HERAFitter

A M Cooper-Sarkar
University of Oxford
ICHEP 2014, Valencia

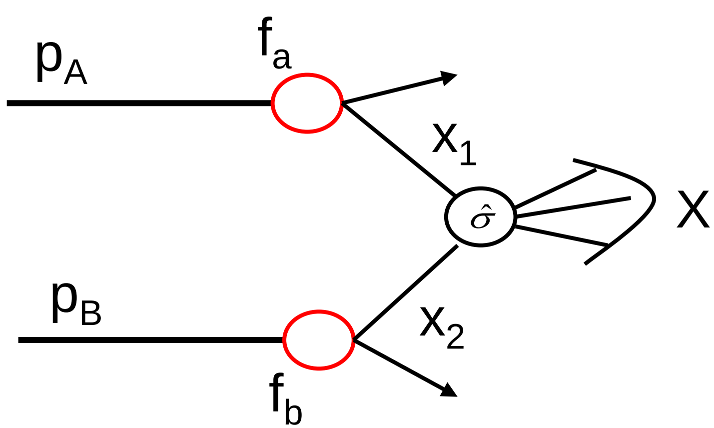
- HERAFitter is an open source QCD framework and it has proved to be a successful platform that is well integrated in the high energy physics community.
- HERAFitter infrastructure has increased the scientific output of the HERA and LHC data, provides a flexible environment for theory benchmarking.
- Stable release: herafitter-1.0.0, can be found at *www.herafitter.org.*

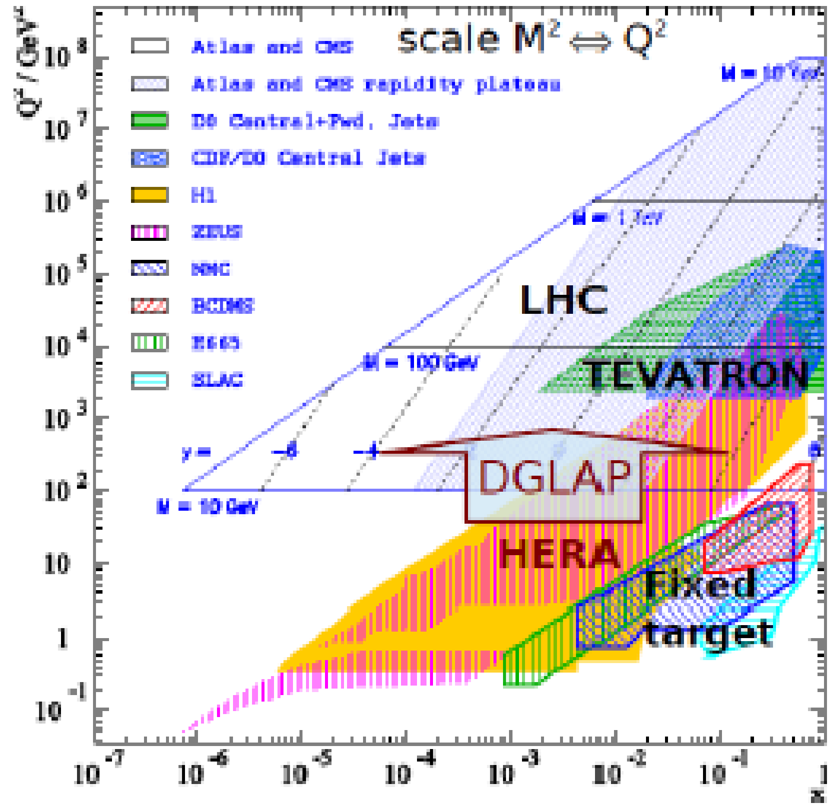
- Develops in a close collaboration of experimentalists from HERA (H1 and ZEUS) an LHC (ALICE, ATLAS and CMS) collaborations and theorists
- $\sim$ 30 developers

Uncertainties on Parton Distribution Functions (PDFs) limit our knowledge of cross sections whether SM or BSM.

$$\sigma_X = \sum_{a,b} \int_0^1 dx_1 dx_2\, f_a(x_1, \mu_F^2)\, f_b(x_2, \mu_F^2)$$
$$\times\ \hat{\sigma}_{ab\to X}\left(x_1, x_2, \{p_i^\mu\}; \alpha_S(\mu_R^2), \alpha(\mu_R^2), \frac{Q^2}{\mu_R^2}, \frac{Q^2}{\mu_F^2}\right)$$

where X=W, Z, D-Y, H, high-E_T jets, prompt-γ and σ is known to some fixed order in pQCD and EW or in some leading logarithm approximation (LL, NLL, …) to all orders via re-summation





Current knowledge of PDFs is dominated by the HERA data
PDFs are evolved up in scale using the DGLAP equations to make predictions for LHC cross -sections

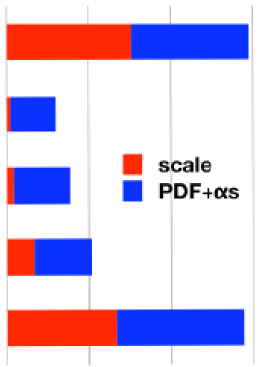
Accurate knowledge of Parton Distribution Functions is essential for precision physics at the LHC

- Uncertainty on PDFs is one of the main uncertainties on the SM Higgs cross section
- And on M_W
- And on predictions for high scale BSM production

