact of different data on the parton density determination. It may also be employed to perform data consistency checks and to test theoretical models.
This short article covers a parallel session contribution of the ACAT2014 conference.

1. Introduction

In the era of hadron colliders the proton parton density functions (PDF) have been gaining increasing attention. While the parton interactions within a hadron cannot be duly described using perturbative approximation, the factorisation theorem helps isolating the hard scattering cross section components which can be determined empirically. Improving precision of PDFs is important to reduce the uncertainty of theoretical predictions for both Standard Model (SM) and new physics processes. The `HERAFitter` [1] is an open source project created to facilitate QCD analysis of experimental data and to improve the current knowledge of PDF.

The PDF extraction procedure in QCD analysis implies fitting the experimental data with theoretical predictions based on parametrised partonic distributions at a starting scale which are then evolved to the scale of the measurements using DGLAP [2, 3, 4, 5, 6] evolution equations. The treatment of the uncertainties and accounting for possible correlations of the precision data is of high importance in order to improve precision of theoretical predictions for high energy processes.

Several groups are currently involved in the PDF fitting analysis: MSTW [7], CT(EQ) [8], NNPDF [9], AB(K)M [10], GJR [11] and HERAPDF [12]. Their methods and distinguishing features are reviewed in [13, 14]. In short, all groups perform global fits of the deep inelastic scattering (DIS), fixed target and $pp(\bar{p})$ collider data, except HERAPDF which relies purely on the HERA data and hence justified usage of the χ^2 tolerance of $\Delta\chi^2 = 1$. The NNPDF approach is notable for their analysis being performed via neural network training in association with a Monte Carlo method to estimate their uncertainties and minimise the input parametric form dependence on PDFs. There are other various differences among PDF groups: higher orders counting, heavy flavour corrections, treatment of α_S, etc.

The interest of achieving a better level of agreement among these PDF groups, as well as an overall better understanding of the proton's PDF is driven by the recent results from the LHC.

2. Overview of the HERAFitter program

`HERAFitter` is an open source QCD fitting framework aimed to provide means of experimental data QCD analysis and extraction of PDFs. The program is designed with a user friendly but highly configurable interface and includes a large number of theoretical models. It allows extensive study of the impact of experimental data. A detailed information about the project status and available releases with documentation can be found in `http://herafitter.org`.

The analysis is based on the χ^2 estimate of how well the theoretical prediction describe experimental data for a given PDF parametrisation. Various ansatz on PDF parametric form can be tested within this framework. The framework is designed such that a user can plug in his own sets of data with corresponding theory prediction, adjust input PDF parametrisation form, select a specific desired scheme for matching the heavy quark thresholds, a calculation order, cuts, etc. The QCD fit output will provide a quantitative evaluation of level of agreement between data and theory, and if desired, the resulting PDF grids in LHAPDF format can be obtained. The package contains elaborate plotting tools for prompt inspection of the results.

The overall analysis approach is schematically shown in the figure 1. The procedure starts with reading the input parameters and definition of initial parametrisation at the starting scale Q_0^2. Initial scale is chosen to be below the charm mass threshold, as evolution is performed starting from a three active quarks as required by the QCDNUM [15] program, which performs the QCD evolution. The minimisation of the χ^2 is performed via the MINUIT [16] package.

