# xFitter (former HERAFitter) Project

## Open Source QCD Fit framework

Ringailė Plačakytė

**University of Hamburg**

on behalf of the xFitter team

# Introduction

Precise knowledge of the PDFs are essential for predictions at hadron colliders

**QCD factorisation:**

$$\sigma \approx \hat{\sigma} \otimes PDF$$

**Experimental Data:**

→ a large variety of data from fixed-target and collider experiments

**Theory:**

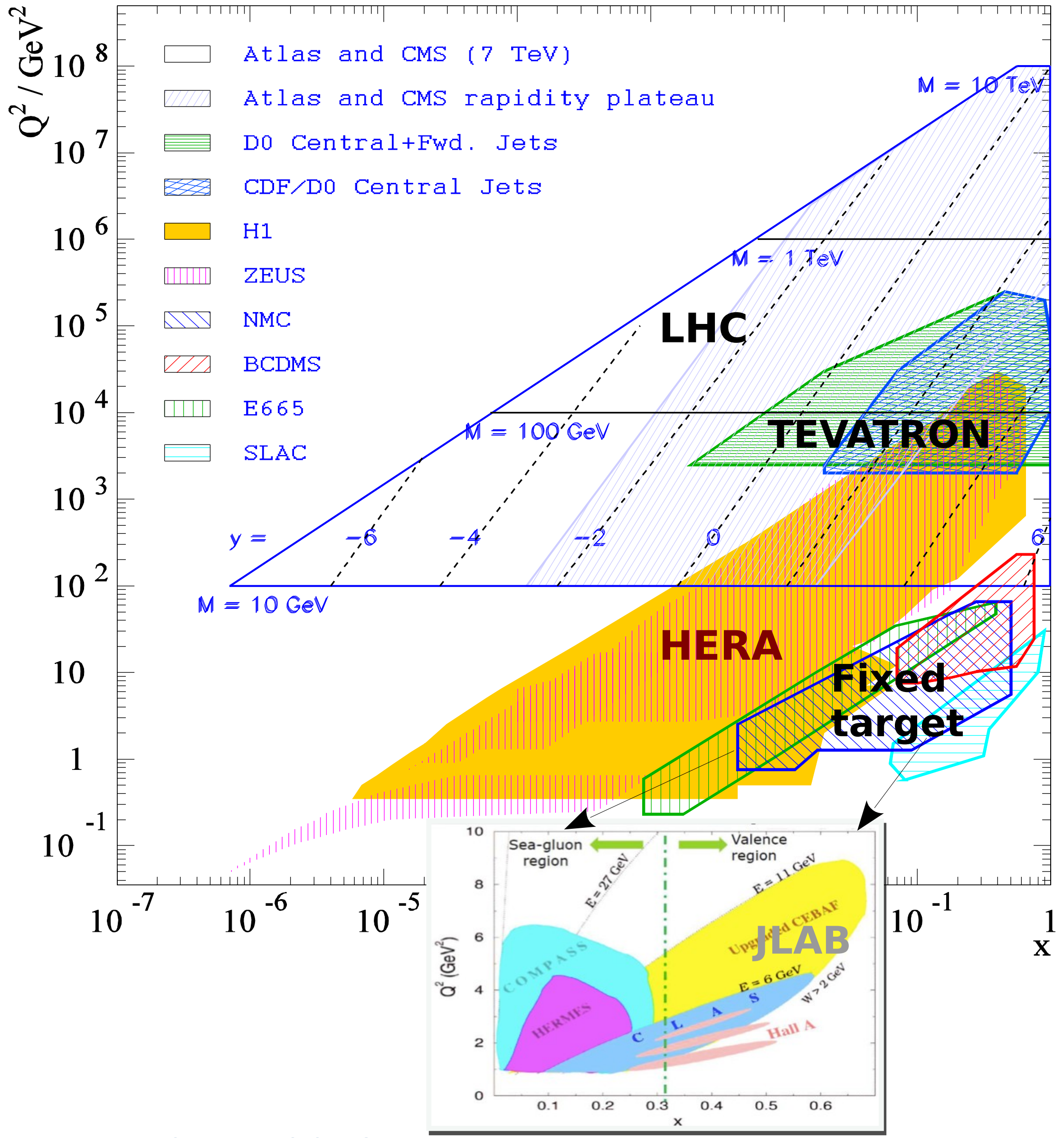
→ intense theoretical developments

**QCD Analysis:**

→ available (unpolarised) PDFs:
CT/CJ,MMHT,NNPDF,ABM,HERAPDF,JR

… and **Tools**

kinematic plane of the experimental data (in x,$Q^2$)



# Tools for PDFs

Available (open-source) tools for the PDF determination/plotting:

xFitter (former HERAFitter): an open-source package that provides a framework for the determination of the PDFs of the proton and for many different kinds of analyses in QCD

EPJC (2015), 75: 304, *xfitter.org*

OPENQCDRAD (ABM collaboration: numerical computation of all hard scattering cross sections (DIS structure function calculation including heavy quark contributions, W and Z production)

PRD86 (2012) 054009, *www-zeuthen.desy.de/˜alekhin/OPENQCDRAD*

ALPOS: an object-oriented data to theory comparison and fitting tool (profit from and exchange with xFitter experience). Access from a public svn repository (via request)

http://desy.de/~britzger/alpos/

QCDNUM: Fast QCD Evolution and Convolution (numerically solves the evolution equations for parton densities and fragmentation functions in pQCD)

Comp.Phys.Com.182:490,2011

APFEL (used by NNPDF): a PDF evolution library, is a computer library specialized in the solution of DGLAP evolution equations up to NNLO in QCD and to LO in QED

Valerio Bertone's talk

arXiv.1310.1394, *apfel.hepforge.org*

TMDlib/TMDplotter: library and plotting tools for Transverse Momentum Dependent parton distributions

Patrick Connor's talk

EPJC 74 (2014) 3220, tmdlib.hepforge.org

# xFitter Project

www.xfitter.org

## 2011 Open Source Revolution:

first open source QCD Fit Platform which started the wave of sharing QCD fit codes

EPJC (2015), 75: 304

- A team of ~30 developers:
  - LHC/HERA/theory/independent
  - several releases since 2011
  - 33 publications that have used the framework [in total]

synergy between experiment and theory groups

**provides a unique QCD framework to address theoretical differences:**

→ benchmark exercises/collaborative efforts/topical studies

**provides means to the experimentalists to optimise the measurements:**

→ assess impact/consistency of new data

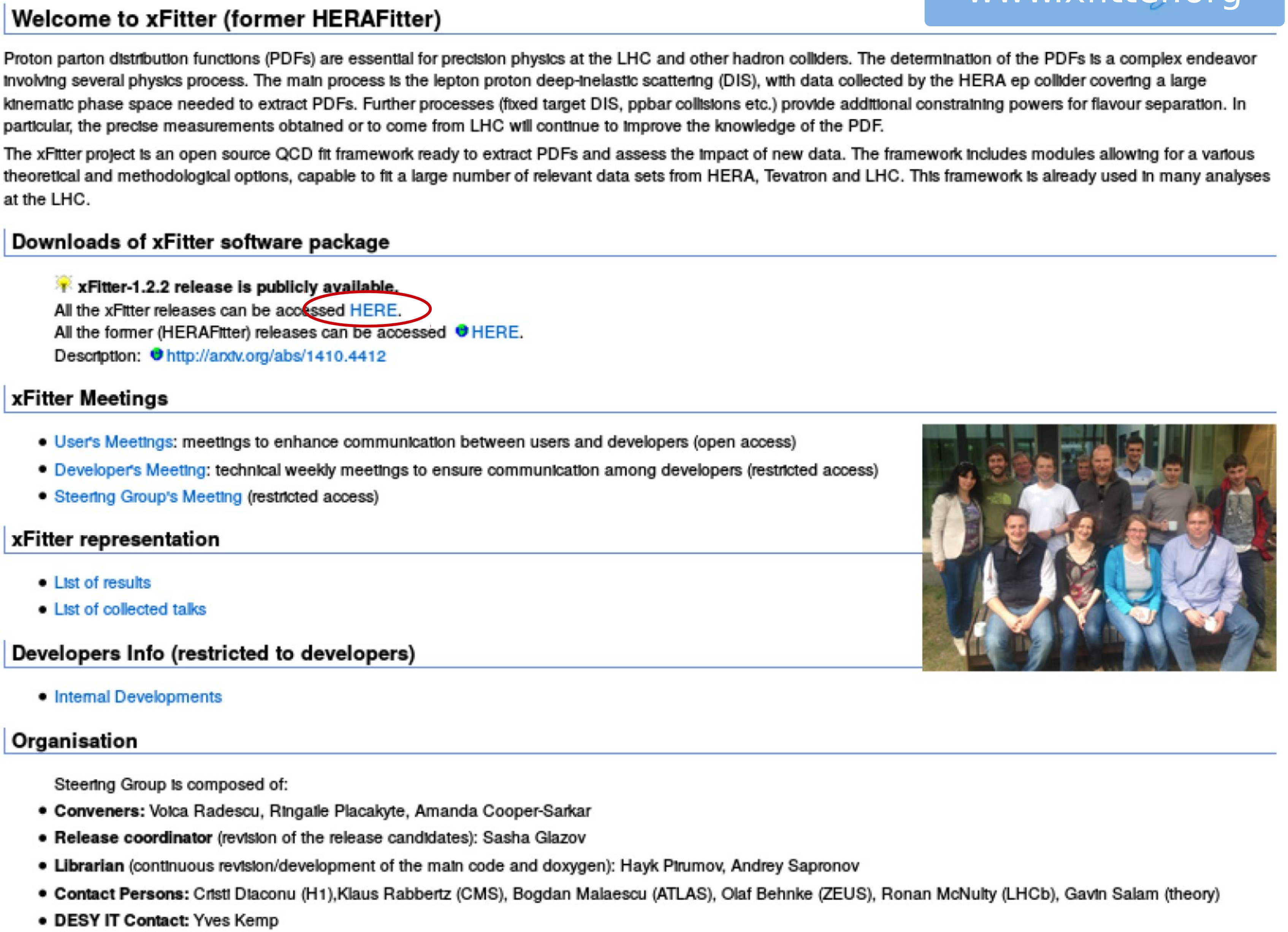
## Dedicated studies [xFitter developers team]:

- method in preserving correlation between PDFs extracted at different orders in pQCD
- address consistency of Tevatron measurement and evaluate their collective impact on valence
- determination of the running mass in $\overline{\text{MS}}$ scheme

# xFitter Project

→ open access, no registration required

www.xfitter.org



## Welcome to xFitter (former HERAFitter)

Proton parton distribution functions (PDFs) are essential for precision physics at the LHC and other hadron colliders. The determination of the PDFs is a complex endeavor involving several physics process. The main process is the lepton proton deep-inelastic scattering (DIS), with data collected by the HERA ep collider covering a large kinematic phase space needed to extract PDFs. Further processes (fixed target DIS, ppbar collisions etc.) provide additional constraining powers for flavour separation. In particular, the precise measurements obtained or to come from LHC will continue to improve the knowledge of the PDF.

The xFitter project is an open source QCD fit framework ready to extract PDFs and assess the impact of new data. The framework includes modules allowing for a various theoretical and methodological options, capable to fit a large number of relevant data sets from HERA, Tevatron and LHC. This framework is already used in many analyses at the LHC.

## Downloads of xFitter software package

**xFitter-1.2.2 release is publicly available.**
All the xFitter releases can be accessed HERE.
All the former (HERAFitter) releases can be accessed HERE.
Description: http://arxiv.org/abs/1410.4412

## xFitter Meetings

- User's Meetings: meetings to enhance communication between users and developers (open access)
- Developer's Meeting: technical weekly meetings to ensure communication among developers (restricted access)
- Steering Group's Meeting (restricted access)

## xFitter representation

- List of results
- List of collected talks

## Developers Info (restricted to developers)

- Internal Developments

## Organisation

Steering Group is composed of:

- **Conveners:** Voica Radescu, Ringaile Placakyte, Amanda Cooper-Sarkar
- **Release coordinator** (revision of the release candidates): Sasha Glazov
- **Librarian** (continuous revision/development of the main code and doxygen): Hayk Pirumov, Andrey Sapronov
- **Contact Persons:** Cristi Diaconu (H1),Klaus Rabbertz (CMS), Bogdan Malaescu (ATLAS), Olaf Behnke (ZEUS), Ronan McNulty (LHCb), Gavin Salam (theory)
- **DESY IT Contact:** Yves Kemp

# Schematic View of the xFitter Program

## Main steps in QCD analysis:

Parametrise PDFs at the starting scale

→ multiple options for functional forms
  → Standard Polynomial, Chebyshev, etc

Evolve to the scale corresponding to data point

→ QCD(DGLAP) evolution codes [QCDNUM, APFEL]
→ TMDs, kt ordered evolution, dipole models, QCD(DGLAP)+QED

Calculate the cross section

Voica's talk

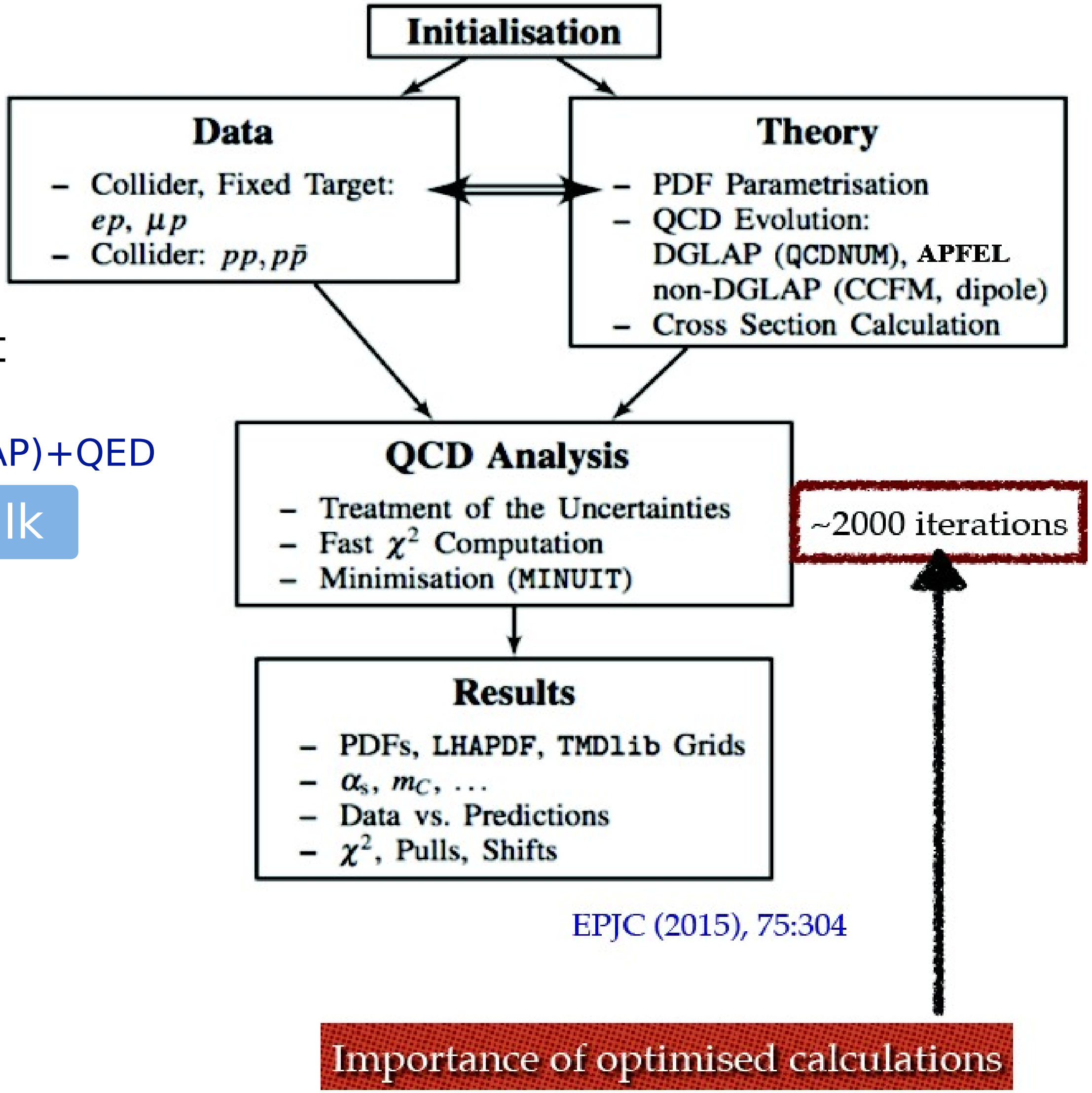
→ various heavy flavour schemes:
  → RT, ACOT, FONLL, FFNS(ABM)
→ fast grid techniques interfaced to DY:
  → APPLGRID, FASTNLO, APFELgrid

Compare with data via $\chi^2$:

→ multiple forms to account for correlations

Minimize $\chi^2$ with respect to PDF parameters

→ profiling, reweighting
→ fit: MINUIT, data driven regularisation



**Initialisation**

**Data**
- Collider, Fixed Target: $ep$, $\mu p$
- Collider: $pp, p\bar{p}$

**Theory**
- PDF Parametrisation
- QCD Evolution: DGLAP (`QCDNUM`), **APFEL** non-DGLAP (CCFM, dipole)
- Cross Section Calculation

**QCD Analysis**
- Treatment of the Uncertainties
- Fast $\chi^2$ Computation
- Minimisation (`MINUIT`)

~2000 iterations

**Results**
- PDFs, `LHAPDF`, `TMDlib` Grids
- $\alpha_s$, $m_C$, ...
- Data vs. Predictions
- $\chi^2$, Pulls, Shifts

EPJC (2015), 75:304

Importance of optimised calculations

# Measure the goodness of the fit: $\chi^2$ function

The goodness of the fit is measured by $\chi^2$ in PDFs

→ various forms are implemented in xFitter:

**Standard $\chi^2$ function**

→ with nuisance parameters:

measurement

$$\chi^2 = \sum_i \frac{(\mu_i - \hat{m}_i)^2}{\Delta_i^2} + \sum_\alpha b_\alpha^2$$

uncorrelated error

sum over correlated systematic sources

theory

nuisance parameter

$$\hat{m}_i = m_i + \sum_\alpha \Gamma_{i\alpha}\, b_\alpha$$

correlated error

→ correlations in the covariance matrix:

$$\chi^2 = (\boldsymbol{\mu} - \boldsymbol{m})^{\mathrm{T}}\, C^{-1}\, (\boldsymbol{\mu} - \boldsymbol{m})$$

$$= \sum_{ij} (\mu_i - m_i)\, C_{ij}^{-1} (\mu_j - m_j)$$

statistical

uncorrelated

correlated

$$C = C^{\mathrm{stat}} + C^{\mathrm{unc}} + C^{\mathrm{syst}}$$

→ mixed form:

$$\chi^2 = \sum_{ij}^{N} \left( D_i - T_i - \sum_k^K r_k \beta_{ik} \right) \mathrm{C}_{ij}^{-1} \left( D_j - T_j - \sum_k^K r_k \beta_{jk} \right)$$

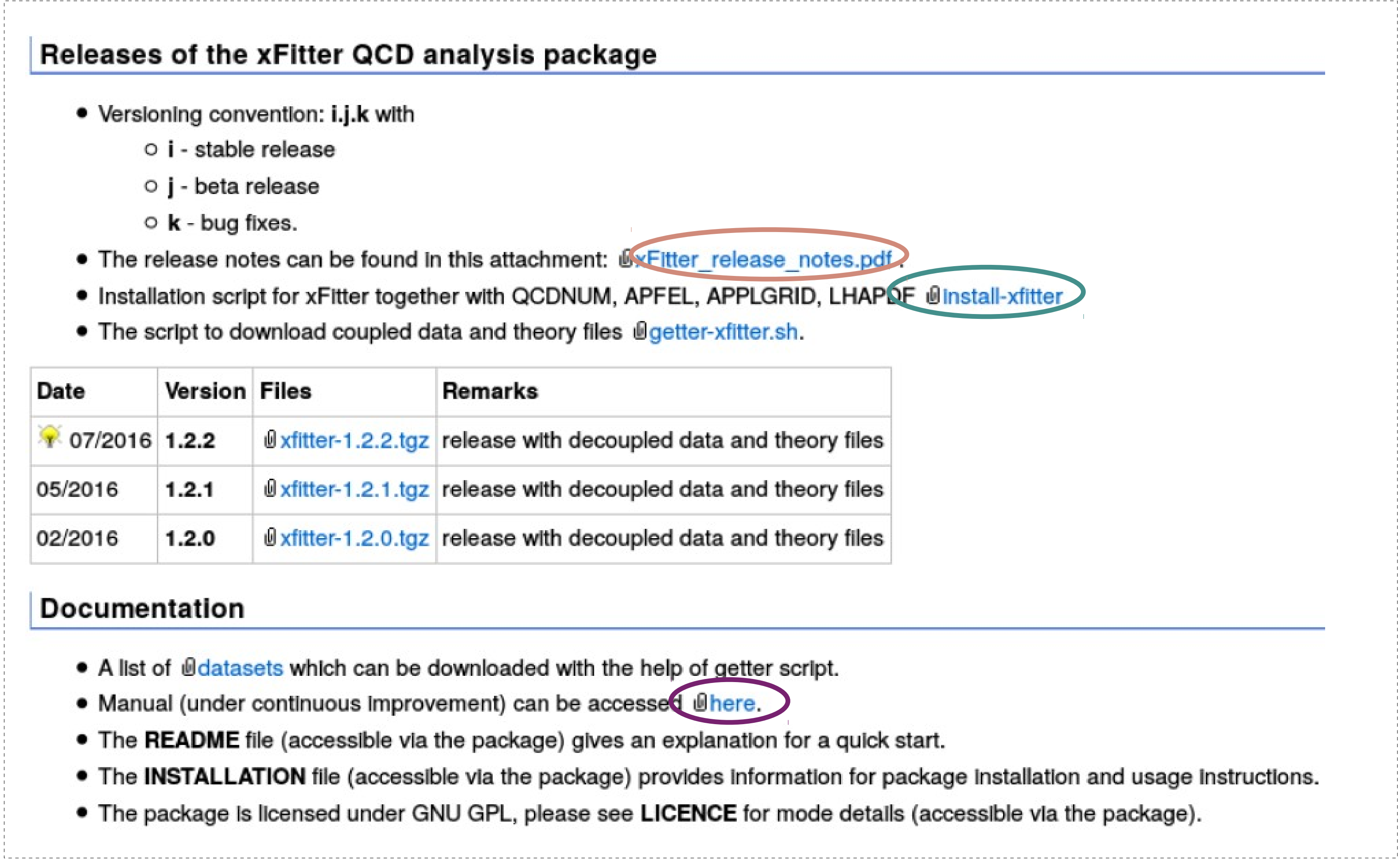
→ *it is important to account statistical and systematic uncertainties of data*

→ *theory (PDF) uncertainties can be accounted for*

# xFitter release xfitter-1.2.2

xFitter /
DownloadPage

https://www.xfitter.org/xFitter/xFitter/DownloadPage



## Releases of the xFitter QCD analysis package

- Versioning convention: **i.j.k** with
  - **i** - stable release
  - **j** - beta release
  - **k** - bug fixes.
- The release notes can be found in this attachment: xFitter_release_notes.pdf.
- Installation script for xFitter together with QCDNUM, APFEL, APPLGRID, LHAPDF install-xfitter
- The script to download coupled data and theory files getter-xfitter.sh.

| Date | Version | Files | Remarks |
|---|---|---|---|
| 07/2016 | **1.2.2** | xfitter-1.2.2.tgz | release with decoupled data and theory files |
| 05/2016 | **1.2.1** | xfitter-1.2.1.tgz | release with decoupled data and theory files |
| 02/2016 | **1.2.0** | xfitter-1.2.0.tgz | release with decoupled data and theory files |

## Documentation

- A list of datasets which can be downloaded with the help of getter script.
- Manual (under continuous improvement) can be accessed here.
- The **README** file (accessible via the package) gives an explanation for a quick start.
- The **INSTALLATION** file (accessible via the package) provides information for package installation and usage instructions.
- The package is licensed under GNU GPL, please see **LICENCE** for mode details (accessible via the package).

By default only final combined HERA I+II data are distributed

→ getter-xfitter.sh script to download data with corresponding theory files

→ in directory 'datasets' located all available files

# xfitter-1.2.2 vs xfitter-1.2.1

Several fixes were applied:

| Release | Date | Description |
|---|---|---|
| `xfitter-1.2.2` | 8.07.2016 | • Fix in profiling due to multiple sign flips, affects also reweighting.<br>• Fix in the output of PDFs, strange is symetrised to $(s+\bar{s})/2$.<br>• Fix in calculation of theory error bands for parametrisation uncertainties for the `--therr` option.<br>• Fix for `has_photon` LHAPDF variable and protection against LHAPDFQ0 with photon PDFs.<br>• Fixes to dipole steering file in input steering file, updated now to current settings.<br>• Added the H1 beauty data to the list of data files<br>• Fix in the default theoretical parameters for HVQMNR to be used not in Fit mode.<br>• Fix on warning message from Fastnlo.<br>• Added examples in the example directory together with the tutorial slides from CTEQ 2016 school.<br>• Fix in configuratuon for `--disable-root` option.<br>• Fix in $\alpha_s$ interpolation and protection in overriding the output directories.<br>• Fix in photon PDF sum rules. |

https://indico.desy.de/contributionDisplay.py?contribId=11&confId=13506

# xfitter-1.2.2 examples (CTEQ school)

https://indico.desy.de/contributionDisplay.py?contribId=11&confId=13506

A list of educational examples are provided in the package - prepared for the CTEQ summer school 2016:

**Exercise 1:** PDF fit
→ learn the basic settings of a QCD analysis, based on HERA data only

**Exercise 2:** Simultaneous PDF fit and $\alpha_s$
→ learn the basic of an $\alpha_s$extraction using H1 jet data

**Exercise 3:** LHAPDF analysis
→ how to estimate impact of a new data without fitting:
→ profiling and reweighting techniques

**Exercise 4:** Plotting LHAPDF files
→ direct visualisation of PDFs from LHAPDF6 using simple python scripts

**Exercise 5:** Equivalence of $\chi^2$ representations
→ understand different $\chi^2$ representations

# New Physics Cases in xFitter

**QED PDFs up to NNLO QCD + LO QED** in FFNS and VFNS are now available via evolutions in:

→ QCDNUM adjusted for DGLAP+QED [R. Sadykov]

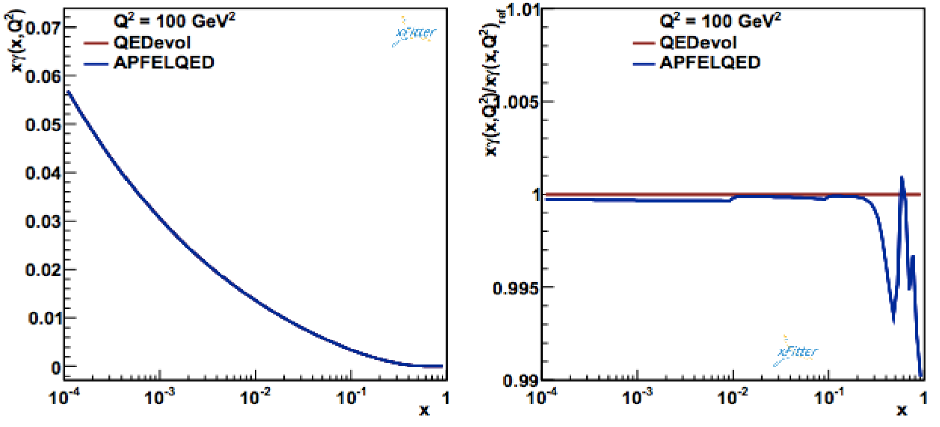
http://www.nikhef.nl/~h24/qcdnum

→ APFEL DGLAP+QED as used by NNPDF2.3 [V. Bertone et al]

https://apfel.hepforge.org

→ plan to add NLO QED, interface APPLGRID to SANC

https://apfel.hepforge.org/mela.html



V. Bertone, R. Sadykov

**New: NLO QCD + QED via APFEL** in xFitter:

ONLY ON GIT
very new!

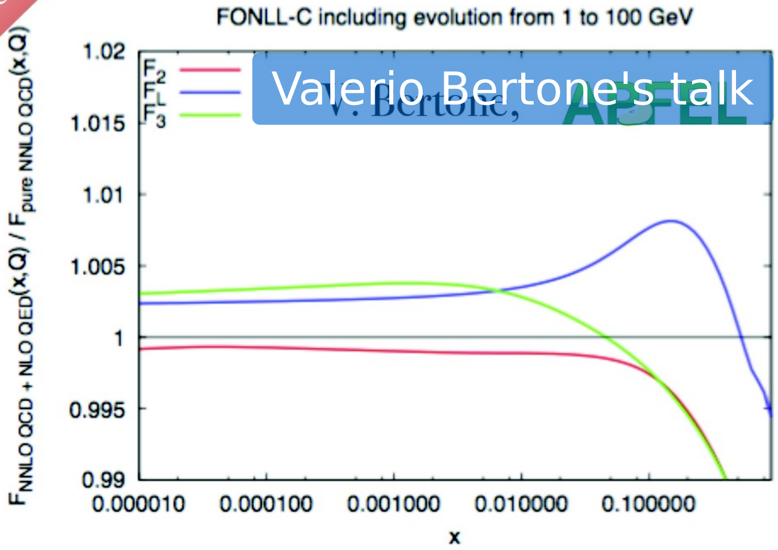
→ at LO QED, no corrections to the SFs are needed

→ at NLO QED, access to new diagrams: $\gamma^*\gamma \to qq$ and $\gamma^*q \to q\gamma$

→ implementing the O($\alpha_s\alpha$) and the O($\alpha^2$) corrections to the DGLAP splitting functions on top of the O($\alpha$) ones

→ implementing O($\alpha_s^2\alpha$) and the O($\alpha^2$), O($\alpha^2\alpha_s$) corrections to $\beta$ functions

→ when including NLO QED corrections, not only the evolution is affected but also the DIS structure functions get corrected



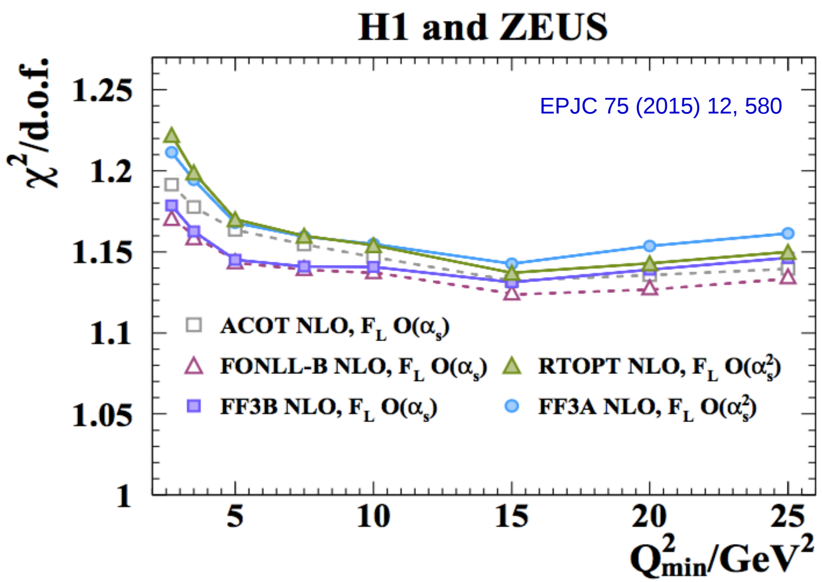
# New Physics Cases in xFitter

Addition of new Heavy Flavour Scheme: **FONLL VFNS**

→ it is available thanks to collaboration with APFEL

→ various FONLL options available via interface to APFEL https://apfel.hepforge.org

→ ABM scheme was up-to-dated to OPENQCDRAD v 2.0b4 http://www-zeuthen.desy.de/~alekhin/OPENQCDRAD

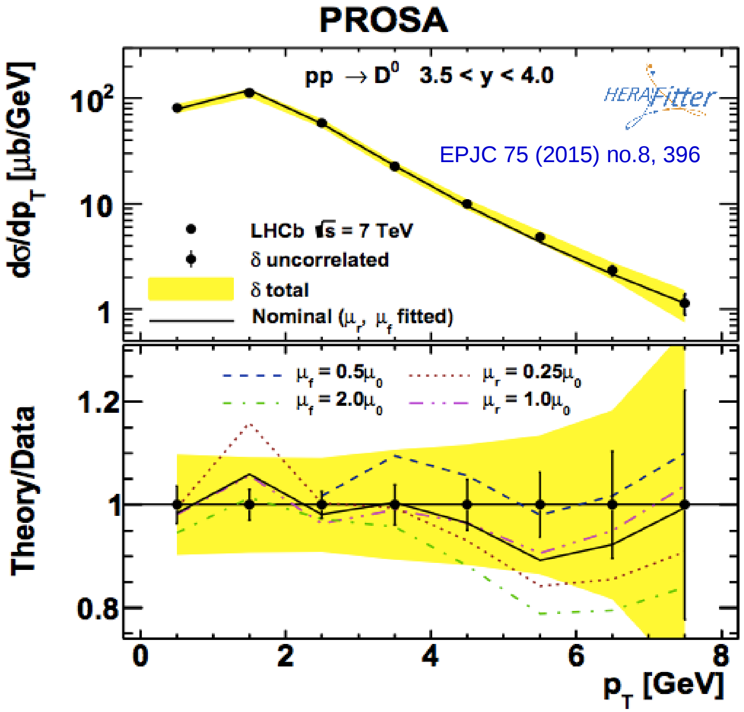


**Interface to Mangano-Nason-Ridolfi** (MNR, NPB 373 (1992) 295) theory code added in xFitter

→ was used for analysing the heavy-flavour production at LHCb and at HERA (via OPENQCDRAD)

  → use of FFNS for accounting of heavy quark masses at NLO

  → added corresponding LHCb data



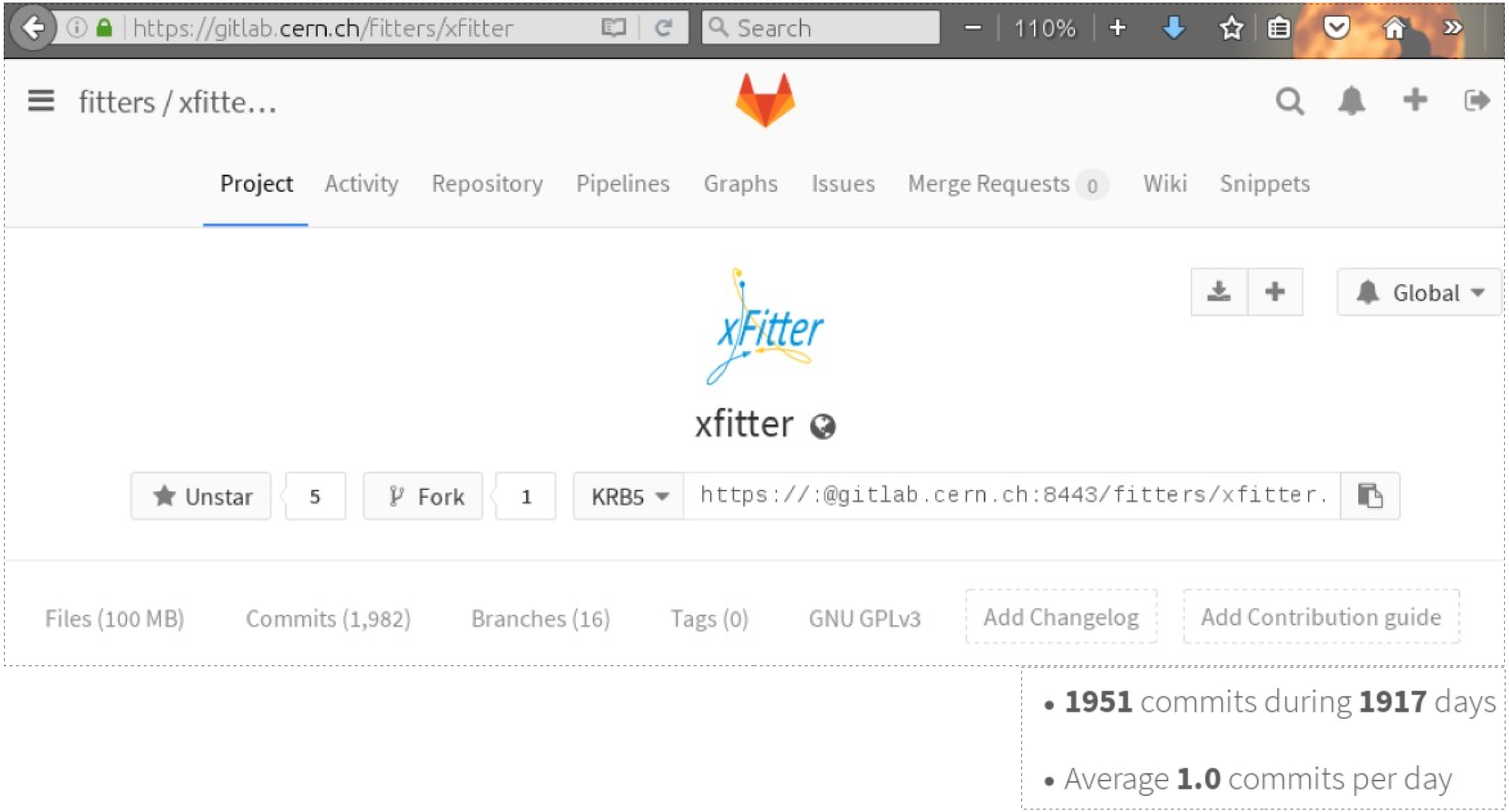
**Added extra reweighing option** using Giele-Keller weights

