HERAFitter Overview

Voica A. Radescu, Ringaile Placakyte (DESY)

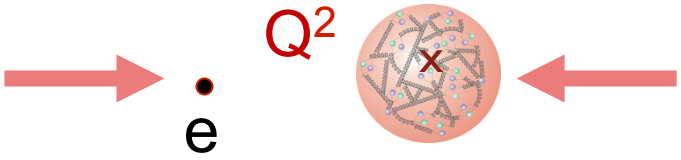

```
------------------------------------------------------------------------------------------

  HH    HH  EEEEEEE  RRRRR      AAA    FFFFFFF  II  TTTTTTTT  TTTTTTTT  EEEEEEE  RRRRR
  HH    HH  EE       RR   RR   AA AA   FF       II     TT        TT     EE       RR   RR
  HHHHHHH   EEEEE    RR   RR  AA   AA  FFFFF    II     TT        TT     EEEEEE   RR   RR
  HH    HH  EE       RRRRR    AAAAAAA  FF       II     TT        TT     EE       RRRRR
  HH    HH  EE       RR  RR   AA   AA  FF       II     TT        TT     EE       RR  RR
  HH    HH  EEEEEEE  RR   RR  AA   AA  FF       II     TT        TT     EEEEEEE  RR   RR

  Version 0.2.1
  http://herafitter.hepforge.org                          herafitter-help@desy.de
------------------------------------------------------------------------------------------
```

“Proton Structure in the LHC Era” school 22-24 October 2012

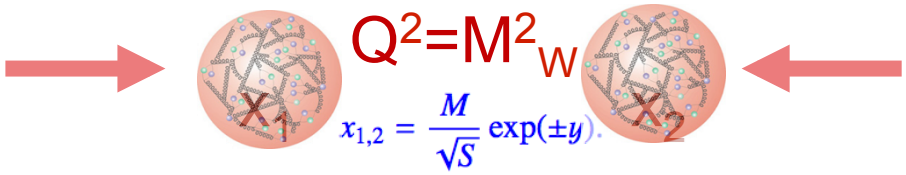
Proton Structure

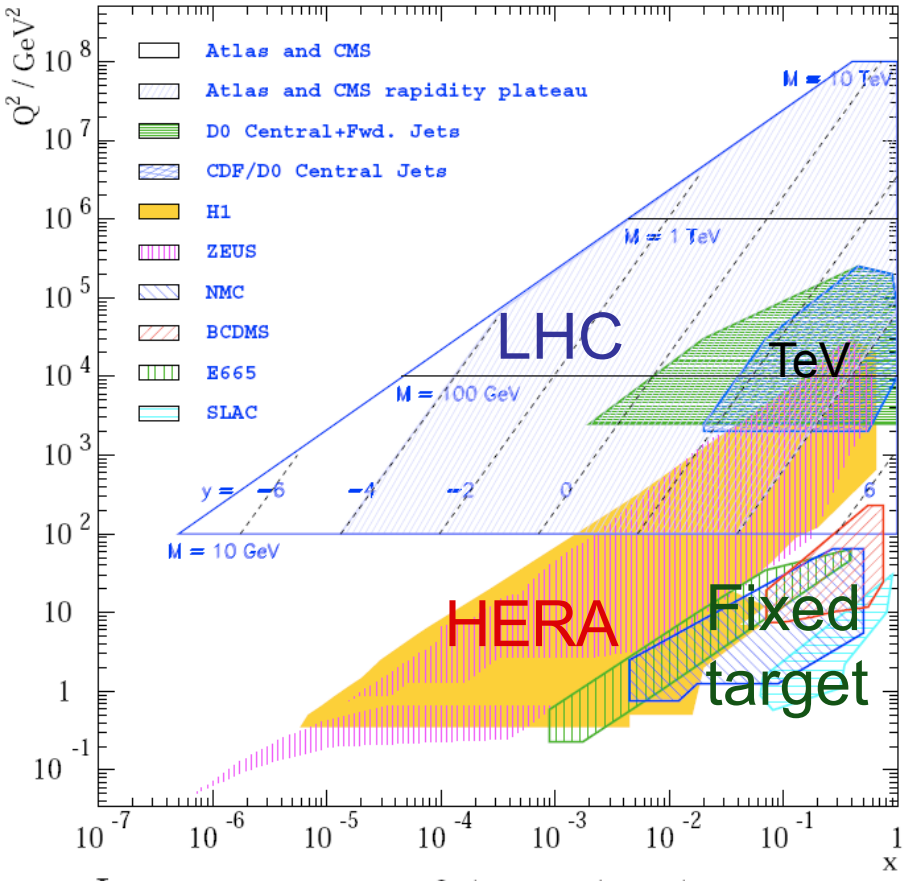
- Factorization theorem states that cross section can be calculated using universal partons x short distance calculable partonic reaction [see S. Alechin's lecture].

- Probing Proton Structure via Deep Inelastic Scattering using elementary particles such as:
  - Neutrinos, muons (fixed target experiments)
  - Electrons (fixed target and collider experiments)



- Knowledge on proton structure can be complemented by the collider experiments at Tevatron and LHC

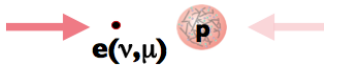




Persistent experimental effort over the last 40 years both by fixed-target and collider experiments around the world supported by the theoretical developments

PDF Constraints

- Main information on proton structure comes from DIS data at HERA (ep):
  - probes linear combination of quarks:
    - CC: provides constraints on valence quarks
    - NC: $F_2 \sim 0.44x(u+\bar{u}+c+\bar{c}) + 0.11x(d+\bar{d}+s+\bar{s}+b+\bar{b})$