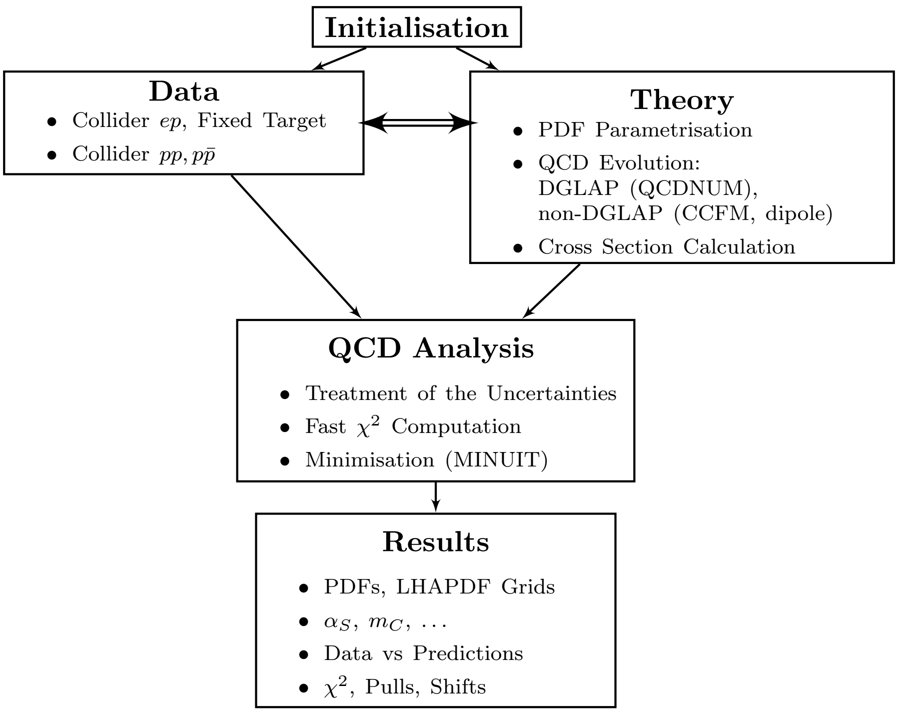

Figure 1. Schematic structure of the `HERAFitter` program.

The experimental data are introduced as a text table which format is defined in a header needed to be specified by the user in the same file. The header includes e.g. reaction name, data table format and theory parameters. The `HERAFitter` program features elaborate correlated data uncertainty treatment which can be presented in various forms: as a covariance matrix, as a table with nuisance parameters or a combination of both. Depending on the choice, corresponding χ^2 definition will be used for minimisation procedure. For a single data set the χ^2 function is defined in a form

$$\chi^2 = \sum_i \frac{\left[m^i - \sum_j \gamma^i_j m^i b_j - \mu^i\right]^2}{\left(\delta_{i,\mathrm{stat}} m^i\right)^2 + \left(\delta_{i,\mathrm{uncor}}\, m^i\right)^2} + \sum_j b_j^2, \qquad (1)$$

Here μ^i is the measured central value at a point i with relative statistical $\delta_{i,stat}$ and relative uncorrelated systematic uncertainty $\delta_{i,unc}$, the quantity m^i is the theoretical prediction, γ^i_j represents relative correlated systematic uncertainties and b_j their shifts. In the case of the covariance matrix, the χ^2 function takes the form

$$\chi^2 = \sum_{ij} (m^i - \mu^i) C^{-1}_{\mathrm{tot}\ ij} (m^j - \mu^j), \qquad (2)$$

where the $C^{-1}_{\mathrm{tot}\ ij}$ is the total covariance matrix given by the sum of the statistical and systematic covariance matrices. Due to flexible implementation in the framework, different approaches can be combined together, e.g. some dataset's systematic uncertainties can be treated using the matrix method, others can be treated using nuisance parameters.

Various types of uncertainties can be handled using following implemented methods: hessian [17], Monte Carlo [18, 19] or offset [20]. The hessian method is a standard way of computing the eigenvectors corresponding to the uncertainty sources by diagonalizing a matrix of second derivatives of the χ^2 distribution. In Monte Carlo method the PDF uncertainties are estimated using pseudo-data replicas randomly generated from the measurement's central values and their systematic and statistical uncertainties. The offset method consists in constraining the PDFs in a flexible parametrisation style.

The current version of the `HERAFitter` framework allows fitting the following types of data: inclusive cross section measurements from HERA DIS and fixed target experiments and Drell–Yan, jet and single-top quark production data from ep, pp and $p\bar{p}$ colliders.

The DIS structure functions may be computed directly in a variety of heavy quark schemes including the fixed-flavour (FFN) and variable flavour number (VFN) schemes. VFN schemes with various treatments for the heavy quark thresholds include the Thorne Roberts (TR) scheme at LO, NLO and NNLO [21, 22] as provided by the MSTW group, the ACOT scheme at LO and NLO as provided by the CTEQ group. The QCDNUM also provides the calculations of the DIS structure functions in the zero-mass VFN and FFN schemes. The FFN scheme is alternatively available via the OPENQCDRAD [23] interface with running mass definition [24].