# xFitter on GitLab

**GitLab** (CERN) is now the main repository of the project

→ open access to download for everyone (read only)

https://gitlab.cern.ch/fitters/xfitter

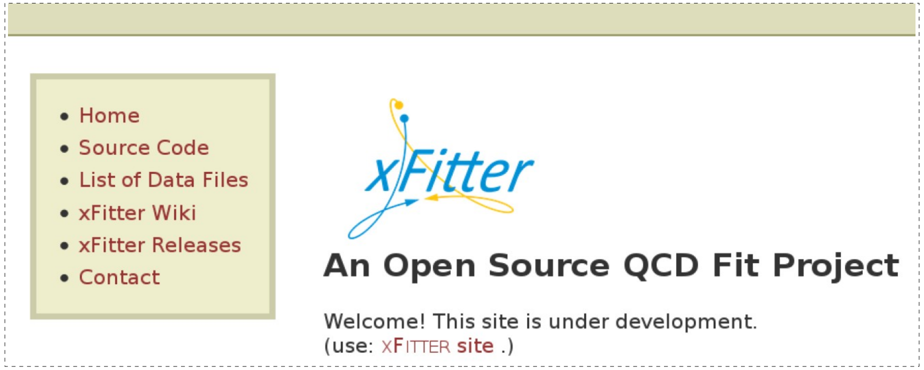


- **1951** commits during **1917** days
- Average **1.0** commits per day

Commits from developers which have no CERN account handled via mirror-GIT public page

# xFitter on Hepforge: data access

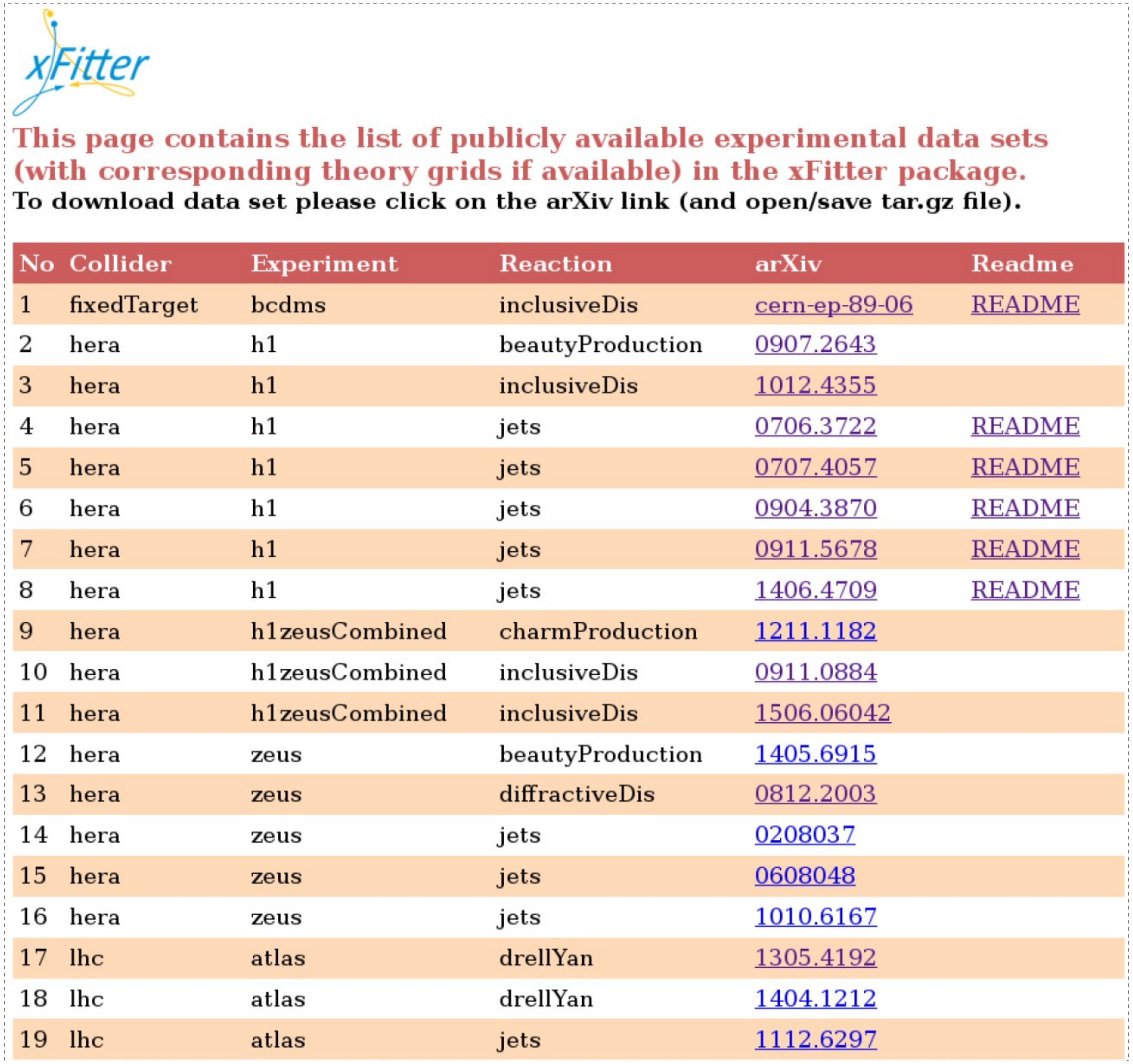
http://xfitter.hepforge.org/



- Home
- Source Code
- List of Data Files
- xFitter Wiki
- xFitter Releases
- Contact

**An Open Source QCD Fit Project**

Welcome! This site is under development.
(use: xFITTER site .)



**This page contains the list of publicly available experimental data sets (with corresponding theory grids if available) in the xFitter package.**

**To download data set please click on the arXiv link (and open/save tar.gz file).**

| No | Collider | Experiment | Reaction | arXiv | Readme |
|---|---|---|---|---|---|
| 1 | fixedTarget | bcdms | inclusiveDis | cern-ep-89-06 | README |
| 2 | hera | h1 | beautyProduction | 0907.2643 | |
| 3 | hera | h1 | inclusiveDis | 1012.4355 | |
| 4 | hera | h1 | jets | 0706.3722 | README |
| 5 | hera | h1 | jets | 0707.4057 | README |
| 6 | hera | h1 | jets | 0904.3870 | README |
| 7 | hera | h1 | jets | 0911.5678 | README |
| 8 | hera | h1 | jets | 1406.4709 | README |
| 9 | hera | h1zeusCombined | charmProduction | 1211.1182 | |
| 10 | hera | h1zeusCombined | inclusiveDis | 0911.0884 | |
| 11 | hera | h1zeusCombined | inclusiveDis | 1506.06042 | |
| 12 | hera | zeus | beautyProduction | 1405.6915 | |
| 13 | hera | zeus | diffractiveDis | 0812.2003 | |
| 14 | hera | zeus | jets | 0208037 | |
| 15 | hera | zeus | jets | 0608048 | |
| 16 | hera | zeus | jets | 1010.6167 | |
| 17 | lhc | atlas | drellYan | 1305.4192 | |
| 18 | lhc | atlas | drellYan | 1404.1212 | |
| 19 | lhc | atlas | jets | 1112.6297 | |

Complementary information about the project (to xfitter.org)

→ possibility to download **data** files (including theory)

→ updated automatically with new data added to svn

will include script to download all data at once

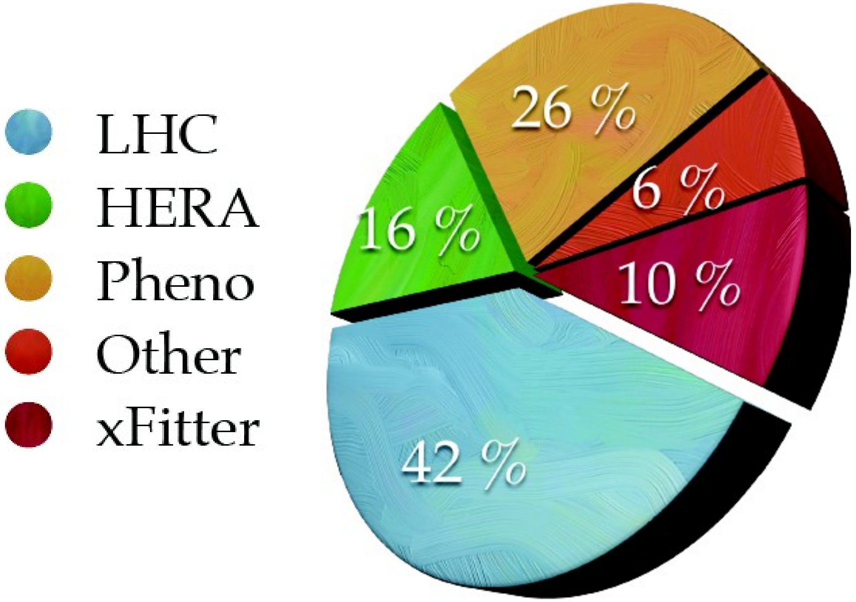
**Your feedback is welcome**

(via email xfitter-help@desy.de)

# Results Obtained with xFitter

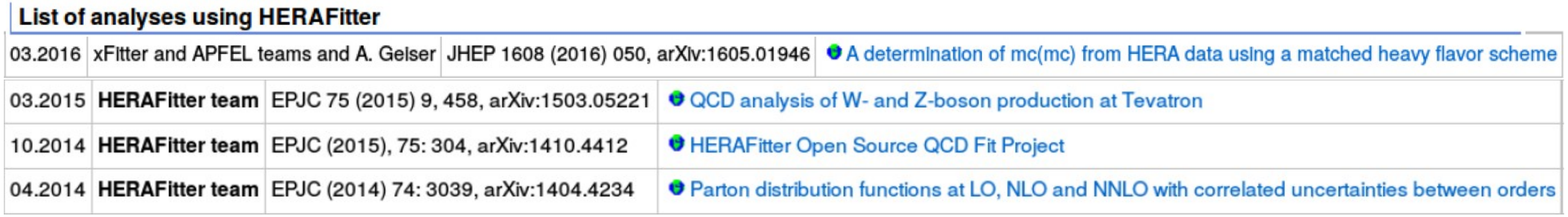
More than **30 public results** obtained using xFitter from the beginning of the project

https://www.xfitter.org/xFitter/xFitter/results



LHC experiments provide the main developments and usage of the xFitter platform

## **xFitter** publications:

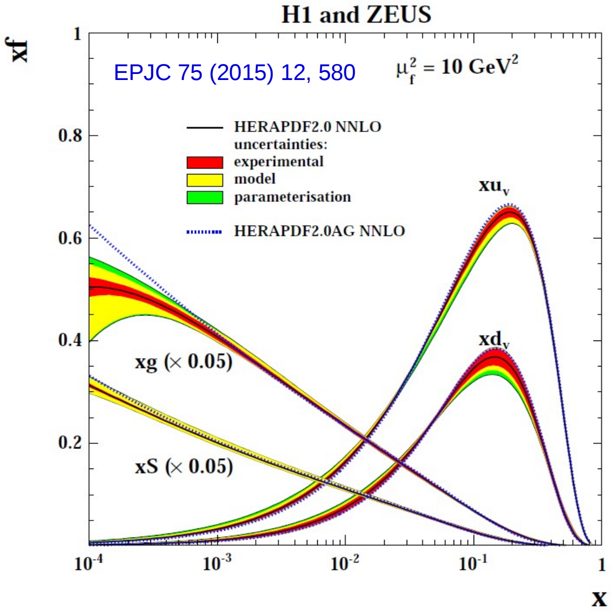


**List of analyses using HERAFitter**

| | | | |
|---|---|---|---|
| 03.2016 | xFitter and APFEL teams and A. Geiser | JHEP 1608 (2016) 050, arXiv:1605.01946 | A determination of mc(mc) from HERA data using a matched heavy flavor scheme |
| 03.2015 | **HERAFitter team** | EPJC 75 (2015) 9, 458, arXiv:1503.05221 | QCD analysis of W- and Z-boson production at Tevatron |
| 10.2014 | **HERAFitter team** | EPJC (2015), 75: 304, arXiv:1410.4412 | HERAFitter Open Source QCD Fit Project |
| 04.2014 | **HERAFitter team** | EPJC (2014) 74: 3039, arXiv:1404.4234 | Parton distribution functions at LO, NLO and NNLO with correlated uncertainties between orders |

# Results Obtained with xFitter: Examples

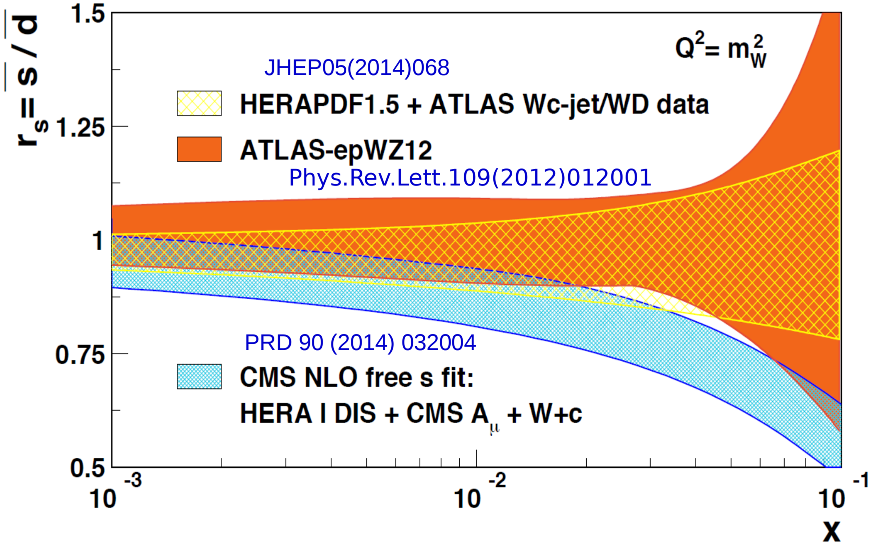
## **DIS inclusive processes in *ep*** (fixed target)



EPJC 75 (2015) 12, 580

## **Drell-Yan processes** (*pp, ppbar*)

→ strange quark density determination

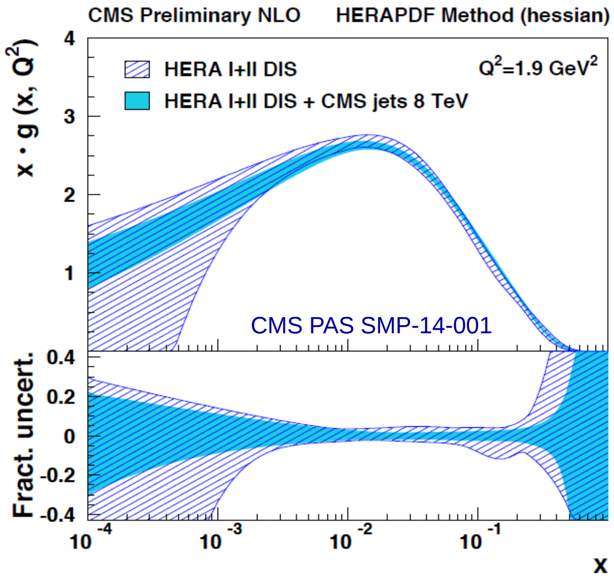


JHEP05(2014)068

Phys.Rev.Lett.109(2012)012001

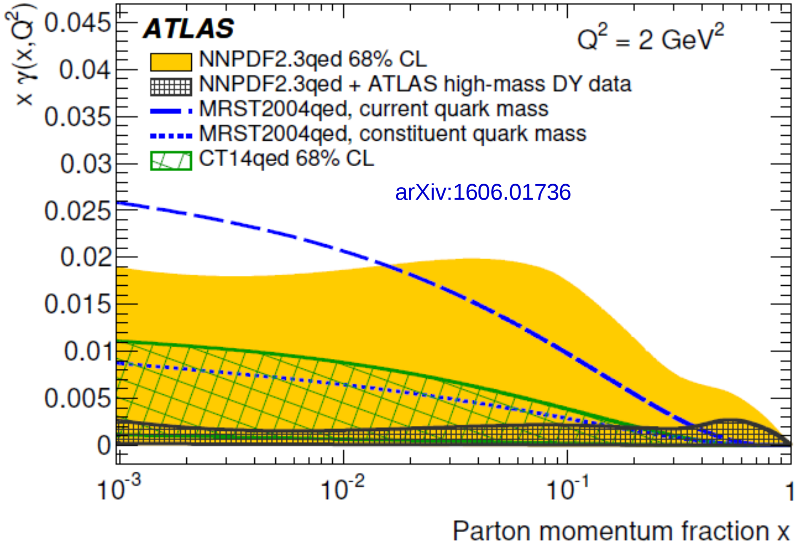
PRD 90 (2014) 032004

## **Jet production** (*ep, pp, ppbar*)



CMS PAS SMP-14-001

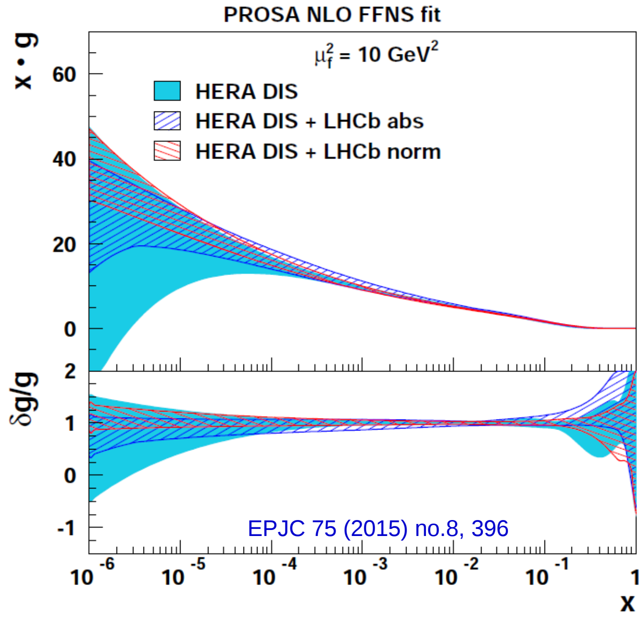
## **DY data sensitivity to photon PDF**