Different PDF groups use different methodologies as well as different data

PDFs and the Higgs

		σ (8 TeV)	uncertainty
NNLL QCD +NLO EW	gg→H	19.5 pb	14.7%
NNLO QCD +NLO EW	VBF	1.56 pb	2.9%
	WH	0.70 pb	3.9%
	ZH	0.39 pb	5.1%
NLO QCD	ttH	0.13 pb	14.4%



HERAFitter is an open source QCD fitting code which allows the comparison of the use of different data sets and different methodology, e.g. choice of heavy flavour scheme, using a common platform

It can be used to extract PDFs ------- it is used for the HERAPDF---------------- and to assess the impact of new data and hence to improve the sensitivity of new measurements to PDF- this is already being done within ATLAS and CMS before finalisation of measurements and publication

How do we determine Parton Distribution Functions ?

Parametrise the parton distribution functions (PDFs) at Q^2_0 (~1-2 GeV^2)- Use (NN)LO QCD DGLAP equations to evolve these PDFs to $Q^2 > Q^2_0$

Construct the measurable cross-sections by convoluting PDFs with (NN)LO matrix elements: make predictions for ~2000 data points across the x,Q^2 plane- Perform χ2 fit to the data

Formalism
(NN)LO DGLAP
MSbar factorisation
Q_0^2
functional form @ Q_0^2
Heavy quark treatment…

$$f_i(x,Q^2) \pm \delta f_i(x,Q^2)$$

$$\alpha_S(M_Z)$$

Who?
ABM, CTEQ, MRST, HERA,JR,NNPDF

Data
DIS (Fixed target and HERA)
Drell-Yan (E605, E772, E866, …)
High E_T jets (Tevatron)
W,Z rapidity and asymmetry (Tevatron)
etc. and NOW LHC data…

http://durpdg.dur.ac.uk/hepdata/
http://projects.hepforge.org/lhapdf/

LHAPDFv5/6

HERAFitter allows the choice of input parametrisation

- HERAPDF/MSTW style
- CTEQ style
- Chebyshevs
- Log Normal

And the choice to input a wide range of data in addition to the HERA-DIS data e.g. LHC data

Measurement at LHC	PDF sensitivity
Jets	high x quarks and gluons (alphas)
Inclusive W, Z and asymmetries	quark flavour separation (u,d,s)
Low and high mass Drell-Yan	quarks at low and high x (u,d)
W + charm	Direct sensitivity to s-quark
Isolated photons	medium - x gluons
Single top	u,d and b quark
ttbar (total, differential)	Medium-x gluon (alphas)
W,Z production with jets	Medium-x gluon
Z+b production	sensitive to b-quark

HERAFitter allows various ways of treating experimental systematic uncertainties

Data should be provided with information on correlated systematic uncertainties this can be in the form of a correlation matrix, or in terms of correlated shifts for each systematic source. These are used in the χ2 minimisation as:

- Nuisance parameter representation
  - Simple form

$$\chi^2_{exp}(m,b) = \sum_i \frac{\left(m_i - \sum_j \gamma^i_j b_j - \mu_i\right)^2}{(\delta_{i;stat}\mu_i)^2 + (\delta_{i;uncor}\mu_i)^2} + \sum_j b_j^2$$

  - Scaled form

$$\chi^2_{exp}(m,b) = \sum_i \frac{\left(m_i - \sum_j \gamma^i_j b_j - \mu_i\right)^2}{\delta^2_{i;stat}\mu_i\left(m_i - \sum_j \gamma^i_j b_j\right) + (\delta_{i;uncor}m_i)^2} + \sum_j b_j^2 + \text{log penalty}$$

- Covariance matrix representation

$$\chi^2_{exp}(m) = \sum_{ij}(m_i - \mu_i)\, C^{-1}_{ij}\,(m_j - \mu_j)$$

- Mixed form (covariance matrix and nuisance parameters):

$$\chi^2_{exp}(m,b) = \sum_{ij}\left(m_i - \sum_k \Gamma^i_k(m_i)b_k - \mu_i\right) C^{-1}_{ij;stat}(m_i,m_j)\left(m_j - \sum_k \Gamma^j_k(m_j)b_k - \mu_j\right) + \sum_k b_k^2$$

Experimental errors can be propagated to the PDFs by several methods

i) Hessian method – nuisance parameters are fitted, χ2 tolerance T >1 can also be used to account for marginally compatible input data sets

ii) Offset method- nuisance parameters are applied as 1σ shifts

iii) MC-method—data points are shifted randomly within their 1σ limits to form MC replicas. This can also allow for asymmetric uncertainties