Theoretical predictions for other processes are provided via the interfaces to fast cross section estimation tools, such as APPLGrid [25] or FastNLO [26]. Evaluation of the differential (or total) cross section is performed by convoluting the parametrised PDFs with pregenerated tables of higher order coefficients (NLO or NNLO). Further corrections may be applied via K-factor tables, for example NNLO QCD computed with FEWZ [27] or NLO EW obtained from SANC [28, 29]. Drell–Yan, jet production in pp, $p\bar{p}$ collisions are the processes with fast cross section estimates available. The HATHOR [30] program interfaced in the `HERAFitter` can be used to calculate complete NNLO total cross section of $t\bar{t}$ production in pp, $p\bar{p}$ collisions while differential NLO cross section for this process is available from MCFM [31, 32] via APPLGrid or FastNLO.

3. Application in QCD analysis of experimental data

The `HERAFitter` program has been 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 [33, 34, 35, 36, 37], inclusive jet production [38], and inclusive photon production [39]. One of the notable results is the strange quark density measurement from 2010 data on single gauge boson production performed by the ATLAS experiment [33]. The analysis had the

HERA-I *ep* data as a basis with additional 30 data points of the ATLAS Drell–Yan measurement fitted with QCD NNLO predictions. The partial χ^2 of the ATLAS data shows significant improvement from 44.5 when the $\bar{s}$ density was tied with the $\bar{d}$-quark to 33.9 when it was released. Figure 2 shows example plots for gluon density measured with ATLAS inclusive jets data and $r_s = \bar{s}/\bar{d}$ value measured from muon charge asymmetry at the CMS experiment. The results of QCD analyses using `HERAFitter` were also published by HERA experiments for inclusive [12, 40] and heavy flavour production measurements [41, 42]. The following phenomenological studies have been performed with `HERAFitter`: a determination of the transverse momentum dependent gluon density using precision HERA data [43], an analysis of HERA data within a dipole model [44], the study of the low-x uncertainties in PDFs determined from the HERA data using different parametrisations [45] and the impact of QED radiative corrections on PDFs [46]. A recent study based on a set of PDFs determined with the `HERAFitter` and addressing the correlated uncertainties between different orders has been published in [47].

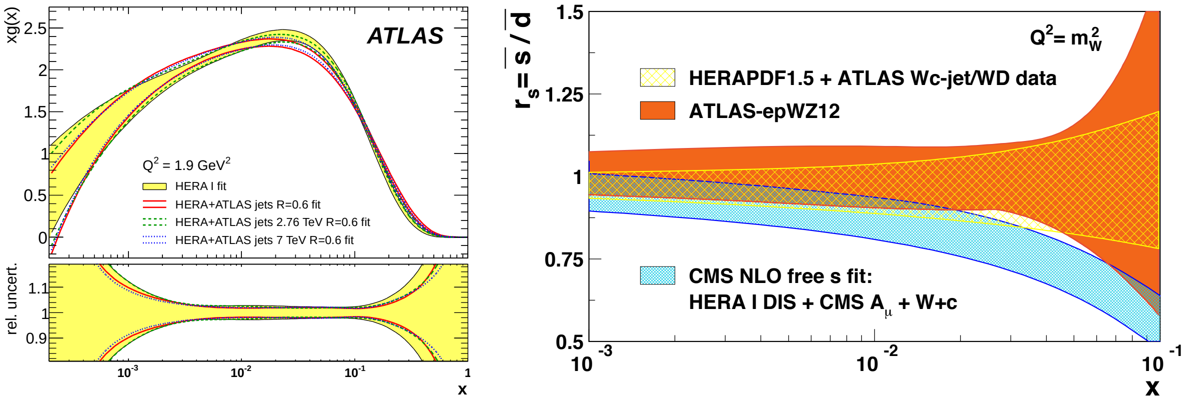

Figure 2. LHC analysis results, obtained with `HERAFitter` program: (*left*) gluon momentum distribution, measured in the ATLAS inclusive jets analysis for $\sqrt{s} = 2.76$ TeV and $\sqrt{s} = 7$ TeV data [38] and (*right*) $r_s = \bar{s}/\bar{d}$ value measurement by the CMS experiment [34].