arXiv:1606.01736

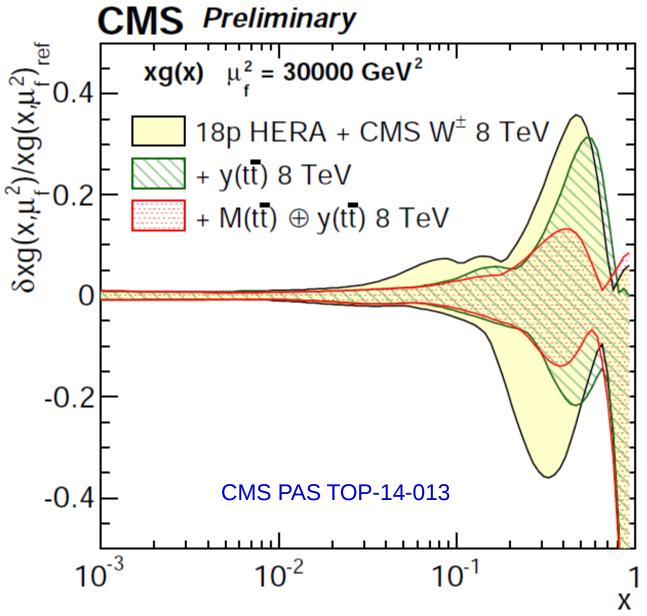
# Results Obtained with xFitter: Examples

## Heavy Quark production *(ep, pp, ppbar)*



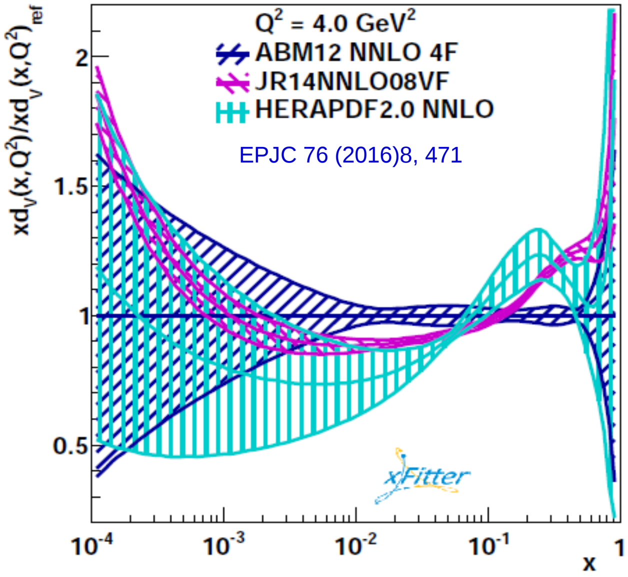
EPJC 75 (2015) no.8, 396

## Top-quark production *(pp, ppbar)*



CMS PAS TOP-14-013

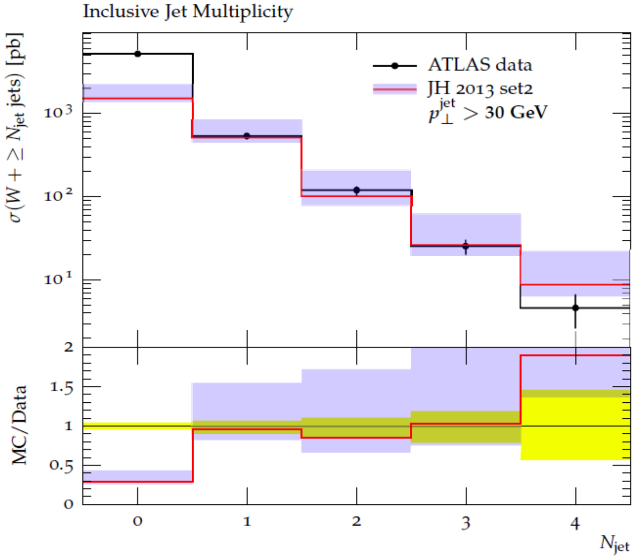
## Evaluation of modern PDFs (benchmarking)



EPJC 76 (2016)8, 471

## TMD PDFs (and applications)

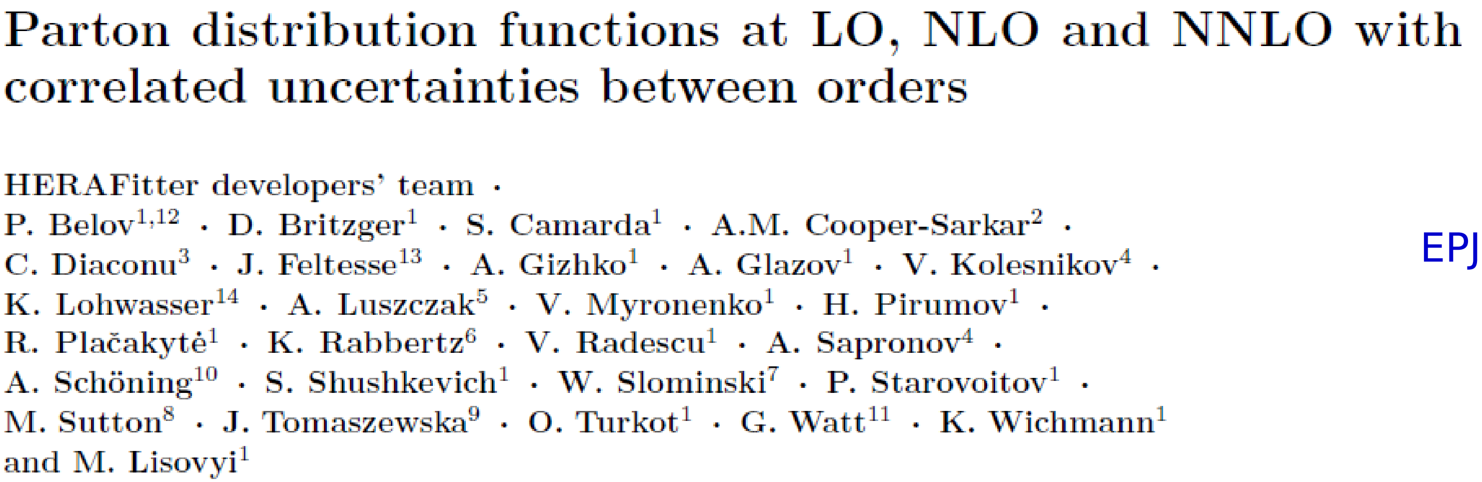
Nucl.Phys. B883 (2014) 1-19



# xFitter Developers Team Publication

The first **xFitter Developers Team** publication

→ addresses the correlations in PDF uncertainties at different orders

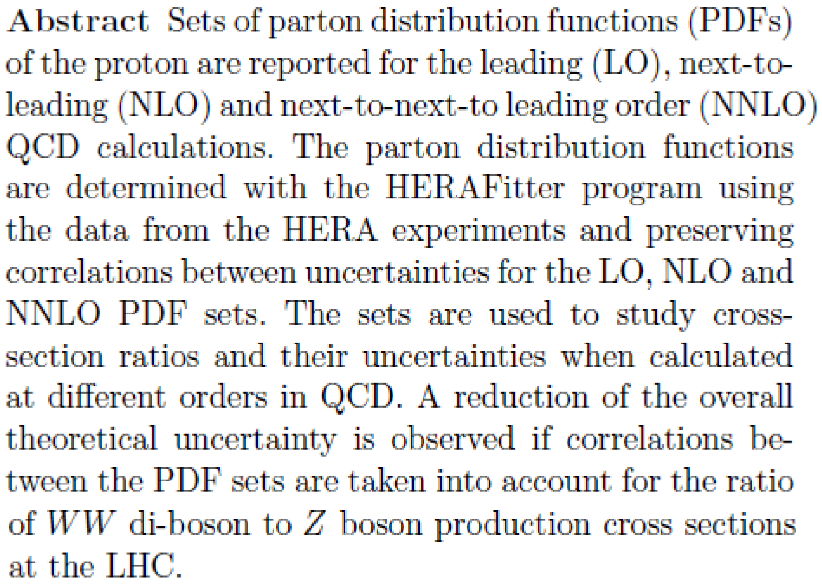


## Parton distribution functions at LO, NLO and NNLO with correlated uncertainties between orders

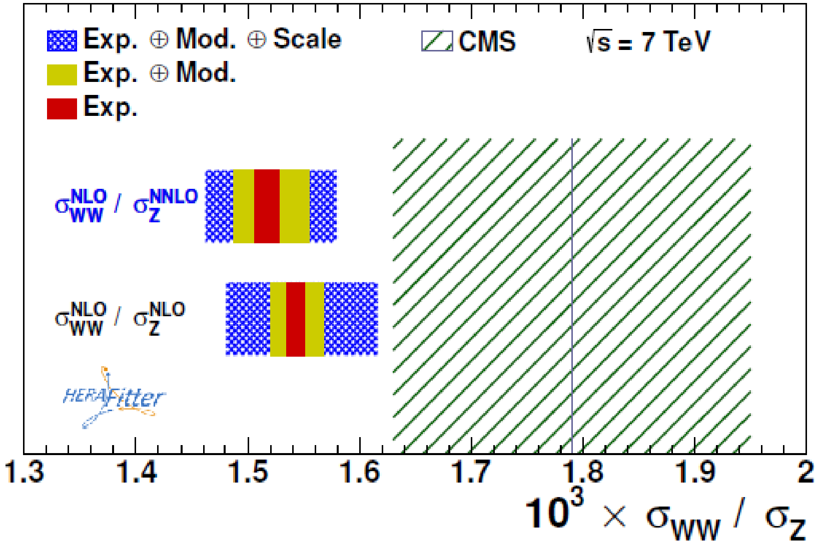
HERAFitter developers' team ·
P. Belov[1,12] · D. Britzger[1] · S. Camarda[1] · A.M. Cooper-Sarkar[2] ·
C. Diaconu[3] · J. Feltesse[13] · A. Gizhko[1] · A. Glazov[1] · V. Kolesnikov[4] ·
K. Lohwasser[14] · A. Luszczak[5] · V. Myronenko[1] · H. Pirumov[1] ·
R. Plačakytė[1] · K. Rabbertz[6] · V. Radescu[1] · A. Sapronov[4] ·
A. Schöning[10] · S. Shushkevich[1] · W. Slominski[7] · P. Starovoitov[1] ·
M. Sutton[8] · J. Tomaszewska[9] · O. Turkot[1] · G. Watt[11] · K. Wichmann[1]
and M. Lisovyi[1]

H. Pirumov,
M. Lisovyi

EPJC (2014) 74: 3039
arXiv:1404.4234



**Abstract** Sets of parton distribution functions (PDFs) of the proton are reported for the leading (LO), next-to-leading (NLO) and next-to-next-to leading order (NNLO) QCD calculations. The parton distribution functions are determined with the HERAFitter program using the data from the HERA experiments and preserving correlations between uncertainties for the LO, NLO and NNLO PDF sets. The sets are used to study cross-section ratios and their uncertainties when calculated at different orders in QCD. A reduction of the overall theoretical uncertainty is observed if correlations between the PDF sets are taken into account for the ratio of *WW* di-boson to *Z* boson production cross sections at the LHC.



# xFitter Developers Team Publication

## QCD analysis of $W$- and $Z$-boson production at Tevatron

- W and Z boson production at Tevatron is valence quark dominated

→ W and Z measurements at Tevatron can be used to improve valence quark PDFs, especially the d-quark type (less constrained by DIS data)

S. Camarda

EPJC 75 (2015) 9, 458

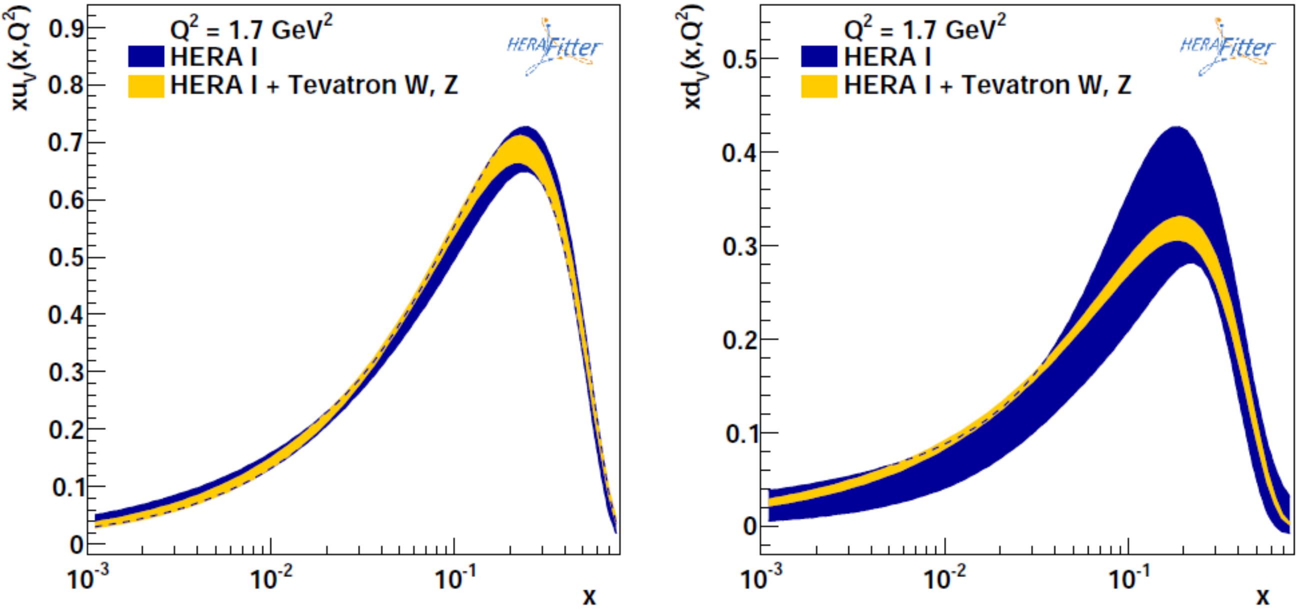
Impact of Tevatron data on PDFs was studied in two ways:

1) preforming a full NLO PDF fit

2) studied by minimizing data to theory $\chi^2$ vs nuisance parameters corresponding to PDF eigenvectors (“profiling”)

→ after the inclusion of Tevatron data, the uncertainties of $u_v$ and $d_v$ become comparable in size

→ data and theory (APPLgrids) files are available in xfitter.hepforge.org



# xFitter Developers Team Publication

## A determination of $m_c(m_c)$ from HERA data using a matched heavy-flavor scheme

V. Bertone

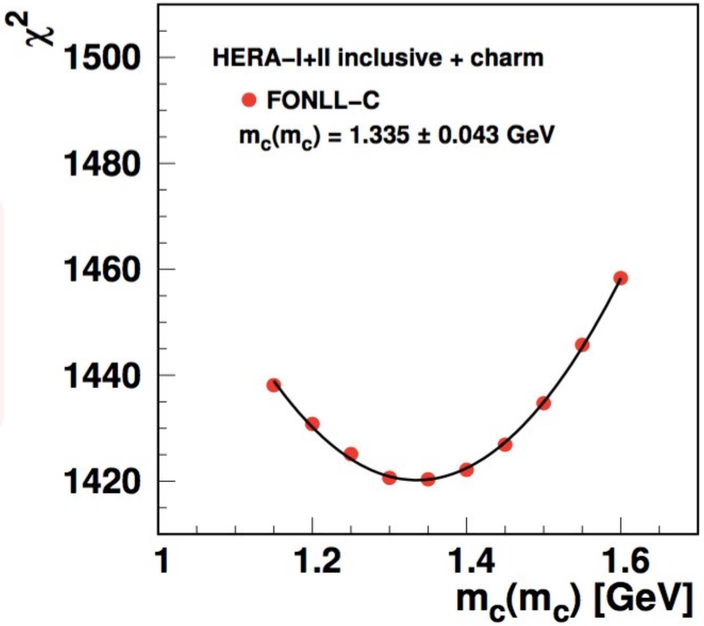
The extraction of $m_c(m_c)$ was performed using FONLL scheme in terms of the MS masses