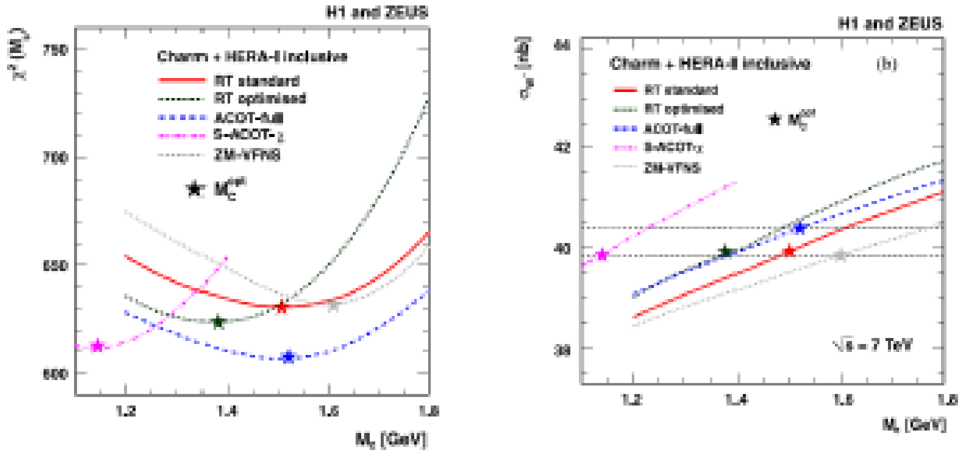
Herafitter allows comparison of different theoretical formalisms

QCDNUM is used for DGLAP evolution, but QCDRAD can be use and APFEL is coming

Coefficient functions can be calculated in **various heavy quark schemes**

- Variable Flavour Number schemes a la MSTW, CTEQ or NNPDF(coming)
- Fixed Flavour Number schemes as used by ABM

Variety of schemes were tested by H1/ZEUS collaborations in F_2^{charm} *analysis*



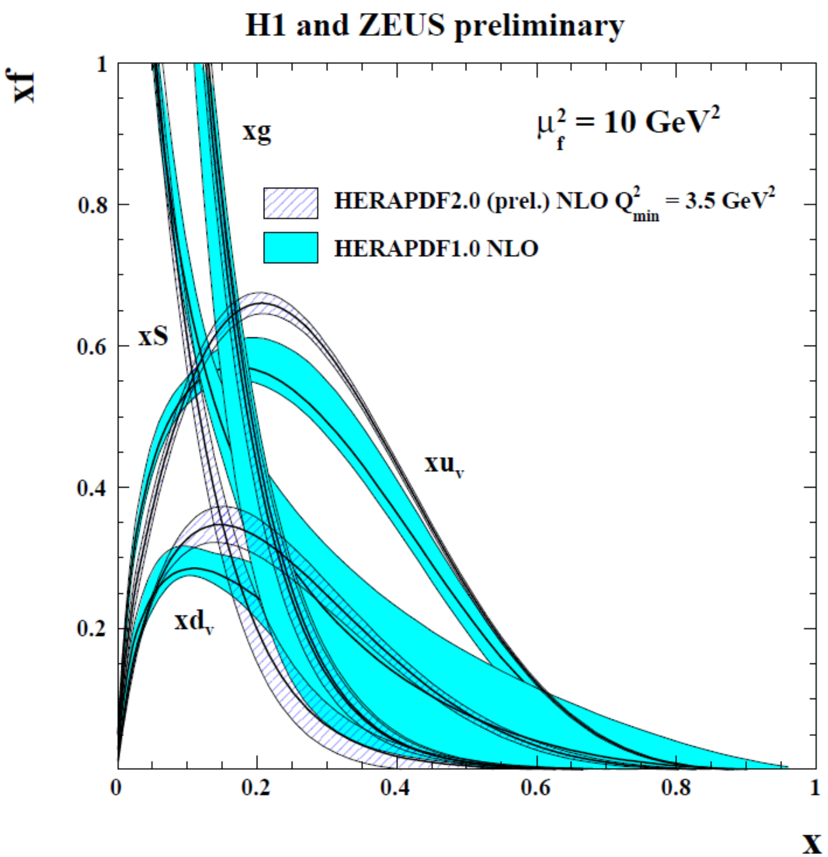
[Eur. Phys. J. C73 (2013) 2311]

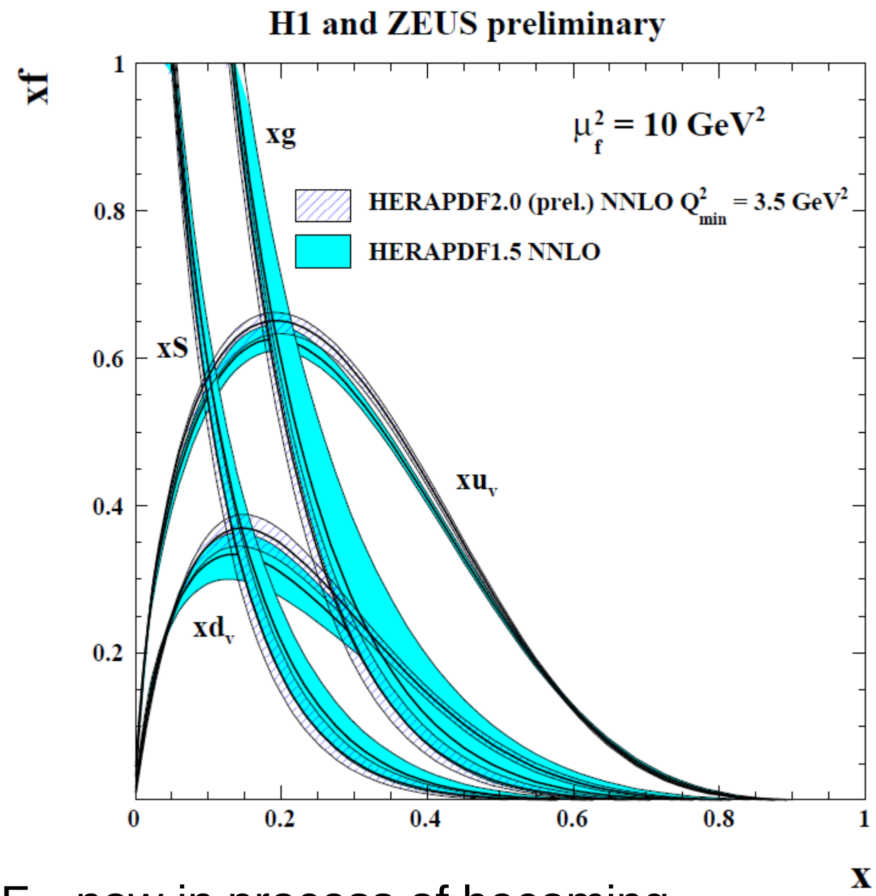
Spread in predictions for *W* and *Z* @LHC is reduced significantly when predictions are evaluated at the optimal m_{charm} determined from F_2^{charm}