The `HERAFitter` framework has been used to produce PDF grids from QCD analyses performed at HERA [12, 48] and at the LHC [49], using measurements from ATLAS [33, 38]. These PDFs can be used to calculate predictions for SM or beyond SM processes. Furthermore, `HERAFitter` provides the possibility to perform various benchmarking exercises [50] and impact studies for possible future colliders as demonstrated by QCD studies at the LHeC [51].

4. Summary
The `HERAFitter` project provides diverse means for QCD interpretation of the experimental measurements at hadron colliders. The source code of the package is freely available under GPL v3 license. The framework design allows robust extraction of proton PDFs and assessment of the data impact on the current PDF determination. The package contains necessary ingredients to study the PDFs incorporating a variety of different data processes and theory calculations and is an optimal platform for QCD related benchmarks.

References
[1] HERAFitter https://www.herafitter.org
[2] Gribov V N and Lipatov L N 1972 *Sov. J. Nucl. Phys.* **15** 438
[3] Gribov V N and Lipatov L N 1972 *Sov. J. Nucl. Phys.* **15** 675
[4] Lipatov L N 1975 *Sov. J. Nucl. Phys.* **20** 94
[5] Dokshitzer Y L 1977 *Sov. Phys. JETP* **46** 641

[6] Altarelli G and Parisi G 1977 *Nucl. Phys. B* **126** 298
[7] Martin A, Stirling W, Thorne R and Watt G 2009 *Eur.Phys.J.* **C63** 189–285 (*Preprint* `0901.0002`)
[8] Gao J, Guzzi M, Huston J, Lai H L, Li Z *et al.* 2014 *Phys.Rev.* **D89** 033009 (*Preprint* `1302.6246`) URL `http://hep.pa.msu.edu/cteq/public/`
[9] Ball R D, Bertone V, Carrazza S, Deans C S, Del Debbio L *et al.* 2013 *Nucl.Phys.* **B867** 244–289 (*Preprint* `1207.1303`) URL `https://nnpdf.hepforge.org/`
[10] Alekhin S, Blümlein J and Moch S 2013 (*Preprint* `1310.3059`)
[11] Jimenez-Delgado P and Reya E 2009 *Phys.Rev.* **D80** 114011 (*Preprint* `0909.1711`) URL `http://www.het.physik.tu-dortmund.de/pdfserver/index.html`
[12] Aaron F *et al.* (H1 and ZEUS Collaborations) 2010 *JHEP* **1001** 109 (*Preprint* `arXiv:0911.0884`)
[13] De Roeck A and Thorne R 2011 *Prog.Part.Nucl.Phys.* **66** 727–781 (*Preprint* `1103.0555`)
[14] Watt G 2011 *JHEP* **1109** 069 (*Preprint* `1106.5788`)
[15] Botje M 2011 *Comput.Phys.Commun.* **182** 490–532 (*Preprint* `1005.1481`)
[16] James F and Roos M 1975 *Comput.Phys.Commun.* **10** 343–367
[17] Pumplin J, Stump D, Brock R, Casey D, Huston J *et al.* 2001 *Phys.Rev.* **D65** 014013 (*Preprint* `hep-ph/0101032`)
[18] Giele W T and Keller S 1998 *Phys.Rev.* **D58** 094023 (*Preprint* `hep-ph/9803393`)
[19] Giele W T, Keller S A and Kosower D A 2001 (*Preprint* `hep-ph/0104052`)
[20] Botje M 2002 *J.Phys.* **G28** 779–790 (*Preprint* `hep-ph/0110123`)
[21] Thorne R S and Roberts R G 1998 *Phys. Rev. D* **57** 6871 (*Preprint* `hep-ph/9709442`)
[22] Thorne R S 2006 *Phys. Rev.* **D73** 054019 (*Preprint* `hep-ph/0601245`)
[23] Alekhin S, Blümlein J and Moch S OPENQCDRAD a program description and the code are available via: http://www-zeuthen.desy.de/∼alekhin/OPENQCDRAD