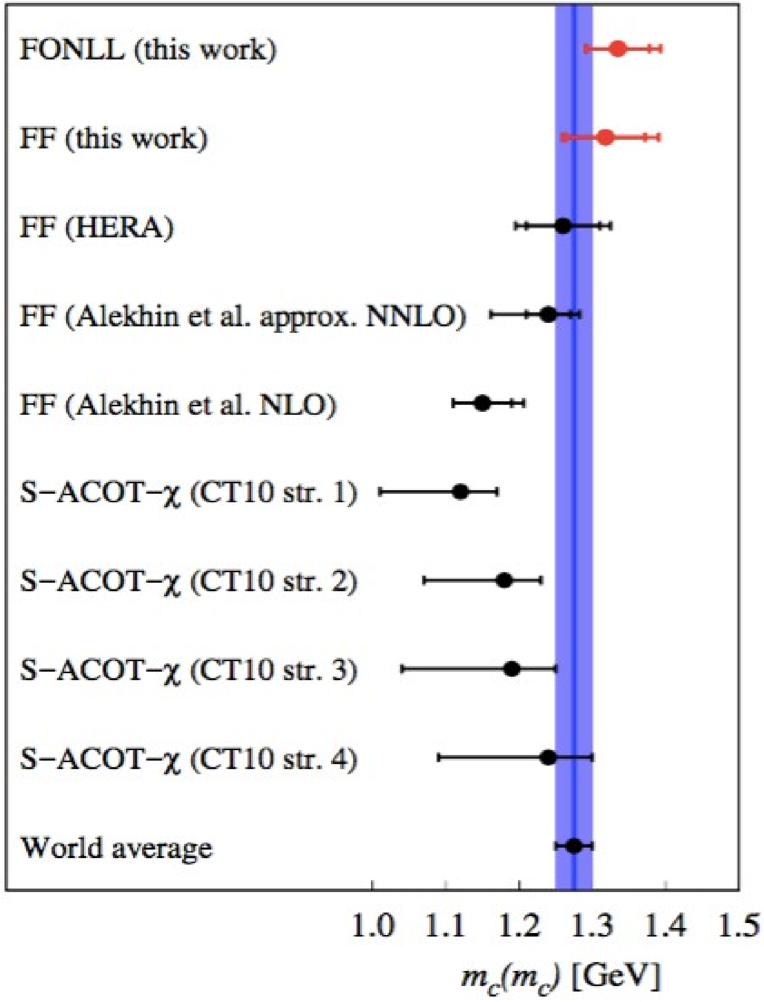
→ combined HERA I + II charm production and DIS cross sections

→ also tested in FFNS (fixed flavour number scheme) at NLO

→ $m_c(m_c)$ value is determined from a parabolic minimum of the global $\chi^2$ scan

→ results is comparable to previous determinations from DIS as well as PDG world average





FONLL-C:

$$m_c(m_c) = 1.335 \pm 0.043(\text{exp})^{+0.019}_{-0.000}(\text{param})^{+0.011}_{-0.008}(\text{mod})^{+0.033}_{-0.008}(\text{th})\ \text{GeV}$$

# Summary

**xFitter** project - a multi-functional QCD framework well integrated into the high energy community (both, experimental and theory) EPJC (2015) 75: 304

→ many active developments thanks to the close collaboration with experiments and theory groups

→ **xfitter-1.2.2** is latest (recommended) release

→ over 30 public results obtained using xFitter (main applications are from LHC)

→ several published dedicated physics studies (developers team publications), more studies are ongoing

→ foreseen future physic (low-x phenomenology, nuclear PDF, etc...) and technical developments (improved user interface for PDF parametrisation form, data cards, python interface, etc...)

*we welcome new ideas and developers :)*

TMDs in xFitter: Voica's talk

# Back-up Slides

# Technical Changes (since xfitter-1.2.0)

- **Change of name of executables:**
  - FitPDF —> xfitter
  - DrawPdfs —> xfitter-draw
  - DrawResults —> xfitter-draw
  - Postproc —> xfitter-process
- Note that in the previous releases there was a theoryfiles directory
  - —> now theoryfiles are stored with datafiles to be in sync

**Installation:**

- xfitter-1.2.0 is compatible with new QCDNUM version > 17.01.10
  - QCDNUM is available now also with autotools installations
  - QCDNUM provides now access to more than standard 13 PDFs, e.g. photon PDF can be added
- Installation of the xfitter-1.2.0 can also be configured via prefix
- Added the possibility to disable root

- Theory formats in xfitter (usage/parsing) have been unified between FASTNLO and APPLGRID
  - old format for FASTNLO is still operational

- Profiling and Reweighing codes now use same general infrastructure

- Possibility to access directly PDFs as stored in LHAPDF (surpassing QCDNUM)
  - LHAPDFNATIVE option added

# Measure the goodness of the fit: $\chi^2$ function

The goodness of the fit is measured by $\chi^2$ in PDFs

→ various forms are implemented in xFitter:

**Standard $\chi^2$ function**

→ same definition rewritten:

$$\chi^2_{\rm exp}\left(\boldsymbol{m},\boldsymbol{b}\right) = \sum_i \frac{\left[m^i - \sum_\alpha \gamma^i_\alpha \mu^i b_\alpha - \mu^i\right]^2}{\left(\delta_{i,\rm stat}\,\mu^i\right)^2 + \left(\delta_{i,\rm uncor}\,\mu^i\right)^2} + \sum_\alpha b_\alpha^2$$

$$\Gamma^i_\alpha = \gamma^i_\alpha\,\mu_i$$

$$\chi^2(\boldsymbol{b}_{\rm exp}, \boldsymbol{b}_{\rm th}) = \sum_{i=1}^{N_{\rm data}} \frac{\left(\sigma_i^{\rm exp} + \sum_\alpha \Gamma^{\rm exp}_{i\alpha} b_{\alpha,\rm exp} - \sigma_i^{\rm th} - \sum_\beta \Gamma^{\rm th}_{i\beta} b_{\beta,\rm th}\right)^2}{\Delta_i^2} + \sum_\alpha b^2_{\alpha,\rm exp} + \sum_\beta b^2_{\beta,\rm th}$$

*Impact of experimental data on PDFs can be studied by minimizing data to theory $\chi^2$ vs nuisance parameters corresponding to PDF eigenvectors ("profiling")*

# PDF Determination: Parametrisation

PDFs are parametrised (at the starting scale $Q^2_0$) using some flexible form

(starting scale choice is arbitrary, often $Q_0 = m_c$)

generic parametrisation form:

$$xf_j(x) = A_j x^{B_j}(1-x)^{C_j} P_j(x)$$

A: overall normalisation
B: small x behavior
C: x→ 1 shape

with $P_j(x) = (1 + \varepsilon_j\sqrt{x} + D_j x + E_j x^2)$ HERAPDF, MSTW/MMHT (Chebyshev polynomials), ABM, JR

or $e^{a_3 x}(1 + e^{a_4} x + e^{a_5} x^2)$ CTEQ, CT (Bernstein polynomials)

→ *parametrisation has to be flexible enough (many free parameters) to avoid bias, however too many parameters may also lead to certain bias (several minima, problems to converge, ...)*

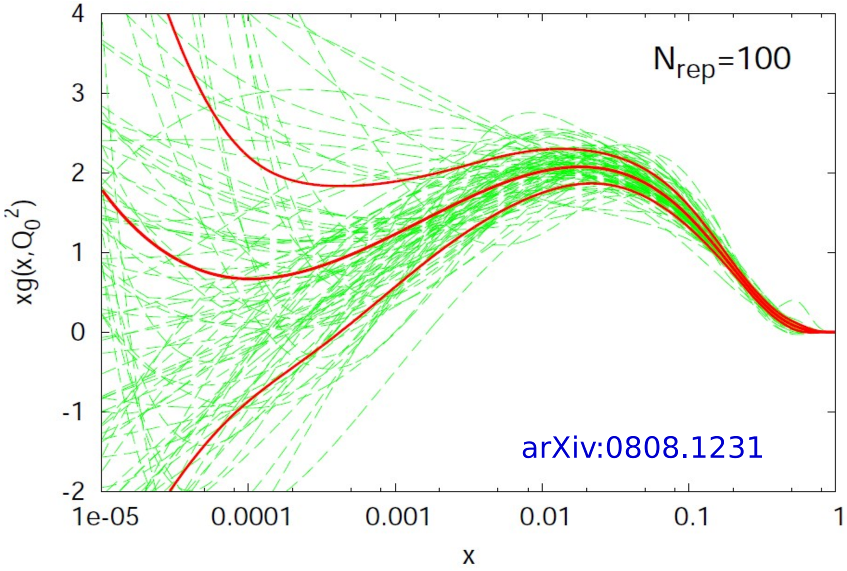
→ *different parametrisations, if carefully chosen, will lead to similar results*

# PDF Determination: Parametrisation

PDFs are parametrised (at the starting scale $Q^2_0$) using some flexible form

parametrised in x-space with the flexible neural network (NN) method:

used by NNPDF collaboration



Basic principle:

→ Monte Carlo (MC) sampling of data (generation of replicas of experimental data)

→ training: set of PDFs parametrised by neural networks on each of the replicas

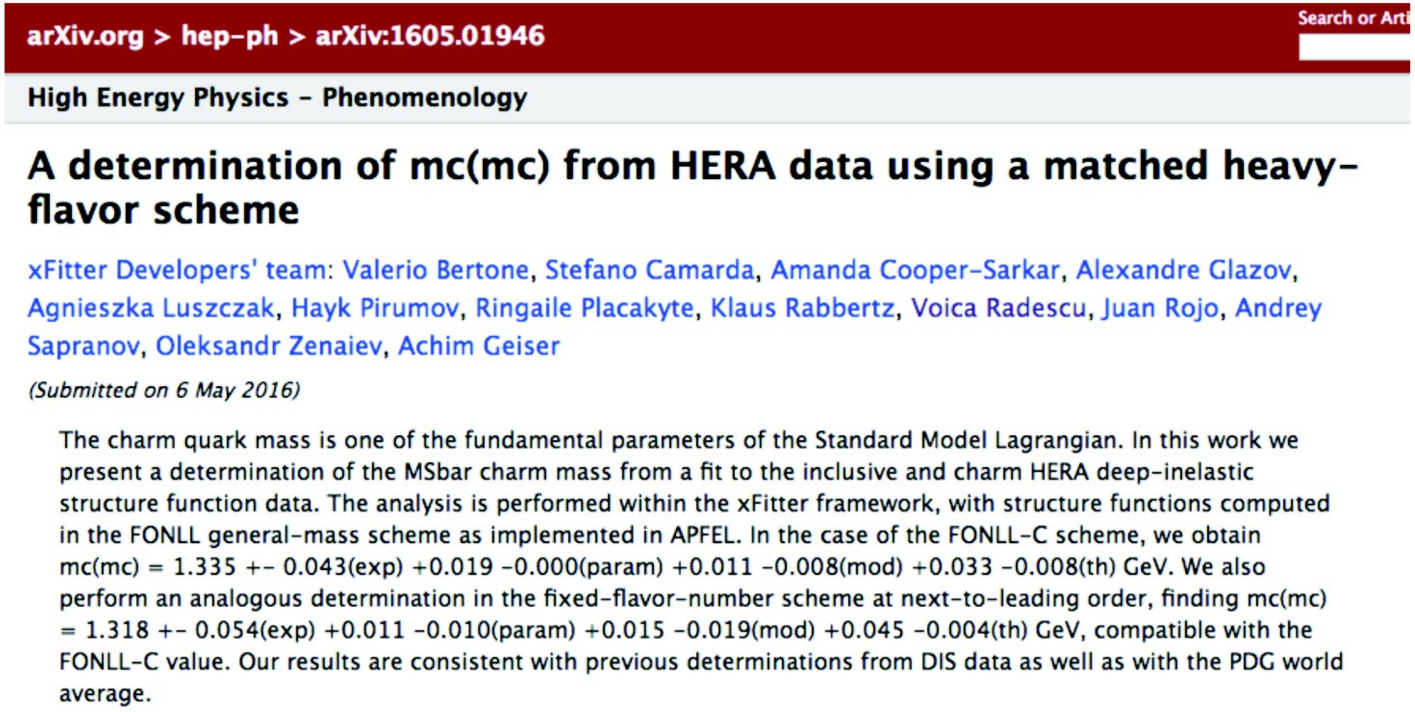
→ validation: fit stops when quality of fit stops improving (determined by random selected validation data)

*advantage*: unbiased parametrisation

*disadvantage*: requires sufficient data

*Uncertainties: Monte Carlo approach (explained later)*

# xFitter Developers Team Publication



arXiv.org > hep-ph > arXiv:1605.01946

Search or Arti

High Energy Physics – Phenomenology

**A determination of mc(mc) from HERA data using a matched heavy-flavor scheme**

xFitter Developers' team: Valerio Bertone, Stefano Camarda, Amanda Cooper-Sarkar, Alexandre Glazov, Agnieszka Luszczak, Hayk Pirumov, Ringaile Placakyte, Klaus Rabbertz, Voica Radescu, Juan Rojo, Andrey Sapranov, Oleksandr Zenaiev, Achim Geiser

*(Submitted on 6 May 2016)*

The charm quark mass is one of the fundamental parameters of the Standard Model Lagrangian. In this work we present a determination of the MSbar charm mass from a fit to the inclusive and charm HERA deep-inelastic structure function data. The analysis is performed within the xFitter framework, with structure functions computed in the FONLL general-mass scheme as implemented in APFEL. In the case of the FONLL-C scheme, we obtain mc(mc) = 1.335 +- 0.043(exp) +0.019 -0.000(param) +0.011 -0.008(mod) +0.033 -0.008(th) GeV. We also perform an analogous determination in the fixed-flavor-number scheme at next-to-leading order, finding mc(mc) = 1.318 +- 0.054(exp) +0.011 -0.010(param) +0.015 -0.019(mod) +0.045 -0.004(th) GeV, compatible with the FONLL-C value. Our results are consistent with previous determinations from DIS data as well as with the PDG world average.

The extraction of $m_c(m_c)$ was performed using FONLL scheme in terms of the $\overline{MS}$ masses → improves perturbative convergence

→ combined HERA I + II charm production and DIS cross sections

→ FONLL-C scheme used – NLO accuracy in the massive sector

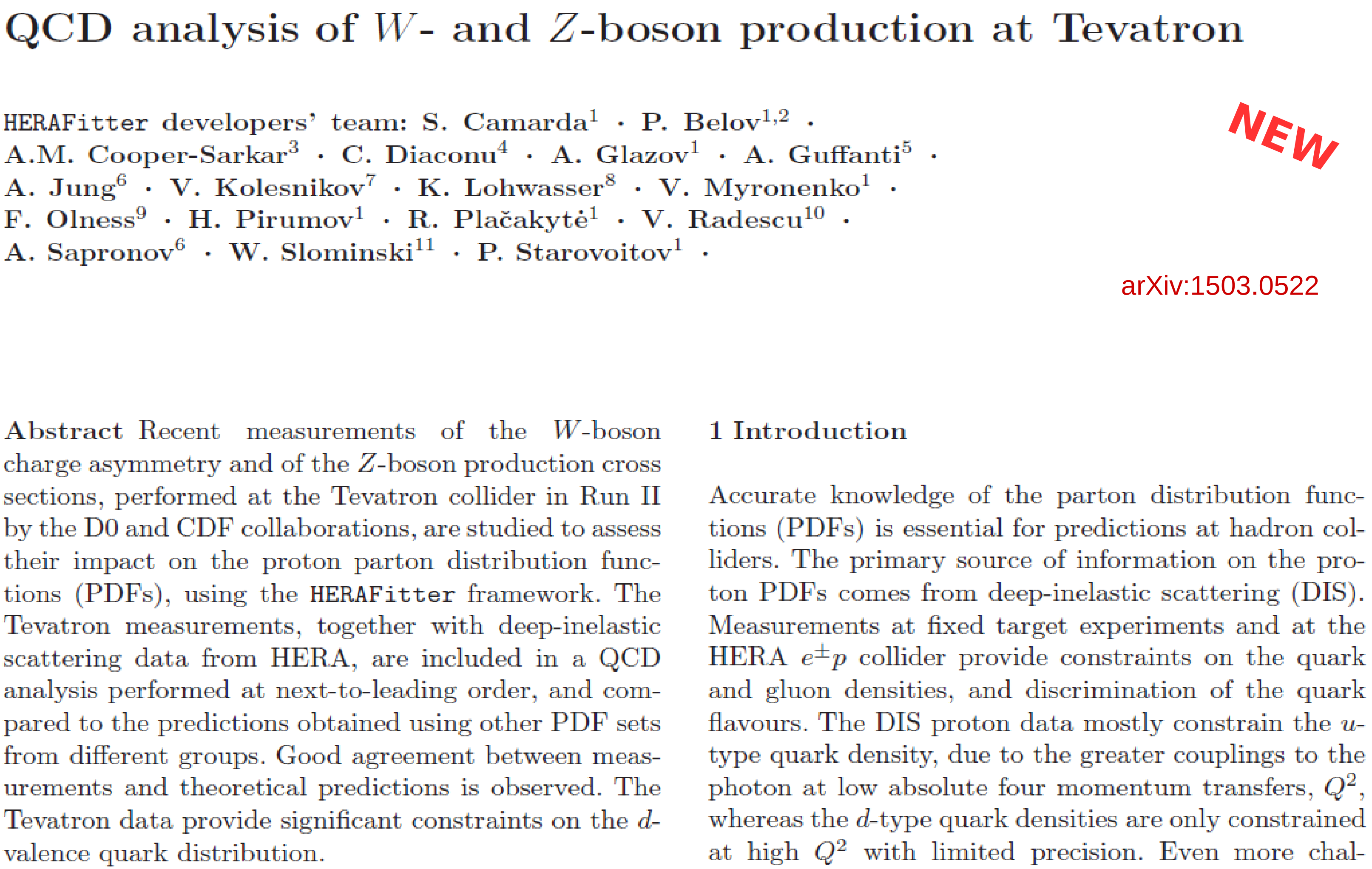
→ also tested in FFNS (fixed flavour number scheme) at NLO

# xFitter Developers Papers

https://www.herafitter.org/HERAFitter/HERAFitter/results

**List of analyses by HERAFitter**

| | | | | |
|---|---|---|---|---|
| NEW 03.2015 | **HERAFitter team** | to be submitted to EPJC, arXiv:1503.05221 | QCD analysis of W- and Z-boson production at Tevatron | Material |
| 10.2014 | **HERAFitter team** | submitted to EPJC, arXiv:1410.4412 | HERAFitter Open Source QCD Fit Project | |
| 04.2014 | **HERAFitter team** | EPJC (2014) 74: 3039, arXiv:1404.4234 | Parton distribution functions at LO, NLO and NNLO with correlated uncertainties between orders | Material |



## QCD analysis of $W$- and $Z$-boson production at Tevatron

HERAFitter developers' team: S. Camarda[1] · P. Belov[1,2] · A.M. Cooper-Sarkar[3] · C. Diaconu[4] · A. Glazov[1] · A. Guffanti[5] · A. Jung[6] · V. Kolesnikov[7] · K. Lohwasser[8] · V. Myronenko[1] · F. Olness[9] · H. Pirumov[1] · R. Plačakytė[1] · V. Radescu[10] · A. Sapronov[6] · W. Slominski[11] · P. Starovoitov[1] ·

**Abstract** Recent measurements of the $W$-boson charge asymmetry and of the $Z$-boson production cross sections, performed at the Tevatron collider in Run II by the D0 and CDF collaborations, are studied to assess their impact on the proton parton distribution functions (PDFs), using the `HERAFitter` framework. The Tevatron measurements, together with deep-inelastic scattering data from HERA, are included in a QCD analysis performed at next-to-leading order, and compared to the predictions obtained using other PDF sets from different groups. Good agreement between measurements and theoretical predictions is observed. The Tevatron data provide significant constraints on the $d$-valence quark distribution.