Tools like FastNLO and Applgrid are implemented to make fast calculations for NLO (and NNLO) cross sections which are otherwise too slow for input to a fit

- NLO jet production can be done by FastNLO or Applgrid interfaced NLOJet++
- NLO Drell-Yan and W,Z production can now be done exactly using Applgrid interfaced to MCFM.
- NNLO Drell-Yan using DYNNLO should be available soon
- Top production, from MCFM at NLO, will be extended to NNLO by DiffTop
- Hathor can also be used for Top

Examples of the use of HERAFitter





The first was the production of the HERAPDF-- now in process of becoming HERAPDF2.0

HERAfitter also contains the HERAverager tool which is used to combine ZEUS and H1 data, and now to combine electron and muon channel data at the LHC

Quantitative comparison of agreement between new data sets and predictions from various PDFs– not just HERAPDF

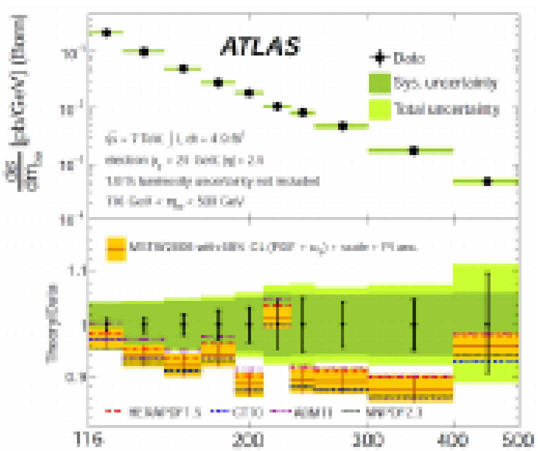
HERAFitter provides a quantitative assessment of level of agreement between data and theory by taking into account theoretical and experimental uncertainties

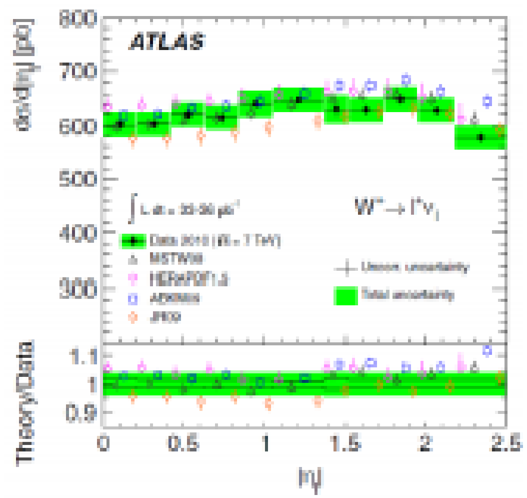
$$\chi^2 = \sum_i \left(\frac{\mu_i - m_i \left[1 + \sum_j b_j^{\text{exp}} \gamma_{ji}^{\text{exp}} + \sum_j b_j^{\text{theo}} \gamma_{ji}^{\text{theo}} \right]}{\Delta_i} \right)^2 + \sum_j (b_j^{\text{exp}})^2 + \sum_j (b_j^{\text{theo}})^2$$

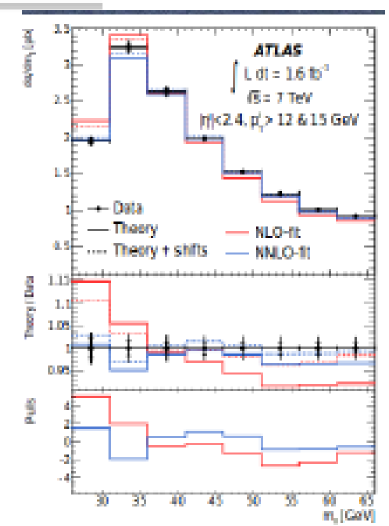
Ex: 30 points from ATLAS WZ 2010 vs NNLO predictions

PDF set	Central PDF	With PDF uncertainties
CT10	34.1	32.0
MSTW08	72.0	49.7
HERAPDF1.5eig	43.1	39.2

Used in ATLAS publications:







Phys. Lett. B 725 (2013) 223

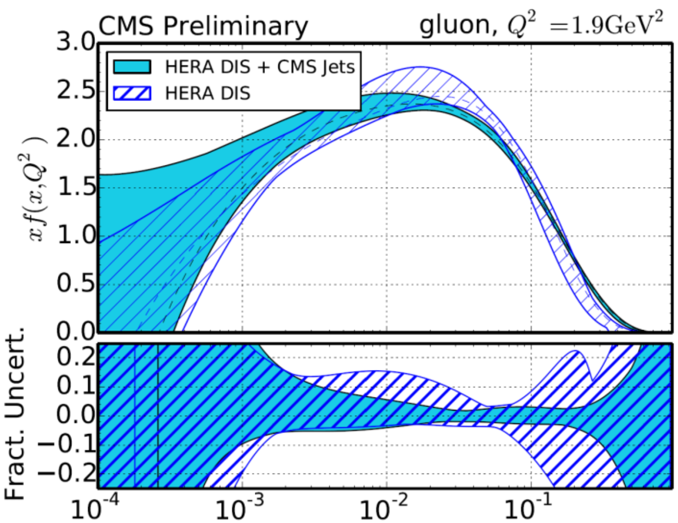
Phys. Rev. D 85 (2012) 072004

arXiv:1404.1212

Inputting new data sets to a PDF fit to assess –and improve- their impact

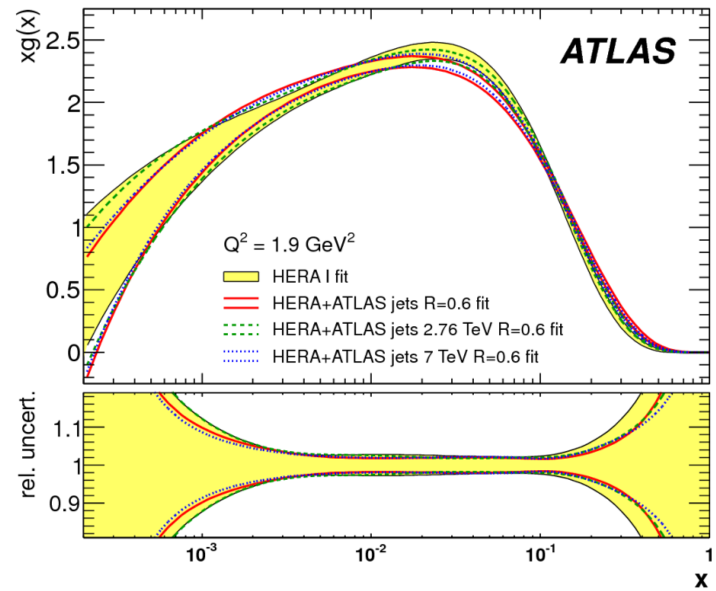
Both CMS and ATLAS have added jet data and W,Z data to the HERA-I data to asses their impact.

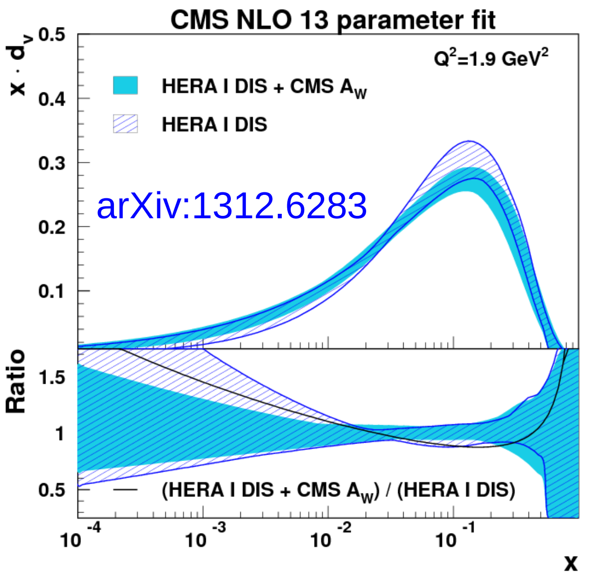
CMS-SMP-12-028



Improving gluon PDF

EPJC(2013)73 2509



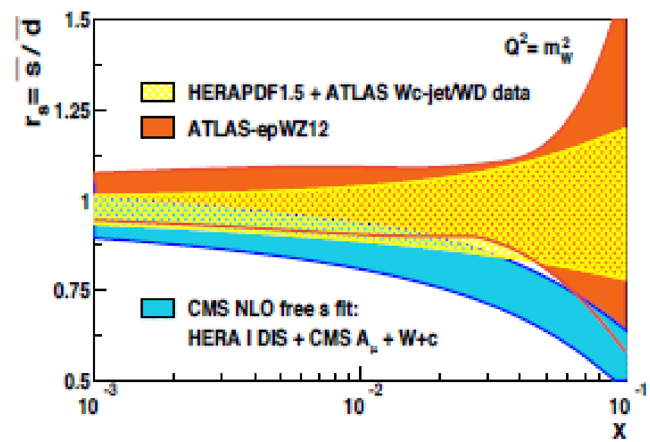


arXiv:1312.6283

Improving valence PDFs

Determining strangeness fraction

arXiv:1402.6263+arxIv:1312.6283



And this can also be done by NNPDF-style reweighting

PDF sets at LO/NLO/NNLO with correlated uncertainties

arXiv:1404.4234

Theoretical predictions are available at different orders

- LO used in parton shower MCs
- NLO for most predictions
- NNLO for a few predictions

Factorisation theorem: $\sigma \approx \hat{\sigma} \otimes \mathrm{PDF}$

Uncertainties come from the PDFs and the sub-process cross-sections
Scale uncertainties affect the sub-process cross-sections more at lower orders