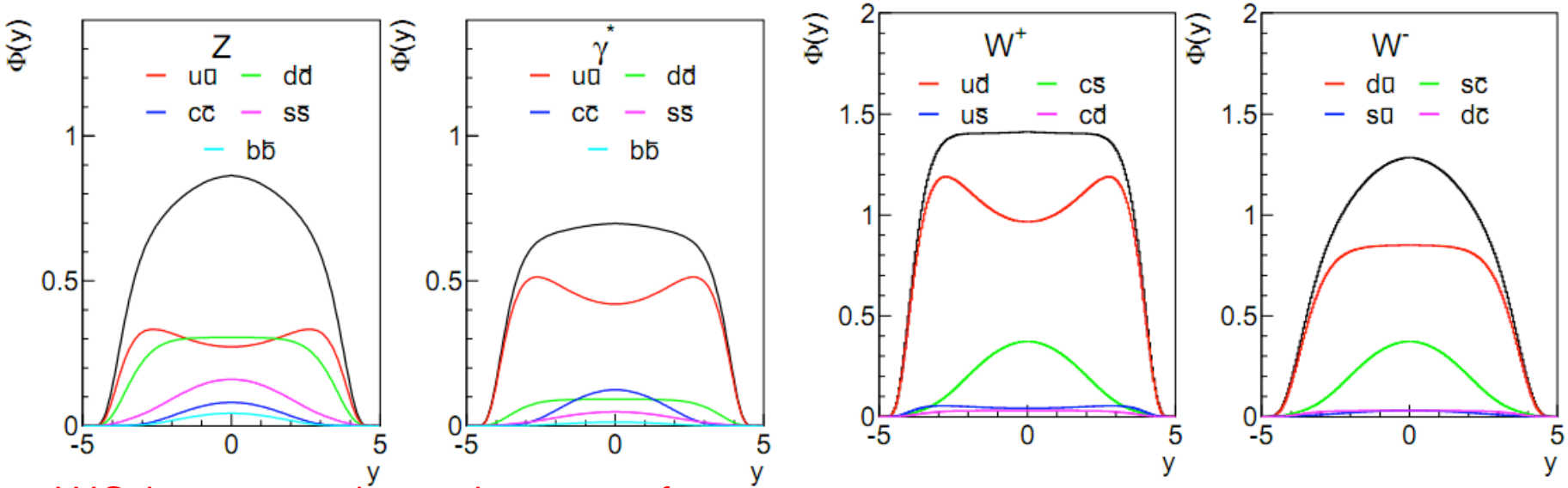
See A.Cooper-Sarkar's lecture

➔ No flavour decomposition of the sea distribution [S=2(ubar+dbar+sbar)]

- Additional constrain come from DY and jet data at the LHC
  - probe a bi-linear combination of quarks

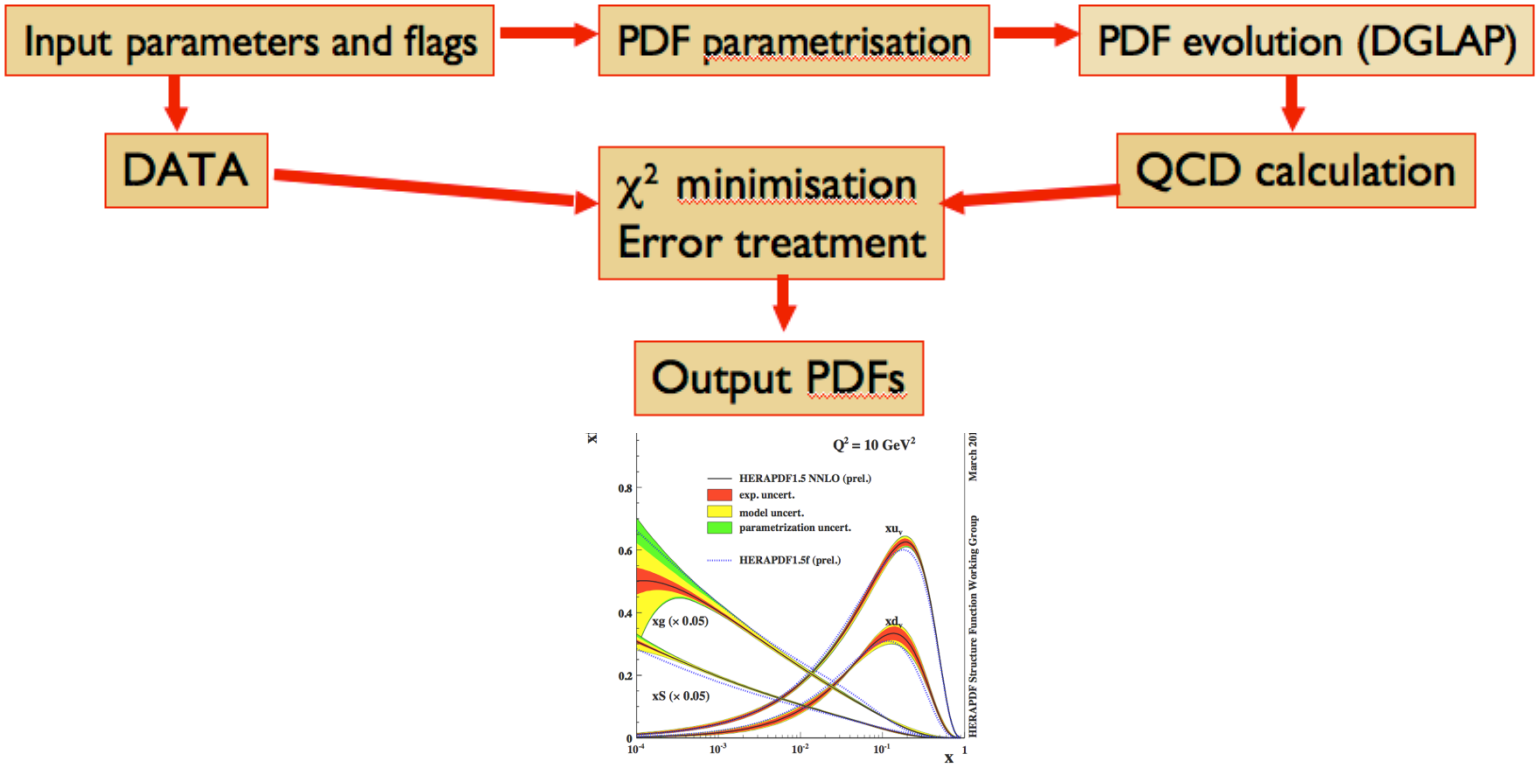
See J.Guimaraes lecture

$$Z \sim 0.29(u\bar{u}+c\bar{c}) + 0.37(d\bar{d}+s\bar{s}+b\bar{b})$$
$$\gamma^* \sim 0.44(u\bar{u}+c\bar{c}) + 0.11(d\bar{d}+s\bar{s}+b\bar{b})$$
$$W^+ \sim 0.95(u\bar{d}+c\bar{s}) + 0.05(u\bar{s}+c\bar{d})$$
$$W^- \sim 0.95(d\bar{u}+s\bar{c}) + 0.05(d\bar{c}+s\bar{u})$$