**1 Introduction**

Accurate knowledge of the parton distribution functions (PDFs) is essential for predictions at hadron colliders. The primary source of information on the proton PDFs comes from deep-inelastic scattering (DIS). Measurements at fixed target experiments and at the HERA $e^{\pm}p$ collider provide constraints on the quark and gluon densities, and discrimination of the quark flavours. The DIS proton data mostly constrain the $u$-type quark density, due to the greater couplings to the photon at low absolute four momentum transfers, $Q^2$, whereas the $d$-type quark densities are only constrained at high $Q^2$ with limited precision. Even more chal-

NEW

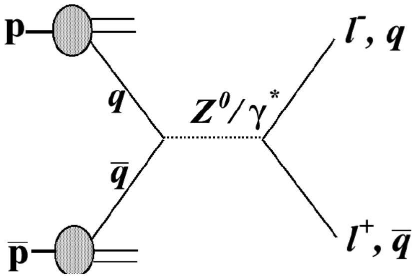
arXiv:1503.0522

# Motivation

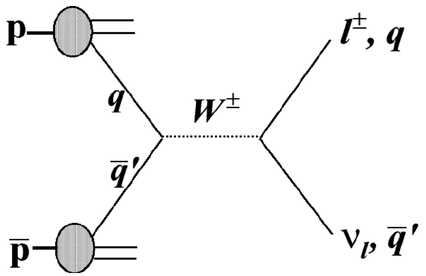
- W and Z boson production at Tevatron is valence quark dominated

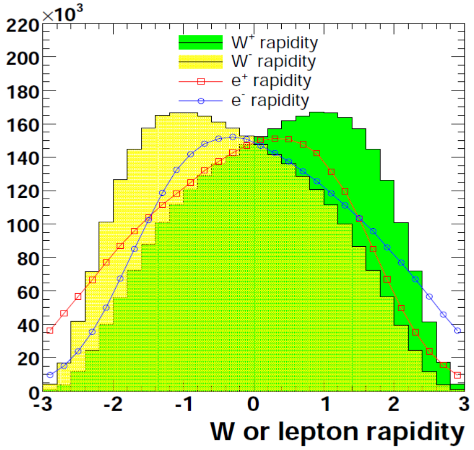
→ W and Z measurements at Tevatron can be used to improve valence quark PDFs, especially the d-quark type (less constrained by DIS data)

- Z rapidity

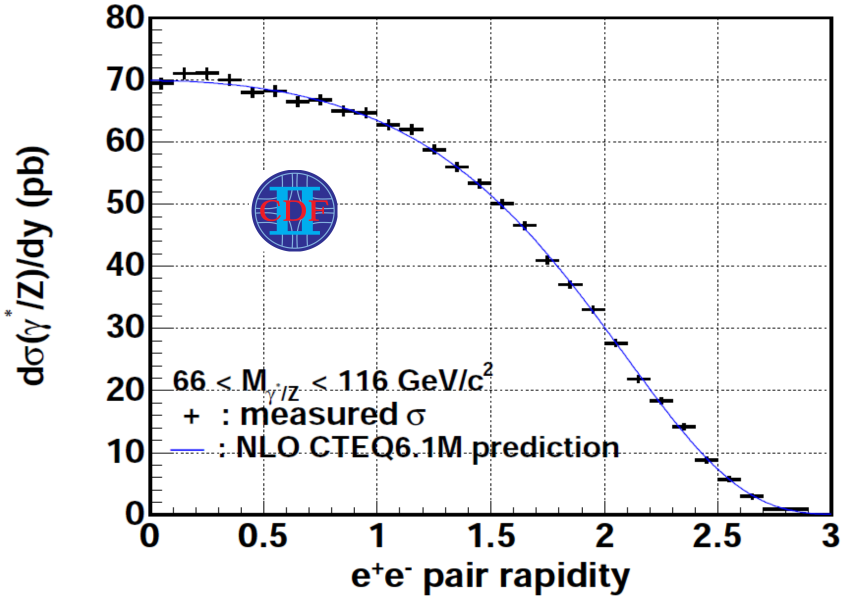


- W and lepton
- charge asymmetry

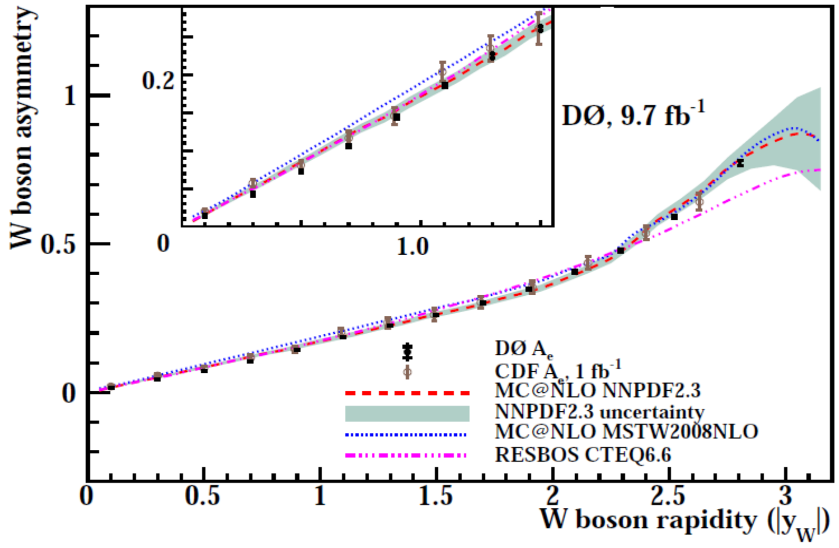




Phys Lett B 692 (2010) 232

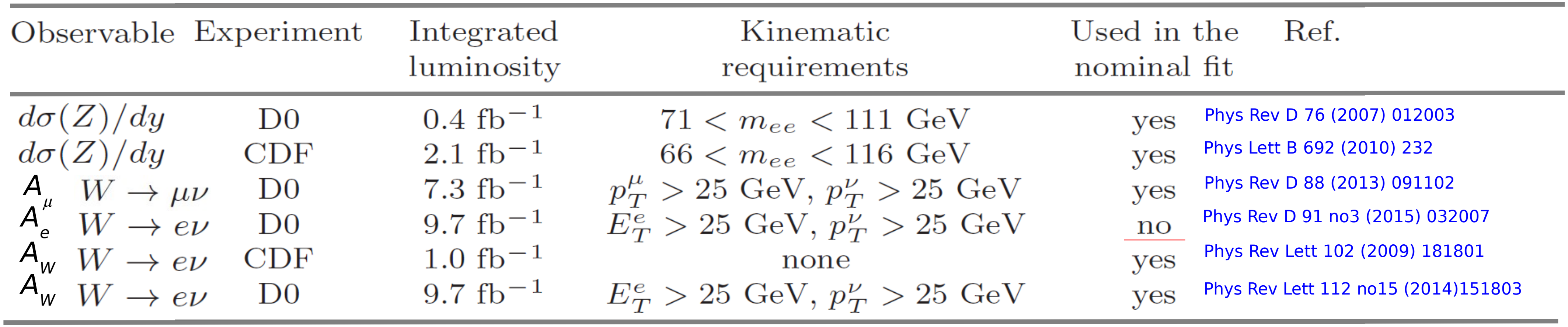


Phys Rev Lett 112 no15 (2014)151803



# W and Z Measurements at Tevatron

Tevatron W and Z production data sets (used in the QCD analysis)



| Observable | Experiment | Integrated luminosity | Kinematic requirements | Used in the nominal fit | Ref. |
|---|---|---|---|---|---|
| $d\sigma(Z)/dy$ | D0 | $0.4\ \mathrm{fb}^{-1}$ | $71 < m_{ee} < 111$ GeV | yes | Phys Rev D 76 (2007) 012003 |
| $d\sigma(Z)/dy$ | CDF | $2.1\ \mathrm{fb}^{-1}$ | $66 < m_{ee} < 116$ GeV | yes | Phys Lett B 692 (2010) 232 |
| $A_\mu$ $W \to \mu\nu$ | D0 | $7.3\ \mathrm{fb}^{-1}$ | $p_T^\mu > 25$ GeV, $p_T^\nu > 25$ GeV | yes | Phys Rev D 88 (2013) 091102 |
| $A_e$ $W \to e\nu$ | D0 | $9.7\ \mathrm{fb}^{-1}$ | $E_T^e > 25$ GeV, $p_T^\nu > 25$ GeV | no | Phys Rev D 91 no3 (2015) 032007 |
| $A_W$ $W \to e\nu$ | CDF | $1.0\ \mathrm{fb}^{-1}$ | none | yes | Phys Rev Lett 102 (2009) 181801 |
| $A_W$ $W \to e\nu$ | D0 | $9.7\ \mathrm{fb}^{-1}$ | $E_T^e > 25$ GeV, $p_T^\nu > 25$ GeV | yes | Phys Rev Lett 112 no15 (2014)151803 |

Revised correlation model:

→ uncertainties of data-driven corrections are treated as bin-to-bin uncorrelated (lepton ID, trigger and charge efficiencies)

(Fast) theoretical predictions: MCFM+APPLGRID

# QCD Analysis Settings

- In the QCD analysis HERA I data sets (JHEP 1001 (2010) 109) used in a simultaneous fit with Tevatron W and Z production data

- Parametrisation functional form:

→ optimisation via paramterisation scan

→ 15 parameter central parametrisation which includes linear and exponential terms

$$
\begin{aligned}
xg(x) &= A_g x^{B_g} (1-x)^{C_g} (1 + D_g x)\,; \\
xu_v(x) &= A_{u_v} x^{B_{u_v}} (1-x)^{C_{u_v}} e^{F_{u_v} x}\,; \\
xd_v(x) &= A_{d_v} x^{B_{d_v}} (1-x)^{C_{d_v}} e^{F_{d_v} x}\,; \\
x\bar{u}(x) &= A_{\bar{u}} x^{B_{\bar{u}}} (1-x)^{C_{\bar{u}}} (1 + D_{\bar{u}} x)\,; \\
x\bar{d}(x) &= A_{\bar{d}} x^{B_{\bar{d}}} (1-x)^{C_{\bar{d}}} (1 + D_{\bar{d}} x)\,.
\end{aligned}
$$

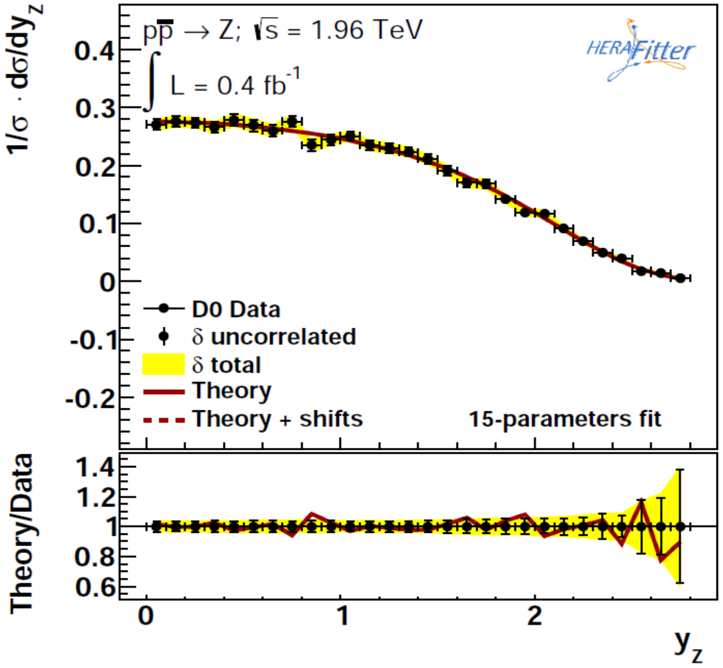
$x\bar{U}=x\bar{u}$ and $x\bar{D}=x\bar{d}+x\bar{s}$ at the starting scale $Q^2$ = 1.7 $\mathrm{GeV}^2$ ($x\bar{s}=r_s x\bar{D}$ with $r_s=1.0$)

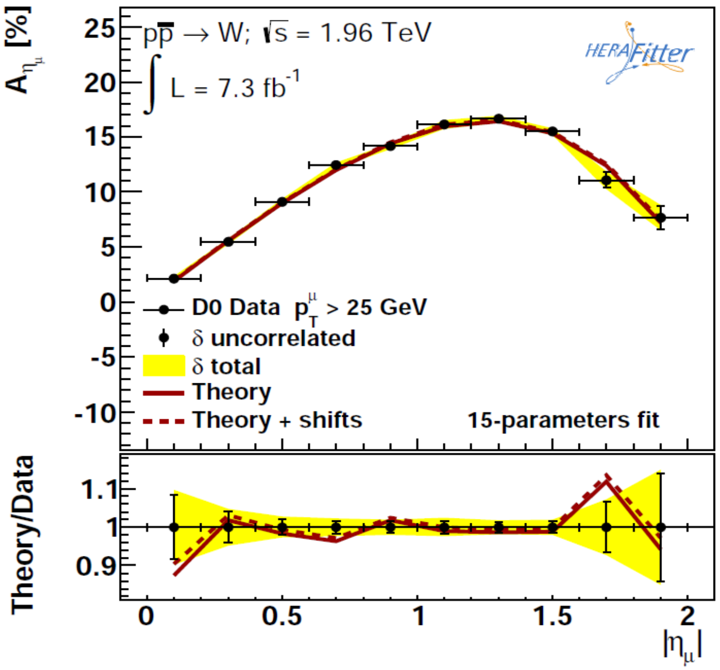
$A_g$, $A_{uv}$, $A_{dv}$ are fixed by the sum rules, $B_{\bar{u}}=B_{\bar{d}}$ and $A_{\bar{u}}=A_{\bar{d}}$

# QCD Analysis Results

Good total and partial (per data set) $\chi^2$ of the fit:

| | HERA I | HERA I + Tevatron W, Z |
|---|---|---|
| Data set | $\chi^2$ / number of points | $\chi^2$ / number of points |
| NC DIS cross sections H1-ZEUS combined $e^-p$. | 112 / 145 | 109 / 145 |
| NC DIS cross sections H1-ZEUS combined $e^+p$. | 326 / 337 | 333 / 337 |
| CC DIS cross sections H1-ZEUS combined $e^-p$. | 20 / 34 | 20 / 34 |
| CC DIS cross sections H1-ZEUS combined $e^+p$. | 27 / 34 | 31 / 34 |
| HERA I correlated $\chi^2$ | 21 | 23 |
| D0 $d\sigma(Z)/dy$ | - | 23 / 28 |
| CDF $d\sigma(Z)/dy$ | - | 32 / 28 |
| D0 muon charge asymmetry in $W \to \mu\nu$ | - | 12 / 10 |
| CDF $W$ charge asymmetry in $W \to e\nu$ | - | 14 / 13 |
| D0 $W$ charge asymmetry in $W \to e\nu$ | - | 8 / 14 |
| Total $\chi^2_{\mathrm{min}}$ / dof | 505 / 535 | 606 / 628 |



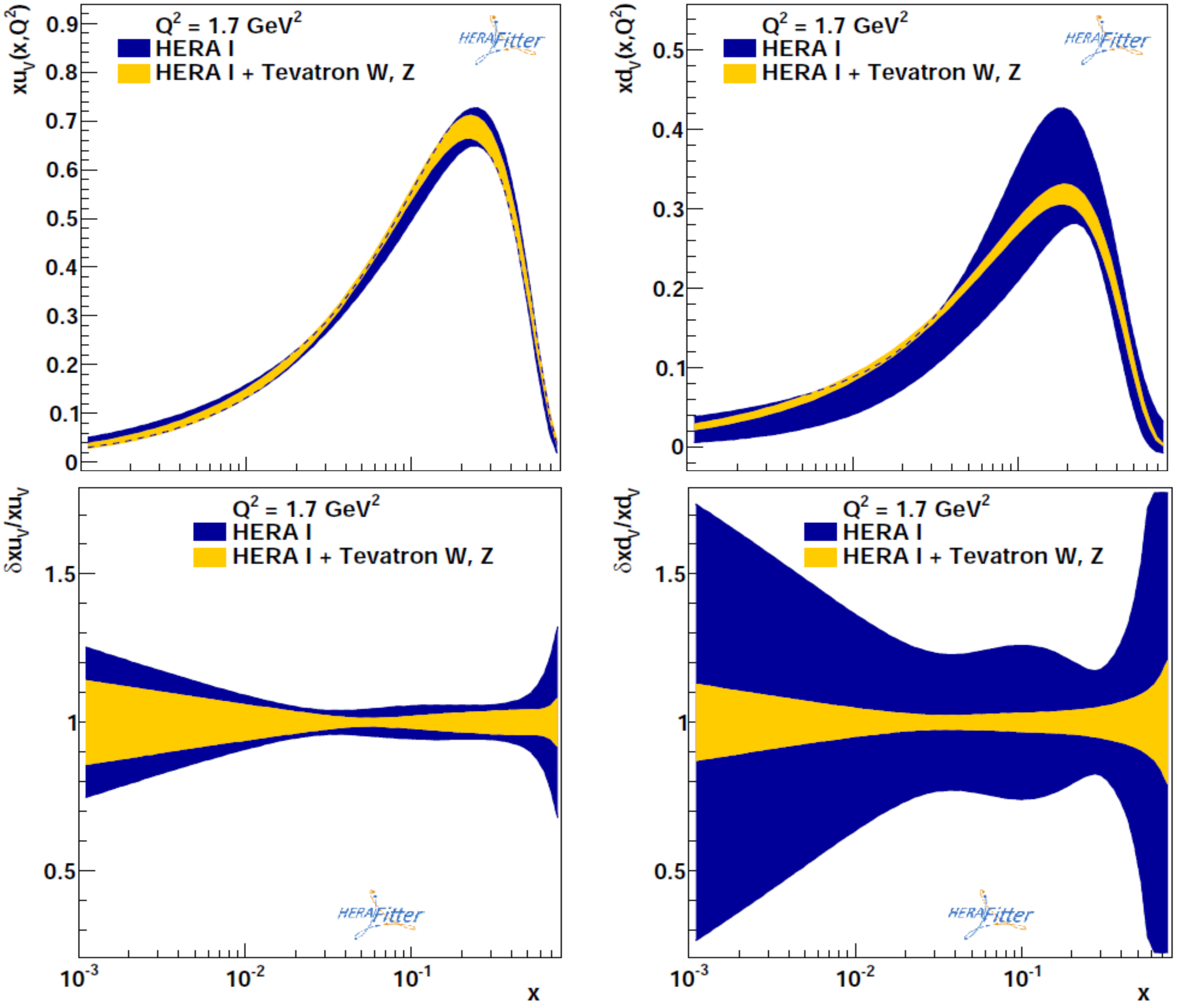


# Impact on Valence Quarks

- Significant impact of the W and Z boson measurements on the valence quarks
- and particularly on the d-type quarks