Ratios are often used as a way of cancelling experimental uncertainties.
But the corresponding theoretical uncertainties may not cancel out

$$\frac{\hat{\sigma}_X^{NLO} \otimes PDF_{NLO}}{\hat{\sigma}_Y^{NLO} \otimes PDF_{NLO}}$$

PDF uncertainties cancel
large scale uncertainty

Large scale uncertainty because NLO calculation

$$\frac{\hat{\sigma}_X^{NLO} \otimes PDF_{NLO}}{\hat{\sigma}_Y^{NNLO} \otimes PDF_{NNLO}}$$

improved scale uncertainty
No cancellation of PDF uncertainty

Improve this by going to NNLO but what if this is only available for ONE of the cross sections?

$$\frac{\hat{\sigma}_X^{NLO} \otimes PDF_{NLO}^{corr}}{\hat{\sigma}_Y^{NNLO} \otimes PDF_{NNLO}^{corr}}$$

PDF uncertainties cancel
improved scale uncertainty

Preserve correlations between PDFs of different orders

This has been done using the HERAPDF formalism and HERA-1 combined data varying the model and parametrization assumptions (as in JHEP 1001, 2010, 109)

MC replica method used to preserve the correlations:

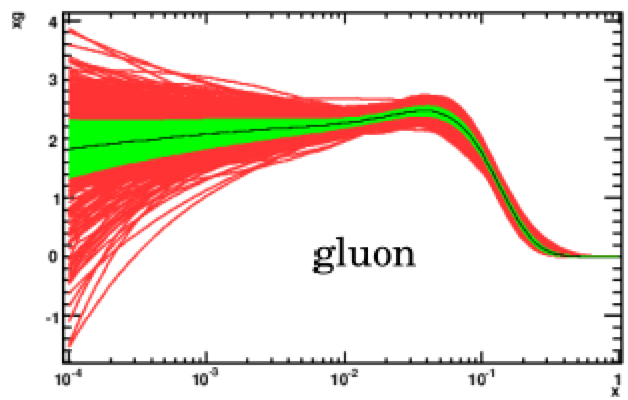
→ 1337 MC replicas of the data fluctuating the inputs within uncertainties using Gaussian prob densities

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

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

model and param uncertainties treated correlated between orders

χ^2/N_{dof} NLO



In practice an eigenvector representation is can be more convenient than MC replicas

The MC replica results can be converted using the method used to extract META-PDFs (arXiv:1401.0013)

→ build the covariance matrix

→ diagonalise matrix and keep only leading eigenvectors

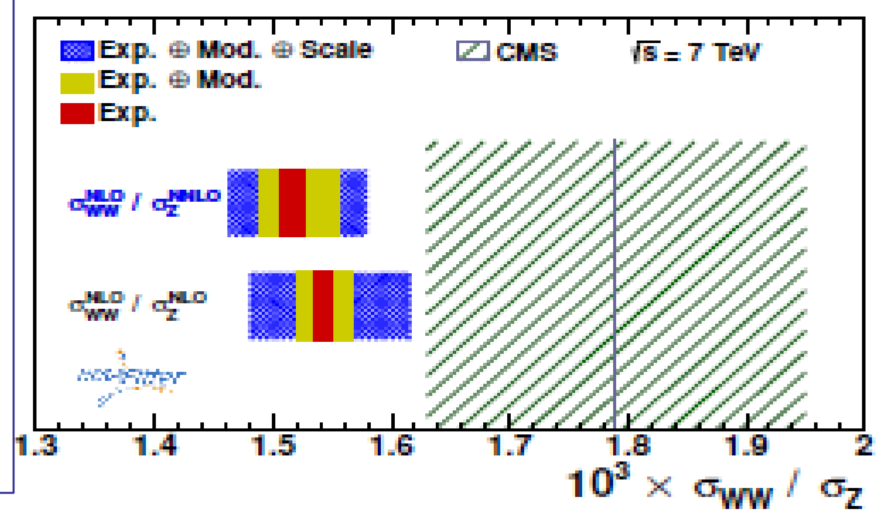
This preserves strong correlations between NLO and NNLO PDFs.

These PDFs have been used to calculate the WW/Z ratio and compare to CMS data arXiv:1306.1126

The total theoretical uncertainty of the calculation is reduced by 30-40% if σ_Z is calculated to NNLO because of reduced scale uncertainties