LHC data can provide complementary information, i.e. flavour decomposition of the quark sea [see Phys.Rev.Lett. 109 (2012) 012001]

Schematics of PDF extractions



- PDFs are extracted from QCD fits to cross section data:
  - Parametrise PDFs at a starting scale by smooth functions with sufficient parameters;
  - Evolve PDFs to other scales by the evolution equations (DGLAP);
  - Compute cross sections for DIS/DY (or other processes) at NLO (NNLO);
  - Calculate χ^2 measure of agreement between data and theory model;
  - Obtain the best estimate of the PDFs by varying the free parameters to minimize χ^2

PDF Fit Analysis Group

- Following Fit groups are active:

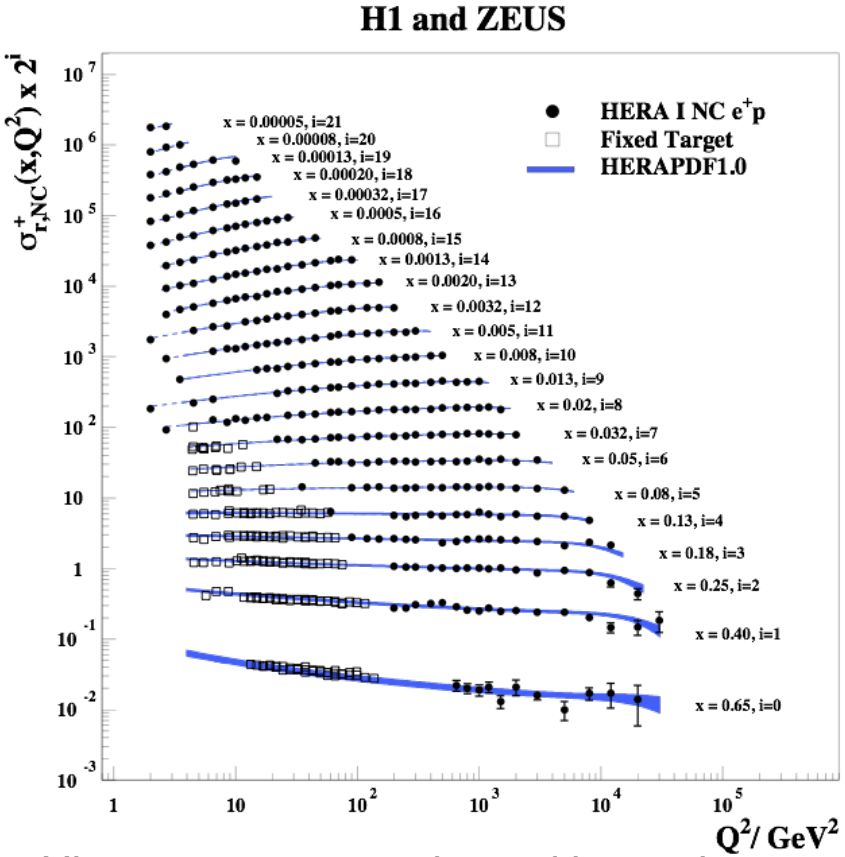
	MSTW08	CTEQ6.6/CT10	NNPDF2.1/2.3	HERAPDF1.0/1.5	ABKM09/ABM11	GJR08/JR09
Evolution	LO	LO	LO	—	—	—
Order	NLO	NLO	NLO	NLO	NLO	NLO
	NNLO	NNLO	NNLO	NNLO	NNLO	NNLO
HF Scheme	RT-GMVF	ACOT-GMVF	FONLL-GMVF	RT-GMVF (*)	BMSN-FFNS	FFNS
α_S NLO	0.120	0.118(f)	0.1191(b)	0.1176(f)	0.118	0.1135
α_S NNLO	0.1171	0.118(f)	0.1174(b)	0.1176(f)	0.1135	0.1124
HERA DIS	not up-to-date	+	+	+/prelim.	partial	+
Fixed target DIS	+	+	+	-	+	+
DY	+	+	+	-	+	+
Tevatron W,Z	some	some	some	-	some	some
Tevatron jets	some	+	+	-	some	some
LHC	-	-	W, Z+jets (NNPDF2.3)	-	-	-

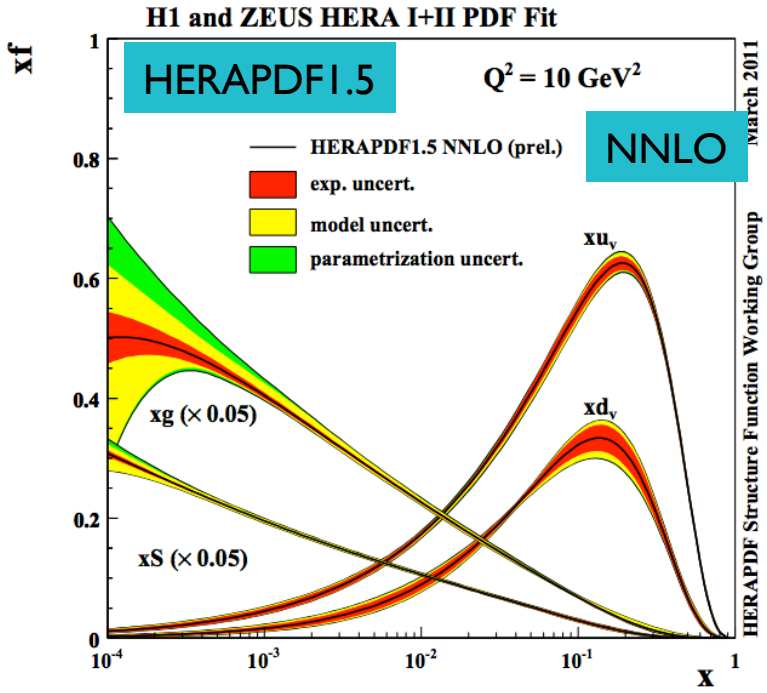
Courtesy S. Glazov, ICHEP2012

- Different data sets
- Different parametrisations
- Different arrangements of the perturbative series
- Different input values for alphas, charm masses
- Different treatment for heavy quark

Motivation for a QCD Fit Platform

- There is valuable expertise in the data combination and treatment of the experimental uncertainties as well as in the QCD fits:





- Ultimate precision is obtained by combining measurements
  - Improvement on Statistical precision
  - Improvement of Systematic precision

➔ QCD Fits within experiments proved to be a very useful tool to interpret data!