→ comparison of PDFs from a fit to the HERA data alone to a fit to the HERA and Tevatron data

→ after the inclusion of Tevatron data, the uncertainties of $u_v$ and $d_v$ become comparable in size



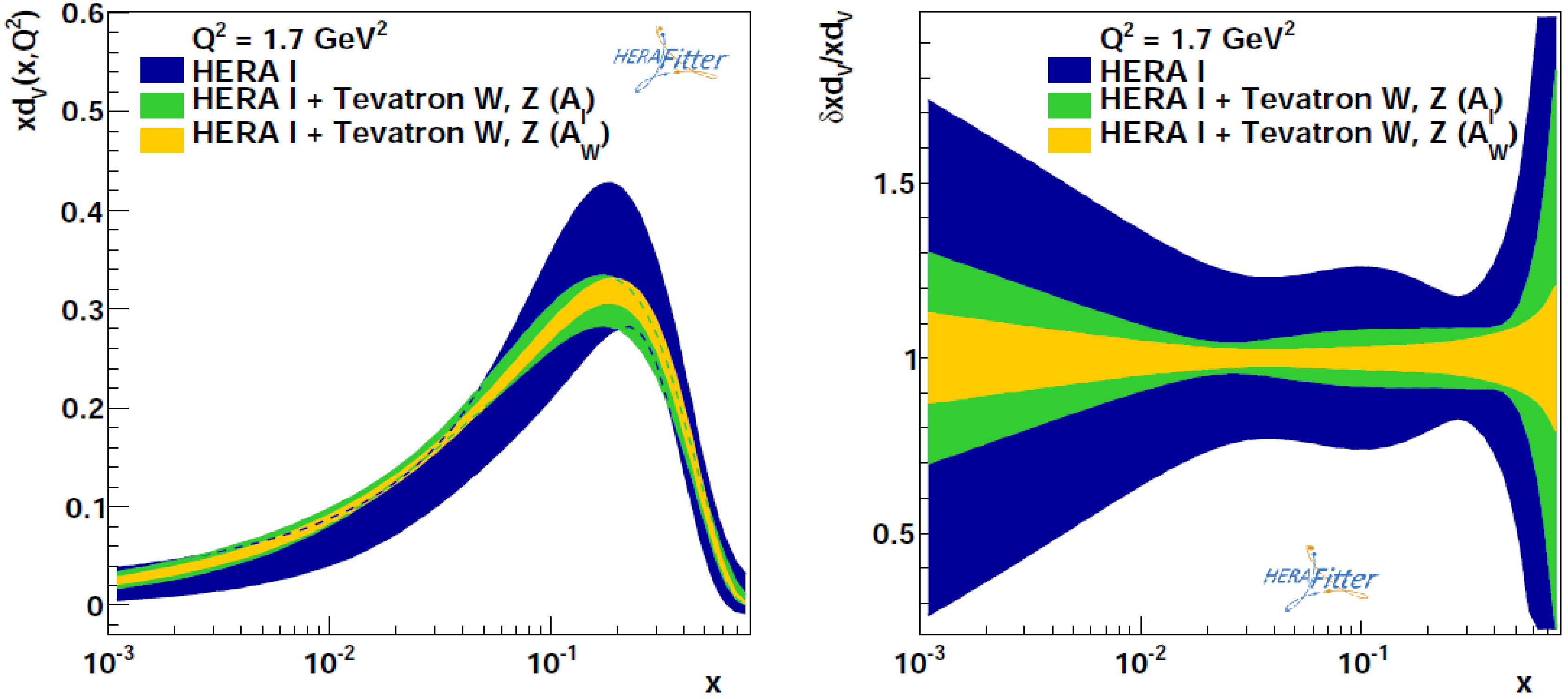
# W vs Lepton Charge Asymmetry

W boson charge asymmetries rely on the reconstruction of the W boson rapidity

→ requires assumptions on the longitudinal momentum of the neutrino

→ model dependence is introduced

→ study of possible bias via alternative fit, excluding W and including lepton asymmetries



→ compatible results but larger uncertainties in a fit with lepton asymmetries

# Comparison of Tevatron Data with Modern PDFs

→ consider only the data sets which are not yet included in the PDF fits

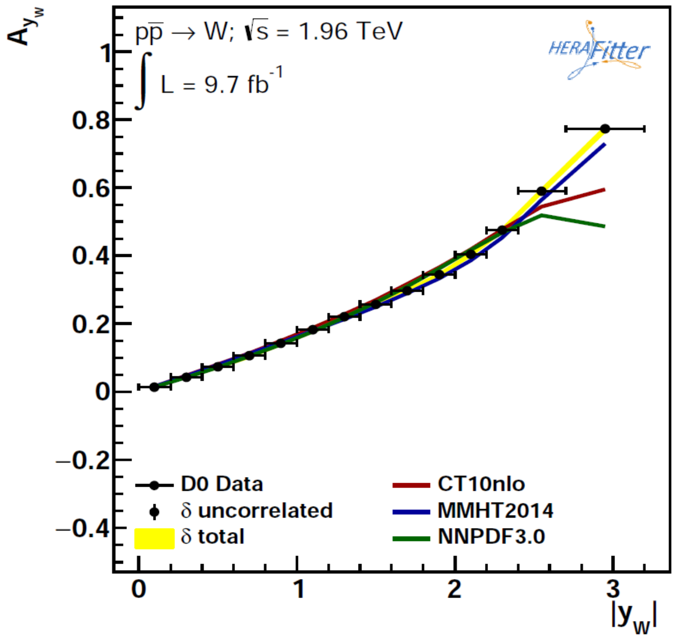
w/wo taking into account theory uncertainties:

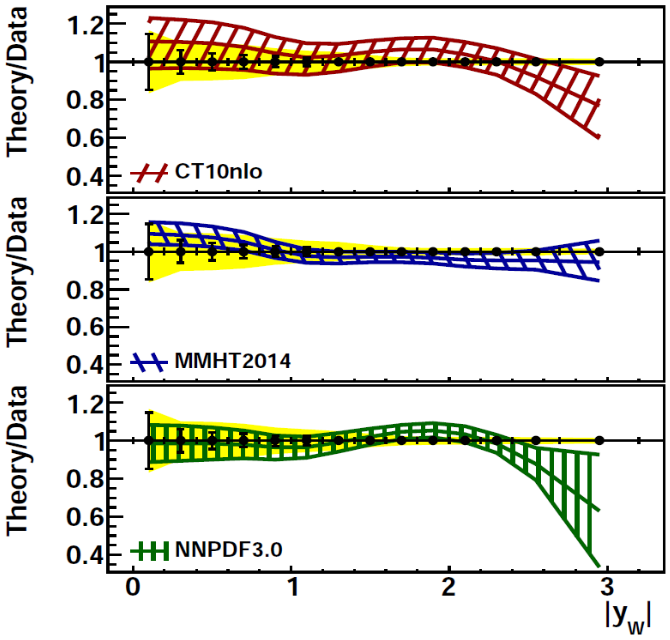
$$\chi^2 = \sum_i \left( \frac{\mu_i - m_i \left[ 1 + \sum_j b_j^{\mathrm{exp}} \gamma_{ji}^{\mathrm{exp}} + \sum_j b_j^{\mathrm{theo}} \gamma_{ji}^{\mathrm{theo}} \right]}{\Delta_i} \right)^2 + \sum_j (b_j^{\mathrm{exp}})^2 + \sum_j (b_j^{\mathrm{theo}})^2$$

| PDF set | CT10nlo | MMHT2014 | NNPDF3.0 |
|---|---|---|---|
| | $\chi^2$ / number of points | $\chi^2$ / number of points | $\chi^2$ / number of points |
| D0 muon charge asymmetry in $W \to \mu\nu$ | 13 / 10 | - | 12 / 10 |
| CDF $W$ charge asymmetry in $W \to e\nu$ | 14 / 13 | - | 15 / 13 |
| D0 $W$ charge asymmetry in $W \to e\nu$ | 8 / 14 | 5/14 | 2 / 14 |
| PDF correlated $\chi^2$ | 3 | 2 | 7 |
| Total $\chi^2$ / dof | 39 / 37 | 7 / 14 | 36 / 37 |
| Total $\chi^2$ / dof without PDFs uncertainties | 369/37 | 25/14 | 906 / 37 |

Good description

→ no tension between Tevatron data observed



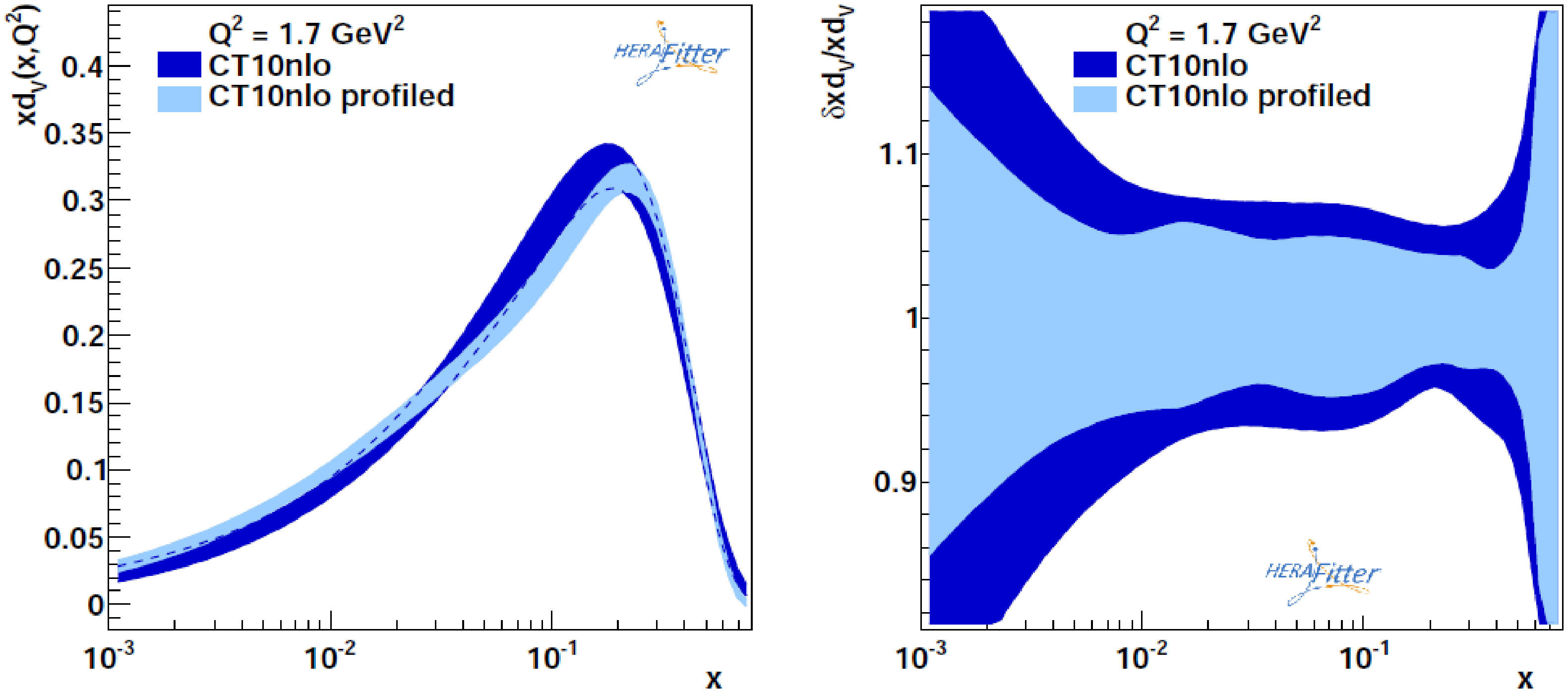


# Impact on PDFs Using Profiling Technique

Impact of Tevatron data on PDFs can be studied by minimizing data to theory $\chi^2$ vs nuisance parameters corresponding to PDF eigenvectors (“profiling”)

$$\chi^2 = \sum_i \left( \frac{\mu_i - m_i\left[1 + \sum_j b_j^{\text{exp}}\gamma_{ji}^{\text{exp}} + \sum_j b_j^{\text{theo}}\gamma_{ji}^{\text{theo}}\right]}{\Delta_i} \right)^2 + \sum_j (b_j^{\text{exp}})^2 + \sum_j (b_j^{\text{theo}})^2$$

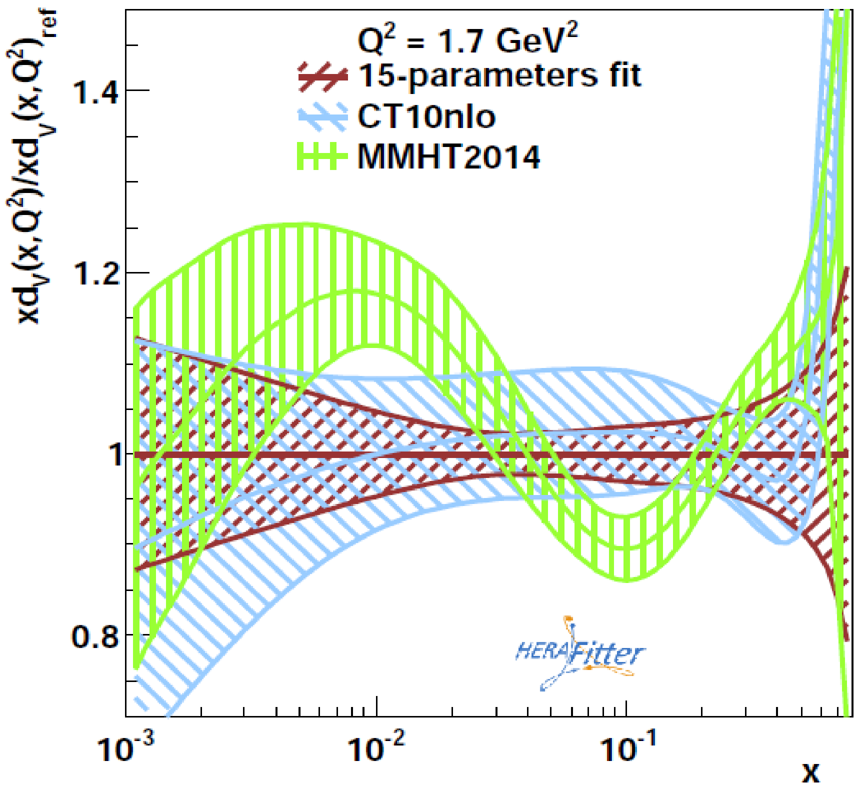
$\mu_i$ - data, $m_i$ – theory, $\beta_j^{\text{theo}}$ – nuisance parameters of theory uncertainties (PDF)
(asymmetric uncertainties are taken into account)

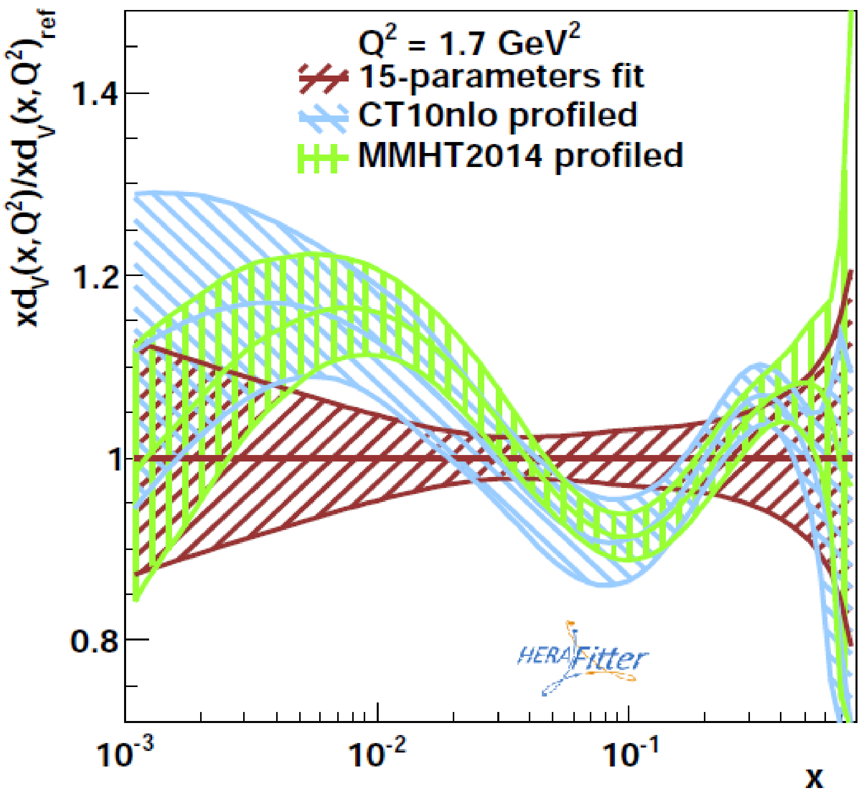


→ similar as in PDF fit case, a significant reduction on the $d_v$ quark is observed