[24] Alekhin S and Moch S 2011 *Phys. Lett.* **B699** 345 (*Preprint* `arXiv:1011.5790`)
[25] Carli T *et al.* 2010 *Eur. Phys. J.* **C66** 503–524 (*Preprint* `arXiv:0911.2985`)
[26] Kluge T, Rabbertz K and Wobisch M 2006 483–486 (*Preprint* `hep-ph/0609285`)
[27] Li Y and Petriello F 2012 *Phys.Rev.* **D86** 094034 (*Preprint* `arXiv:1208.5967`)
[28] Andonov A, Arbuzov A, Bardin D, Bondarenko S, Christova P *et al.* 2006 *Comput.Phys.Commun.* **174** 481–517 (*Preprint* `hep-ph/0411186`)
[29] Bondarenko S G and Sapronov A A 2013 *Comput.Phys.Commun.* **184** 2343–2350 (*Preprint* `1301.3687`)
[30] Aliev M, Lacker H, Langenfeld U, Moch S, Uwer P *et al.* 2011 *Comput.Phys.Commun.* **182** 1034–1046 (*Preprint* `arXiv:1007.1327`)
[31] Campbell J M and Ellis R K 2010 *Nucl. Phys. Proc. Suppl.* **205-206** 10–15 (*Preprint* `arXiv:1007.3492`)
[32] Campbell J M and Ellis R K 2012 Top-quark processes at NLO in production and decay report FERMILAB-PUB-12-078-T (*Preprint* `1204.1513`)
[33] Aad G *et al.* (ATLAS Collaboration) 2012 *Phys.Rev.Lett.* **109** 012001 (*Preprint* `1203.4051`)
[34] Chatrchyan S *et al.* (CMS Collaboration) 2014 *Phys.Rev.* **D90** 032004 (*Preprint* `1312.6283`)
[35] Aad G *et al.* (ATLAS Collaboration) 2013 *Phys. Lett.* **B725** 223 (*Preprint* `arXiv::1305.4192`)
[36] Aad G *et al.* (ATLAS Collaboration) 2014 *JHEP* **1406** 112 (*Preprint* `1404.1212`)
[37] Aad G *et al.* (ATLAS Collaboration) 2014 *JHEP* **1405** 068 (*Preprint* `1402.6263`)
[38] Aad G *et al.* (ATLAS Collaboration) 2013 *Eur.Phys.J.* **73** 2509 (*Preprint* `arXiv:1304:4739`)
[39] Aad G *et al.* (ATLAS Collaboration) 2013 A study of the sensitivity to the proton parton distributions of the inclusive photon production cross section in *pp* collisions at 7 TeV measured by the ATLAS experiment at the LHC Tech. Rep. ATL-PHYS-PUB-2013-018 CERN Geneva
[40] Aaron F *et al.* (H1 Collaboration) 2012 *JHEP* **1209** 061 (*Preprint* `arXiv:1206.7007`)
[41] Abramowicz H *et al.* (H1 and ZEUS Collaborations) 2013 *Eur. Phys. J.* **C73** 2311 (*Preprint* `arXiv:1211.1182`)
[42] Abramowicz H *et al.* (ZEUS Collaboration) 2014 (*Preprint* `1405.6915`)
[43] Hautmann F and Jung H 2014 *Nuclear Physics B* **883** 1 – 19 (*Preprint* `1312.7875`)
[44] Luszczak A and Kowalski H 2014 *Phys.Rev.* **D89** 074051 (*Preprint* `1312.4060`)
[45] Glazov A, Moch S and Radescu V 2011 *Phys. Lett. B* **695** 238–241 (*Preprint* `arXiv:1009.6170`)
[46] Sadykov R 2014 (*Preprint* `1401.1133`)
[47] HERAFitter Developers Team and Lisovyi M 2014 (*Preprint* `arXiv:1404.4234`)
[48] H1prelim-13-141 and ZEUS-prel-13-003, H1prelim-10-142 and ZEUS-prel-10-018, H1prelim-11-042 and ZEUS-prel-11-002 HERAPDF1.5LO, NLO and NNLO available via: http://lhapdf.hepforge.org/pdfsets
[49] ATLAS NNLO epWZ12 available via: http://lhapdf.hepforge.org/pdfsets
[50] Butterworth J, Dissertori G, Dittmaier S, de Florian D, Glover N *et al.* 2014 (*Preprint* `1405.1067`)
[51] Abelleira Fernandez J L *et al.* (LHeC Study Group) 2012 *Journal of Phys.* **G** 075001 (*Preprint* `arXiv:1206.2913`)