BUT ONLY because the PDF uncertainties at NLO and NNLO are correlated

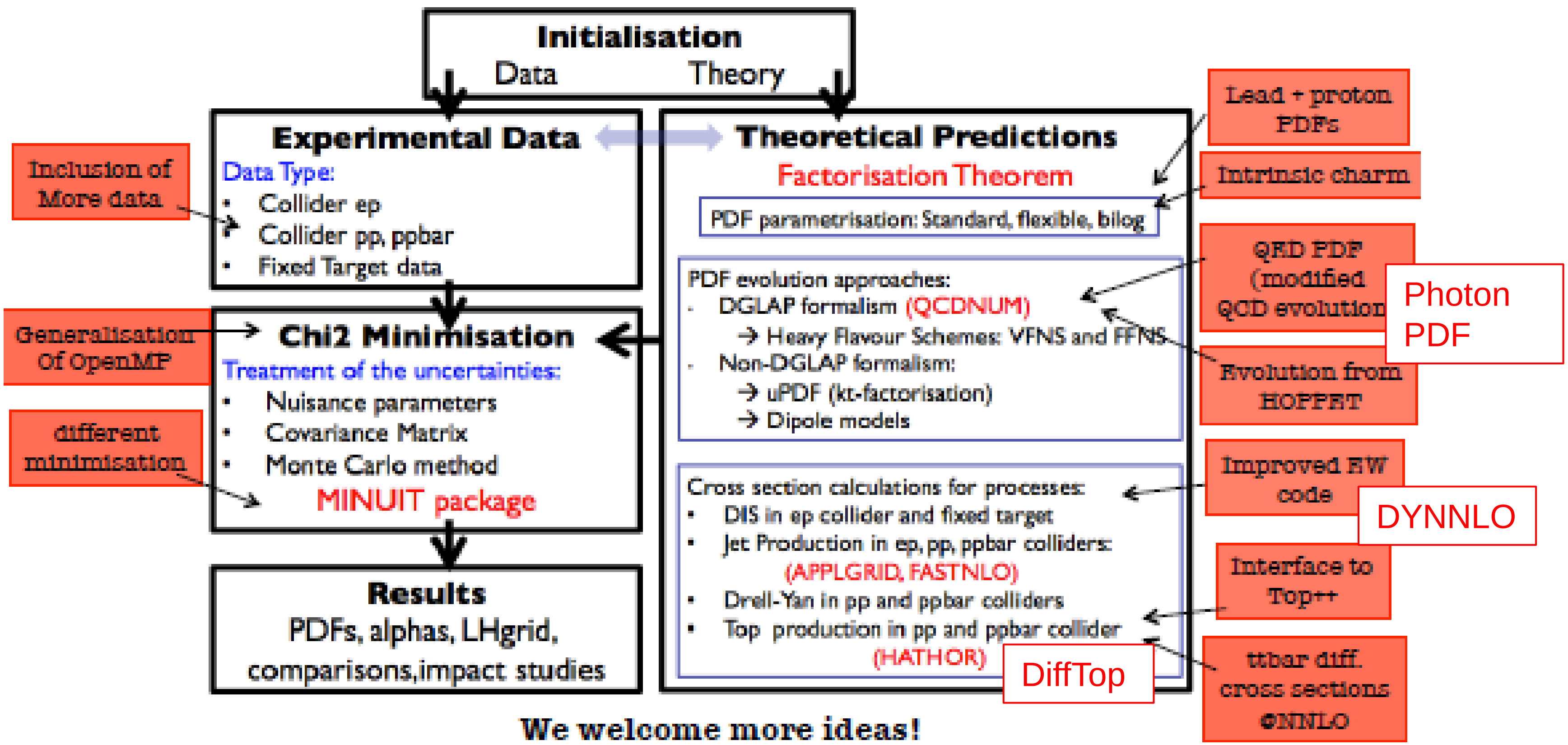
(arXiv:1306.1126)



HERAFitter Prospects

HERAFitter has a modular structure facilitating fast developments

- Many new developments are planed to be implemented in future releases:



We welcome more ideas!

HERAFitter

- Oct — **Proton Structure in the LHC era event @ DESY**
- …
- Apr — **First paper by HERAFitter developers' team**
- Mar — HERAFitter plenary talk @ QCD Moriond

2014

- Dec — **First HERAFitter Stable Release**
- Sep — Award winning poster at the conferences
- Mar — Third HERAFitter Beta Release

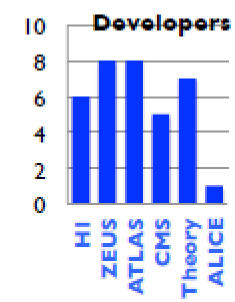
2013

- Oct — **First PDF School based on HERAFitter**
- May — Second HERAFitter Beta Release
- Mar — **First LHC paper using HERAFitter**
- Feb — First HERAFitter Workshop

2012

- Nov — First HERAFitter Invited presentation
- Oct — Presented to the LHC Community
- Sep — **First HERAFitter Beta Release**

2011



- Standard Polynomials:

$$xf(x) = Ax^B(1-x)^C P_i(x),$$

- Log-Normal Distributions:

$$xf(x) = x^{p-b\log(x)}(1-x)^{q-\log(1-x)}$$

- Chebyshev Polynomials:

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

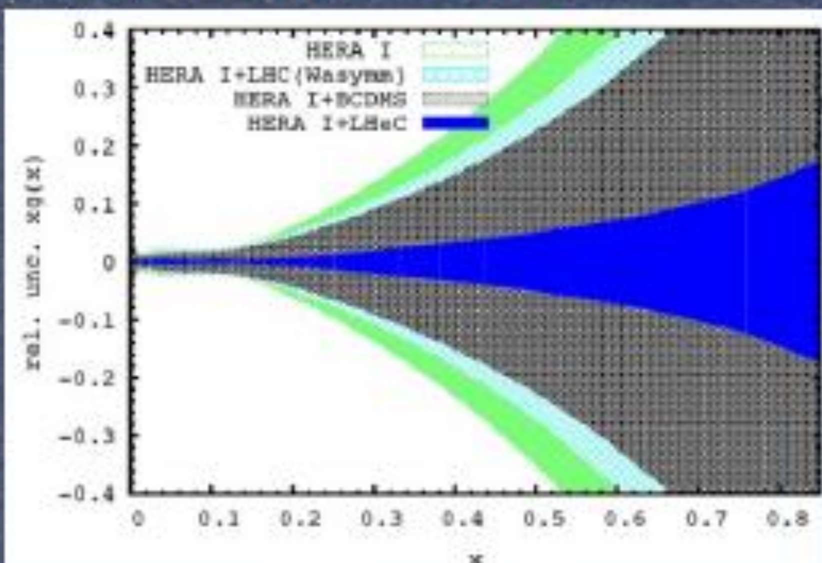
$$xS(x) = (1-x)\sum_{i=0}^{N_S-1} A_{S_i} T_i\left(-\frac{2\log x - \log x_{min}}{\log x_{min}}\right).$$