# Impact on PDFs Using Profiling Technique

Original CT and MMHT PDFs → Profiled CT and MMHT PDFs





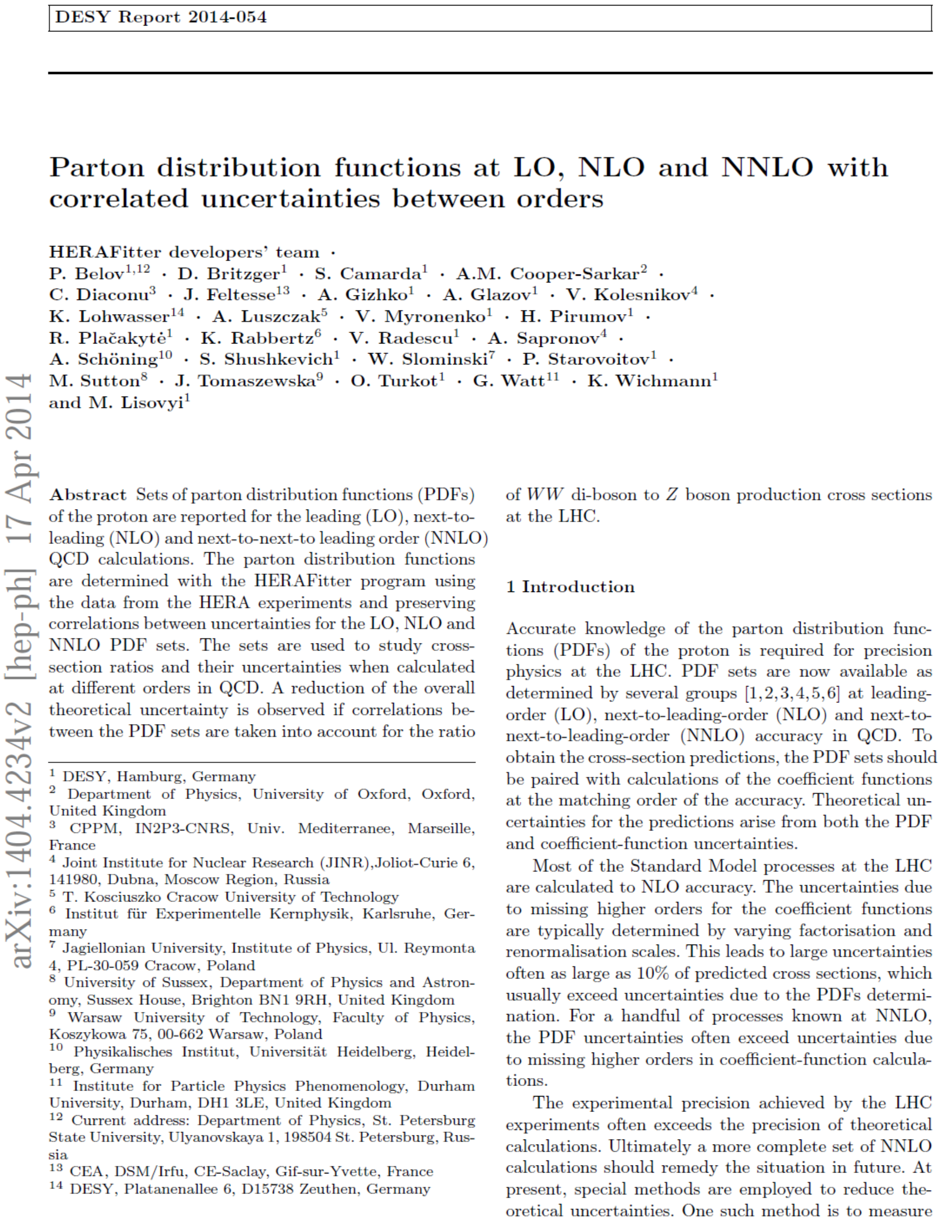
*only the measurements that are not included in the each of PDF sets are considered for the profiling*

→ improved agreement of the d-valence distribution between the MMHT2014 and CT10nlo PDF sets

The first xFitter Developers Team publication:

EPJC (2014) 74: 3039

PDFs at LO, NLO, NNLO with correlated uncertainties between orders



DESY Report 2014-054

arXiv:1404.4234v2 [hep-ph] 17 Apr 2014

# Parton distribution functions at LO, NLO and NNLO with correlated uncertainties between orders

**HERAFitter developers' team** ·
**P. Belov**[1,12] · **D. Britzger**[1] · **S. Camarda**[1] · **A.M. Cooper-Sarkar**[2] ·
**C. Diaconu**[3] · **J. Feltesse**[13] · **A. Gizhko**[1] · **A. Glazov**[1] · **V. Kolesnikov**[4] ·
**K. Lohwasser**[14] · **A. Luszczak**[5] · **V. Myronenko**[1] · **H. Pirumov**[1] ·
**R. Plačakytė**[1] · **K. Rabbertz**[6] · **V. Radescu**[1] · **A. Sapronov**[4] ·
**A. Schöning**[10] · **S. Shushkevich**[1] · **W. Slominski**[7] · **P. Starovoitov**[1] ·
**M. Sutton**[8] · **J. Tomaszewska**[9] · **O. Turkot**[1] · **G. Watt**[11] · **K. Wichmann**[1]
**and M. Lisovyi**[1]

**Abstract** Sets of parton distribution functions (PDFs) of the proton are reported for the leading (LO), next-to-leading (NLO) and next-to-next-to leading order (NNLO) QCD calculations. The parton distribution functions are determined with the HERAFitter program using the data from the HERA experiments and preserving correlations between uncertainties for the LO, NLO and NNLO PDF sets. The sets are used to study cross-section ratios and their uncertainties when calculated at different orders in QCD. A reduction of the overall theoretical uncertainty is observed if correlations between the PDF sets are taken into account for the ratio of $WW$ di-boson to $Z$ boson production cross sections at the LHC.

## 1 Introduction

Accurate knowledge of the parton distribution functions (PDFs) of the proton is required for precision physics at the LHC. PDF sets are now available as determined by several groups [1,2,3,4,5,6] at leading-order (LO), next-to-leading-order (NLO) and next-to-next-to-leading-order (NNLO) accuracy in QCD. To obtain the cross-section predictions, the PDF sets should be paired with calculations of the coefficient functions at the matching order of the accuracy. Theoretical uncertainties for the predictions arise from both the PDF and coefficient-function uncertainties.

Most of the Standard Model processes at the LHC are calculated to NLO accuracy. The uncertainties due to missing higher orders for the coefficient functions are typically determined by varying factorisation and renormalisation scales. This leads to large uncertainties often as large as 10% of predicted cross sections, which usually exceed uncertainties due to the PDFs determination. For a handful of processes known at NNLO, the PDF uncertainties often exceed uncertainties due to missing higher orders in coefficient-function calculations.

The experimental precision achieved by the LHC experiments often exceeds the precision of theoretical calculations. Ultimately a more complete set of NNLO calculations should remedy the situation in future. At present, special methods are employed to reduce theoretical uncertainties. One such method is to measure

---

[1] 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, Ul. 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] CEA, DSM/Irfu, CE-Saclay, Gif-sur-Yvette, France
[14] DESY, Platanenallee 6, D15738 Zeuthen, Germany

# Motivation

Predictions for various processes at LHC are available at different orders in pQCD

→ PDFs and coefficient functions ideally should be paired at the same order

→ uncertainties rise from both, PDFs and coefficient functions, missing higher order uncertainties often determined by the scale variation

Ratios of cross sections are used to reduce theoretical uncertainties:

| | |
|---|---|
| $\frac{\mathrm{NLO}_X}{\mathrm{NLO}_Y}$ | PDF uncertainties cancel<br>large scale uncertainty |
| $\frac{\mathrm{NLO}_X}{\mathrm{NNLO}_Y}$ | improved scale uncertainty<br>No cancellation of PDF uncertainty |
| $\frac{\mathrm{NLO}_X + \mathrm{NNLO\ PDF}}{\mathrm{NNLO}_Y + \mathrm{NNLO\ PDF}}$ | PDF uncertainties cancel<br>improved scale uncertainty<br>not clear definition in pQCD |
| $\frac{\mathrm{NLO}_X + \mathrm{NLO\ PDF}_{\mathrm{corr}}}{\mathrm{NNLO}_Y + \mathrm{NNLO\ PDF}_{\mathrm{corr}}}$ | PDF uncertainties cancel<br>improved scale uncertainty |

Requires preserved correlations in PDF uncertainties at different orders

# QCD Analysis Settings

EPJC (2014) 74: 3039

A QCD fit to HERA I data (JHEP 1001, 2010, 109) using xFitter

Parametrisation:

$$x\mathrm{u_v}(x) = A_{\mathrm{u_v}}\, x^{B_{\mathrm{u_v}}}\, (1-x)^{C_{\mathrm{u_v}}}\, (1+E_{\mathrm{u_v}} x^2),$$
$$x\mathrm{d_v}(x) = A_{\mathrm{d_v}}\, x^{B_{\mathrm{d_v}}}\, (1-x)^{C_{\mathrm{d_v}}},$$
$$x\overline{\mathrm{U}}(x) = A_{\overline{\mathrm{U}}}\, x^{B_{\overline{\mathrm{U}}}}\, (1-x)^{C_{\overline{\mathrm{U}}}},$$
$$x\overline{\mathrm{D}}(x) = A_{\overline{\mathrm{D}}}\, x^{B_{\overline{\mathrm{D}}}}\, (1-x)^{C_{\overline{\mathrm{D}}}},$$
$$x\mathrm{g}(x) = A_{\mathrm{g}}\, x^{B_{\mathrm{g}}}\, (1-x)^{C_{\mathrm{g}}} + A'_{\mathrm{g}}\, x^{B'_{\mathrm{g}}}\, (1-x)^{C'_{\mathrm{g}}}$$

\+ additional constraints and assumptions

→ other setting largely follow JHEP 1001, 2010, 109

model and parametrisation uncertainties treated correlated between orders

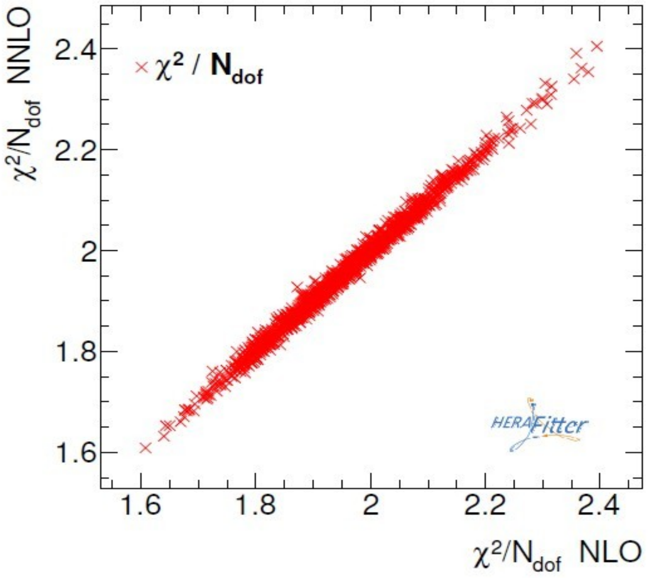
## MC replica method used to preserve the correlations:

→ 1337 MC replicas of the data fluctuating the inputs within uncertainties using Gaussian prob densities

→ perform a consistent fit of PDFs at different orders to each replica

central PDF = average over replicas,
PDF uncertainty = RMS over replicas

correlation of $\chi^2$ between orders



# Eigenvector Representation

EPJC (2014) 74: 3039

## Eigenvector representation is often more convenient than MC replica

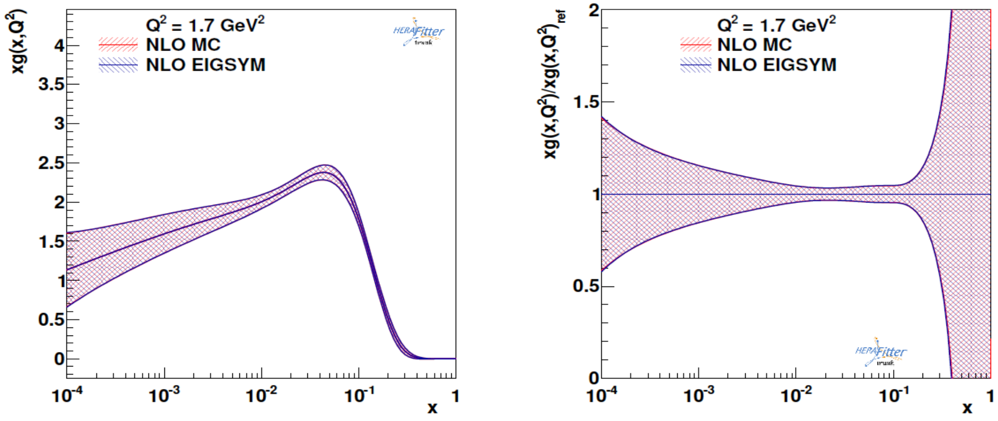
employ the method suggested for extraction of META PDFs (arXiv:1401.0013)

→ build the covariance matrix

→ diagonalise matrix and keep only leading eigenvectors

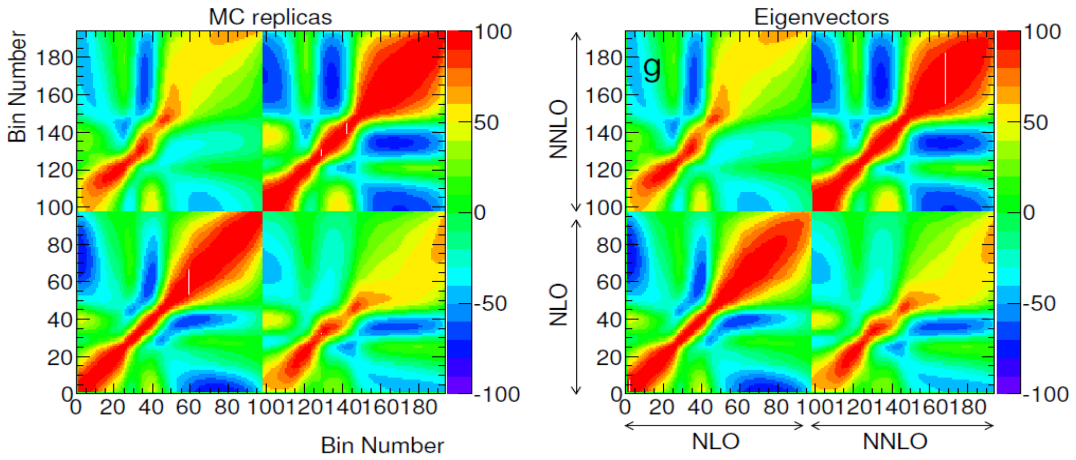
→ evolve tabulated values (central and eigenvectors) using DGLAP evolution

Comparison of PDFs determined with MC method and its eigenvector representation



→ very good agreement between PDFs

Correlation coefficients among PDFs:



→ PDFs show high degree of correlation at neighboring x bins (smoothness of parametrisation)

→ strong correlation between NLO and NNLO PDFs

# Comparison to Data: WW/Z ratio

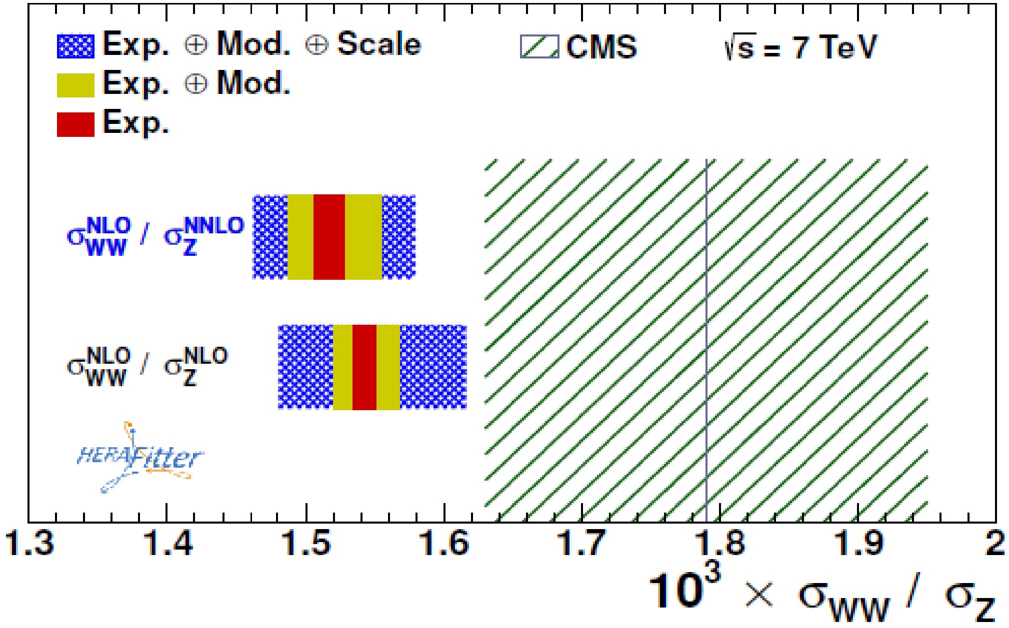
EPJC (2014) 74: 3039

(arXiv:1306.1126)

Used case example: WW/Z ratio measurement from CMS

→ build the cross section ratio using the correlated PDFs

| Ratio | Value $\times 10^{-3}$ | Exp. PDF $\times 10^{-3}$ | Mod. PDF $\times 10^{-3}$ | Scale $\times 10^{-3}$ |
|---|---|---|---|---|
| $\frac{\sigma_{WW}^{NLO}}{\sigma_{Z}^{NLO}}$ | 1.543 | $\pm 0.008$ | $^{+0.023}_{-0.021}$ | $^{+0.069}_{-0.058}$ |
| $\frac{\sigma_{WW}^{NLO}}{\sigma_{Z}^{NNLO}}$ | 1.517 | $\pm 0.010$ | $^{+0.036}_{-0.027}$ | $^{+0.050}_{-0.046}$ |



→ predictions agree with the data within 1-2σ

→ the total theoretical uncertainty is reduced by 30-40%

→ mixed-order calculations with correlated PDFs help to reduce PDF and scale uncertainties