Motivation for a QCD Fit Platform (II)

- Data from HERA and LHC reach ~1% accuracy. The data are correlated point-to-point and across different processes due to common detector effects.

  → Treatment of the correlations very important [see Tuesday's lectures]

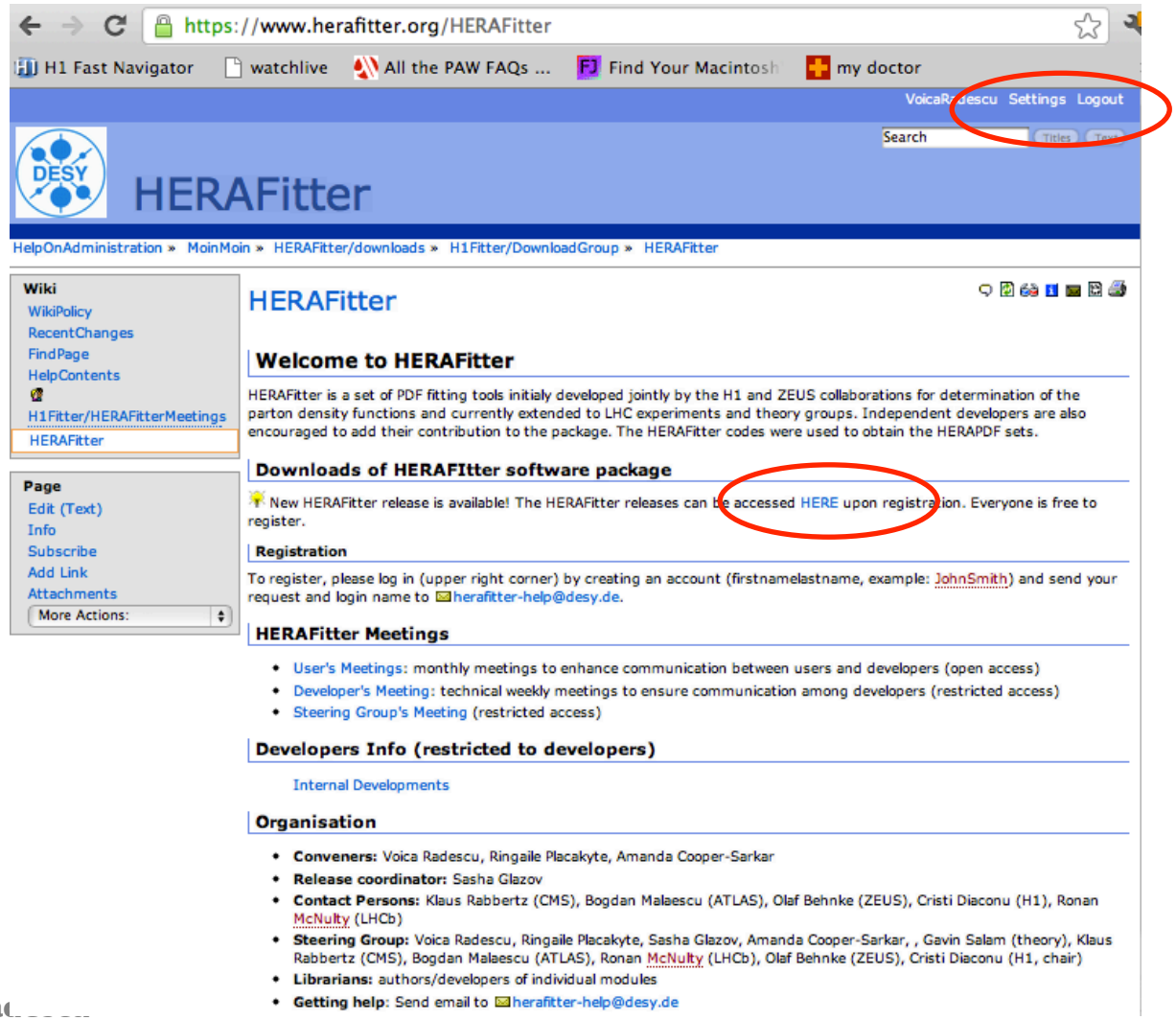
- Theoretical calculations for DIS and DY processes are available to NNLO accuracy in QCD (and NLO in EW). However, calculations, e.g. FEWZ programs for W,Z production are not fast, taking days to reach percent accuracy:

  - effect of PDFs in these calculations can be factorised, leading to fast computation tools: FastNLO, APPLGRID [see Tuesday's lectures]

  → Need a tool which combines the data and theory together.