- Use of External PDFs via LHAPDF interface to construct theoretical predictions.

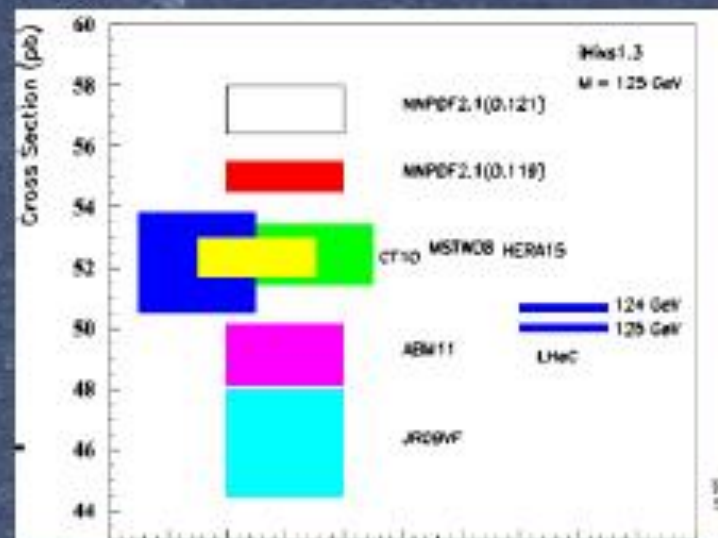
Sensitivity Studies

Platform can be used for sensitivities studies of the potential of future colliders:

- LHeC ep simulated data was used for to study sensitivity to PDFs:



- The output in LHAPDF format can be used for Higgs predictions:



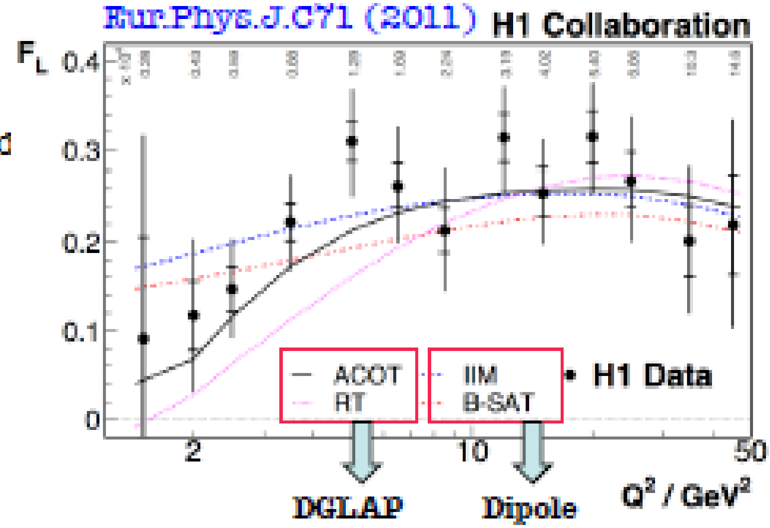
(Journal of Phys. G 39 (2012))

Going beyond DGLAP

As an alternative to DGLAP, HERAFitter includes also Dipole models:

- Studied by the H1 collaboration in comparing different models on FL:
  - Dipole Models implemented in HERAFitter:
    - GBW model: first model
    - IIM (based on BK-equation)
    - BGK (based on GBW, but gluon evolved using DLGAP)
  - DGLAP Models:
    - RT as used by MSTW group
    - ACOT as used by CTEQ group

Eur.Phys.J.C71 (2011)

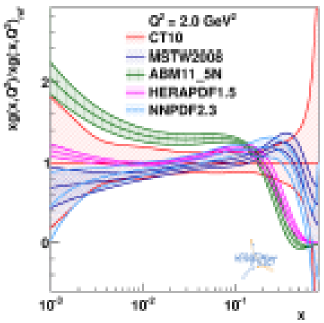


Unintegrated PDFs based on the kT-factorisation (CCFM) evolution.

- applicable only to NC ep scattering:
  https://www.herafitter.org/HERAFitter/HERAFitter/HERAFitterMeetings/Meeting2012-Oct-29?action=AttachFile&do=get&target=updf.pdf

Diffractive DIS PDF fits.

PDF Sensitivity study on prompt photon



First result with stable release HERAFitter 1.0.0

Assess:

- **Compatibility** between data and NLO predictions
- **Sensitivity** to the gluon PDF

HERAFitter 1.0.0 stable release provides automatic tools for:

- χ^2 comparison with and without PDF uncertainties
- PDF plots
- Data plots