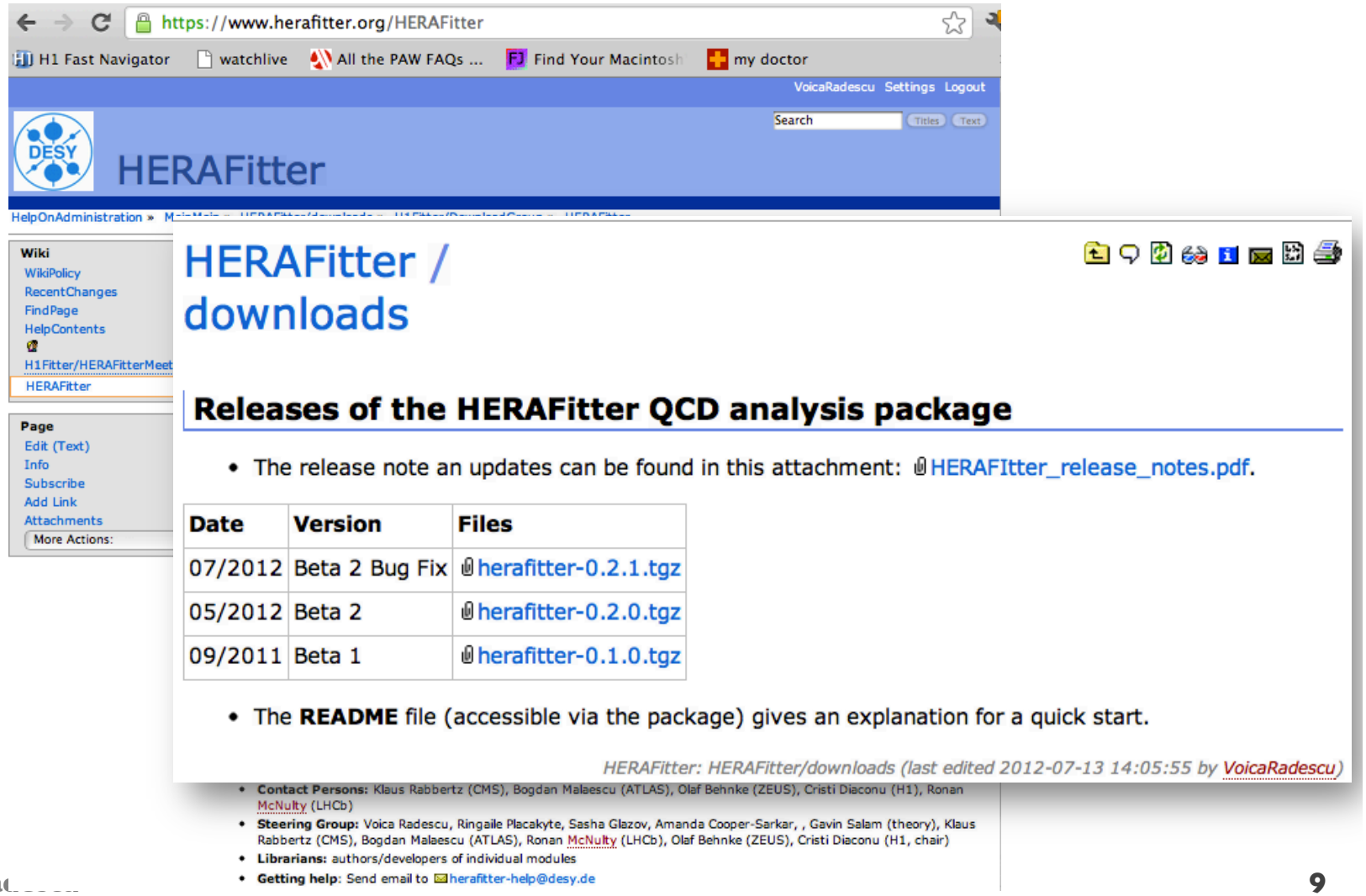
HERAFitter Package

- A ready platform to analyse new data and their impact.
- The beta releases can be accessed through the HEPFORGE site
  http://projects.hepforge.org/herafitter [it requires the QCDNUM package [M. Botje] for evolution]
  - Accessible to anyone for download via registration to feedback users

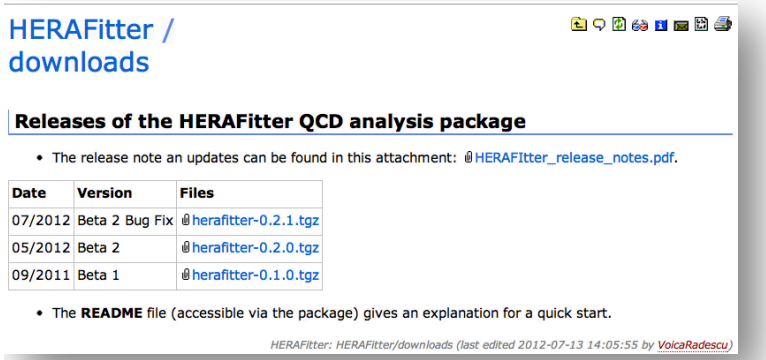


HERAFitter Package

- A ready platform to analyse new data and their impact.
- The beta releases can be accessed through the HEPFORGE site
  http://projects.hepforge.org/herafitter [it requires the QCDNUM package [M. Botje] for evolution]
  - Accessible to anyone for download via registration to feedback users



HERAFitter Package Installation



HERAFitter / downloads

Releases of the HERAFitter QCD analysis package

- The release note an updates can be found in this attachment: HERAFItter_release_notes.pdf.

Date	Version	Files
07/2012	Beta 2 Bug Fix	herafitter-0.2.1.tgz
05/2012	Beta 2	herafitter-0.2.0.tgz
09/2011	Beta 1	herafitter-0.1.0.tgz

- The **README** file (accessible via the package) gives an explanation for a quick start.

HERAFitter: HERAFitter/downloads (last edited 2012-07-13 14:05:55 by VoicaRadescu)

- The HERAFitter code uses automake tools to configure and build the package:

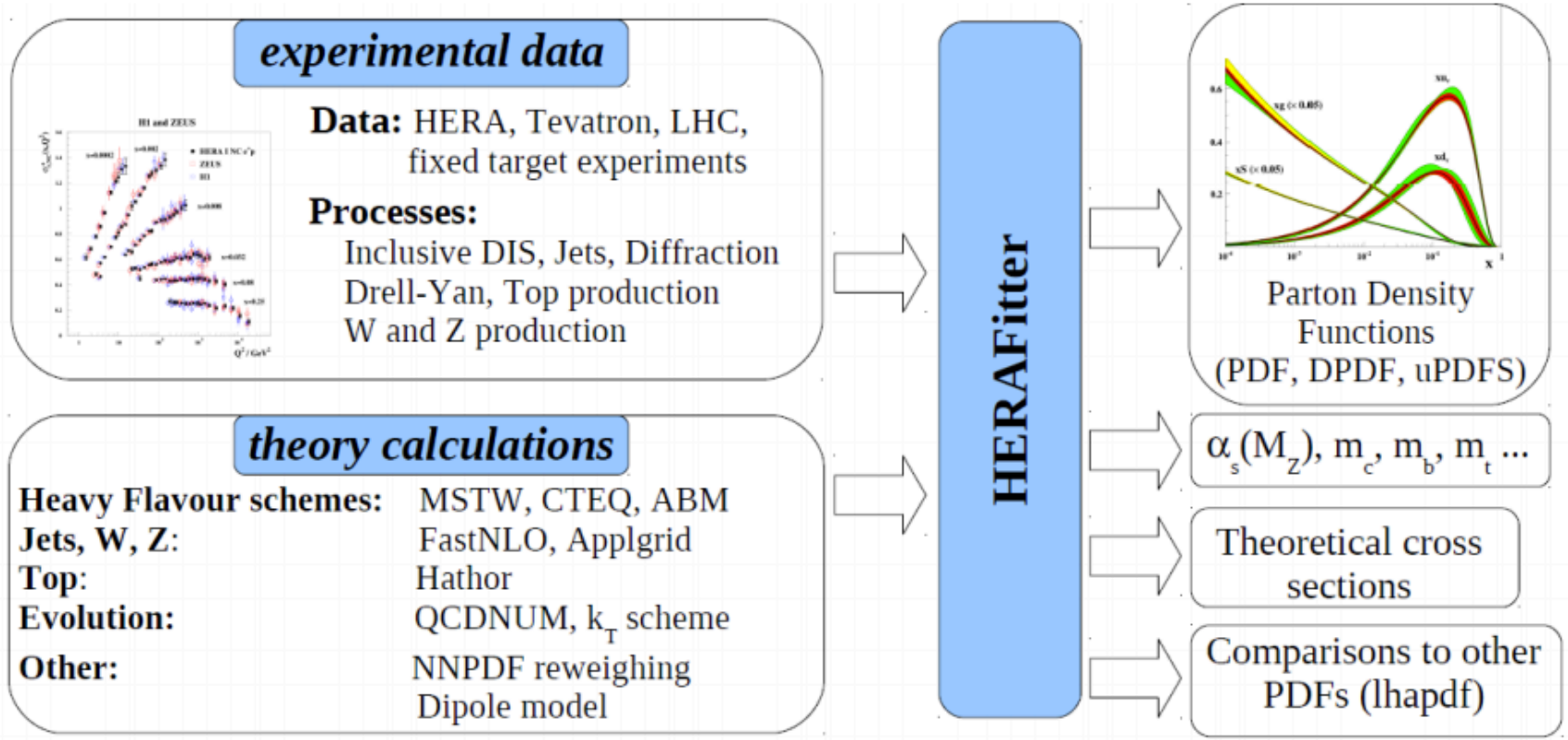
```
./configure
make
make install
```

  - Configure options: (./configure –help)

```
--enable-trapFPE        Stop of floating point errors (default=no)
--enable-checkBounds    add -fbounds-check flag for compilation (default=no)
--enable-nnpdfWeight    use NNPDF weighting (default=no)
--enable-lhapdf         use lhapdf (default=no)
--enable-applgrid       use applgrid for fast pdf convolutions (default=no)
--enable-hathor         use hathor for ttbar cross section predictions
                        (default=no)
```

  - Currently the pre-requested packages are high energy physics specific CERNLIB, QCDNUM
- Tested on the virtual machine (UBUNTU) ☺.

HERAFitter Functionality



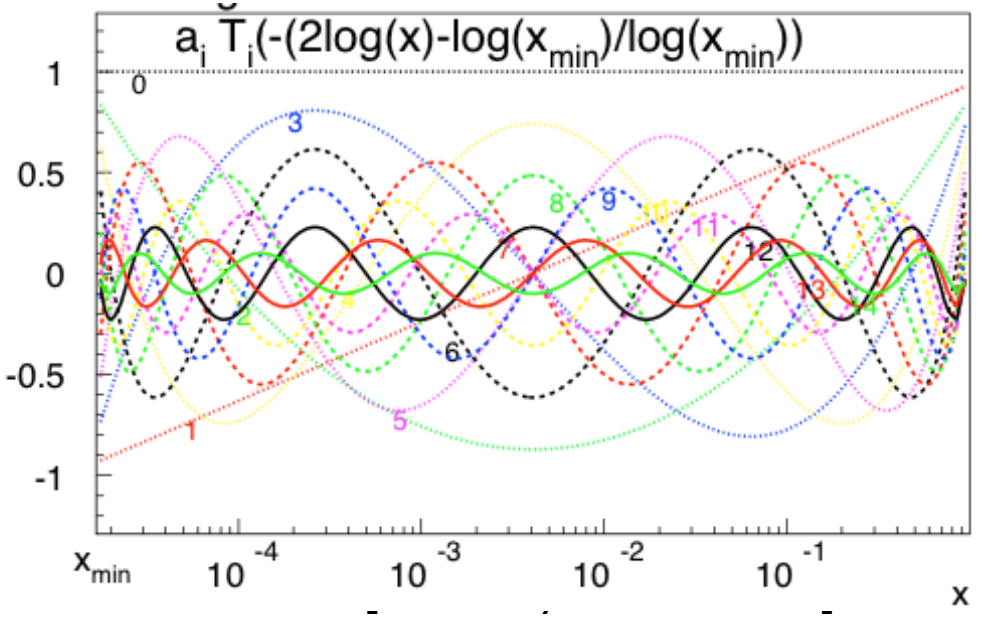
- Includes various methods for:
  - Error propagation: Hessian vs Monte Carlo: benchmark with NNPDF [PDF4LHC Report arXiv:1101:0536
  - Experimental Error treatment: Correlated, Uncorrelated, Offset [see Mandy's, Sasha's slides]
  - Parametrisation:
    - Standard Functional form (a la MSTW, CTEQ, ABM)
    - Chebyshev polynomials [PLB27193]
  - Chisquare definitions

HERAFitter Functionality

$$x f(x, Q_O^2) = A x^B (1-x)^C (1 + Dx + Ex^2)$$ STANDARD

$$x f(x) = \sum_{i=0}^{N-1} a_i T_i \left(\frac{-(2 \log x - \log x_{min})}{\log x_{min}} \right)$$ Chebyshev

$T_0(x) = 1$
$T_1(x) = x$
$T_2(x) = 2x^2 - 1$
$T_3(x) = 4x^3 - 3x$
$T_4(x) = 8x^4 - 8x^2 + 1$
$T_5(x) = 16x^5 - 20x^3 + 5x$
$T_6(x) = 32x^6 - 48x^4 + 18x^2 - 1$
$T_7(x) = 64x^7 - 112x^5 + 56x^3 - 7x$
$T_8(x) = 128x^8 - 256x^6 + 160x^4 - 32x^2 + 1$
$T_9(x) = 256x^9 - 576x^7 + 432x^5 - 120x^3 + 9x.$



- Parametrisation:
  - Standard Functional form (a la MSTW, CTEQ, ABM)
  - Chebyshev polynomials [PLB27193]
- Chisquare definitions

HERAFitter Functionality

Various Chisquares options:

- This definition corrects for possible biases introduced by the low statistics samples:
  - Uncertainties are scaled with expected number of events rather than observed
  - Logarithm term arises from the transition of likelihood to chisquare
- Correlated sources are taken into account (and a penalty term arises in the chisquare)

$$\chi^2 = \sum_i \frac{\left[\mu_i - m_i\left(1 - \sum_j \gamma_j^i b_j\right)\right]^2}{\delta_{i,\mathrm{unc}}^2 m_i^2 + \delta_{i,\mathrm{stat}}^2 \mu_i m_i \left(1 - \sum_j \gamma_j^i b_j\right)} + \sum_j b_j^2 + \sum_i \ln \frac{\delta_{i,\mathrm{unc}}^2 m_i^2 + \delta_{i,\mathrm{stat}}^2 \mu_i m_i}{\delta_{i,\mathrm{unc}}^2 \mu_i^2 + \delta_{i,\mathrm{stat}}^2 \mu_i^2}$$

Measured | Theoretical | Correl. sources | Correl. shifts | Log likelihood transition

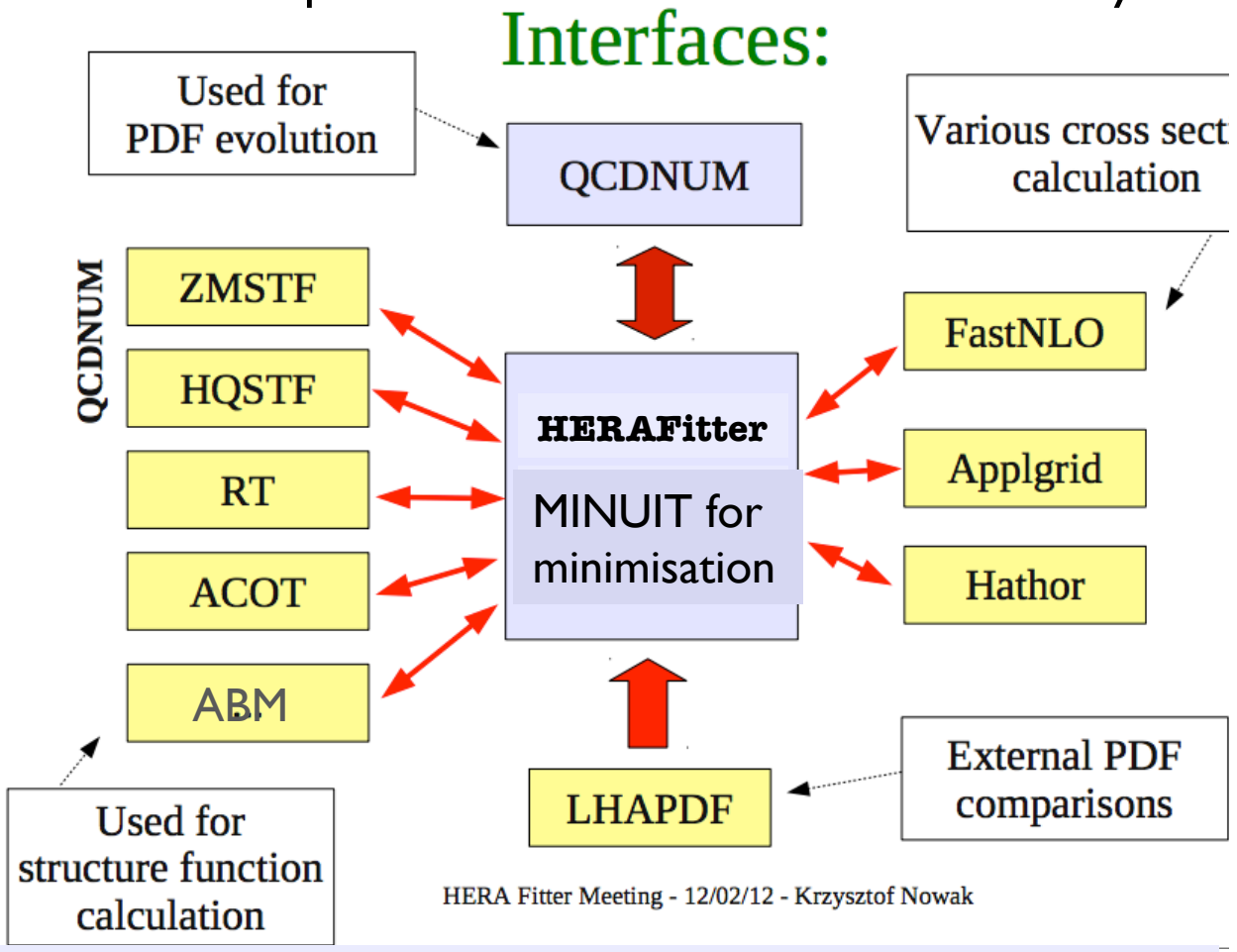
- Simplified options:
  - No scaling of the errors
  - Including scaling of the errors but no log term

(NOT IN BETA RELEASE: covariance matrix)

HERAFitter Structure

Modular Structure with reduced external dependencies:

- new developments can be added in a modular way



- Beta releases contain a set of tools for its use at the LHC experiments
  - It can produce out of the box HERAPDF1.0, (ATLAS strange determination)
  - It contains additional data from HERA, Fixed target, Tevatron, LHC

HERAFitter Physics Cases

- Determination of proton PDFs from HERA data
  - Inclusive NC and CC processes
    - involving low Q2 phenomenology (DIPOLE vs DGLAP models)
    - Mixed DGLAP-Dipole fits
  - DIS charm data (submitted to journal)
  - Inclusive DIS jets * (PDF + alphas)
  - Diffractive PDF fits
- Production of W, Z at LHC: additional lever arm to constrain PDFs
  - Inclusive Differential W, Z cross sections
  - Drell Yan at low and higher masses
  - Jet production * (PDF + alphas)
  - W+charm
- Studies concerning different treatment of correlations (Hessian vs MC vs Offset):
- Top production at LHC:
  - ttbar cross sections
  - Ratio of top/antitop cross sections
- Further developments:
  - QED evolution, photon PDFs
  - Benchmarking of theories
  - Fits using kt evolution
  - Nuclear PDFs

HERAFitter Package Specifics

- The software code is a mixture of C++ and Fortran codes. The core interfaces are provided in the Fortran part of the code.

- Central steering file to define input data, fitting parameters steering.txt
- Central steering file to define input PDF parameters minuit.in.txt
- Central steering file to define input ew constants ewparam.txt

- Package includes a ready to use data sets from various experiments:

bcdms hera lhc tevatron

- Inclusion of new data tables for existing processes should be possible without code recompilation. Data are provided as text files with a specified header and the main body, as a table.

[See K.Nowak's tutorial too]

Performance

HERAFitter Package is designed to be able to provide a fast feedback to the user:

- Use of Bayesian reweighing technique [see A. Guffanti's lecture]
- Full NLO (or NNLO for DIS) calculations vs FAST options based on kfactors

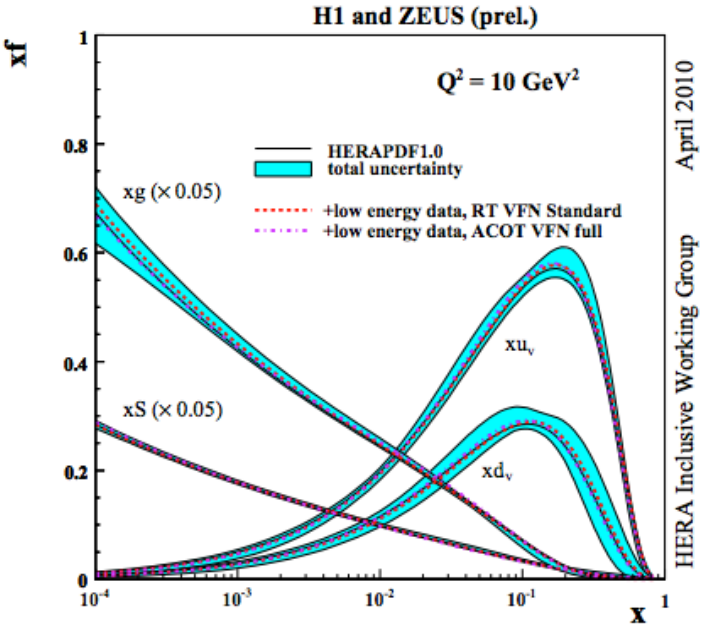
  - RT scheme (NLO or NNLO): ~3h for ~1500 iterations
  - RT FAST:
    - Theory (NLO) = {NLO(RT)/NLO(ZMVFNS)}_kfactor x NLO(ZMVNS) ~10min for ~2500 iterations
    - Theory (NNLO) = {NNLO(RT)/NNLO(ZMVFNS)}_kfactor x NNLO(ZMVNS) ~10min for ~2500 iterations

    basically it takes the speed of QCDNUM (few miliseconds per iterations)

  - ACOT Full:
    - Theory = {NLO(ACOT Full)/LOmassive(ACOT Full)}_kfactor x LOmassive(ACOT Full)~30min

(Exercise: test the difference between RT and ACOT scheme)

HERAFitter Outputs

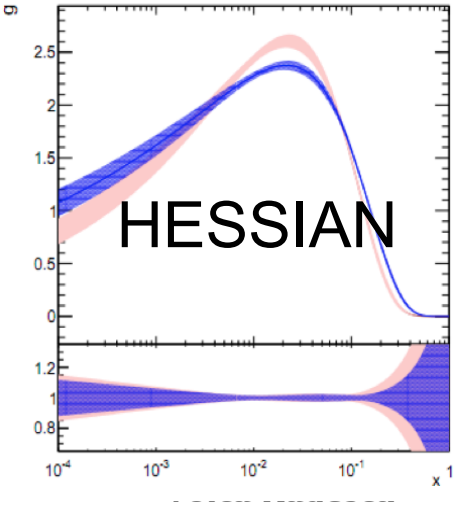


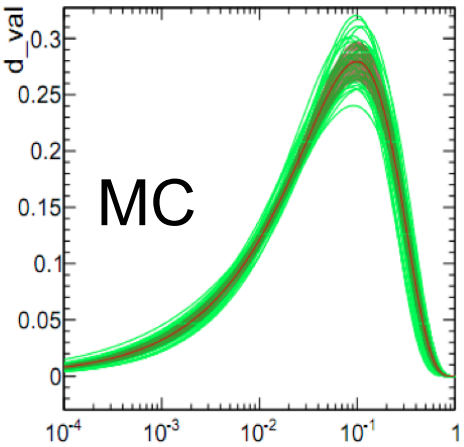
Output contains basic text (and graphic) information stored in output directory:

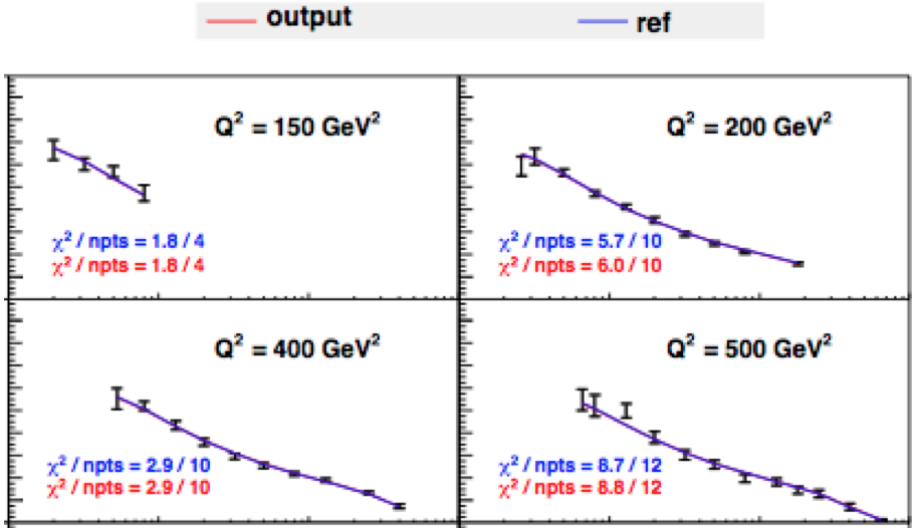
- Quality of the fit (chisquares, pulls)
- Resulting PDFs:
  - text and HERALHGRID LHAPDF format grids ready to plug into the MC generators
- Hessian vs MC replicas error estimation

[Phys.Rev.D65:014011,2001 and arXiv:1101:0536

[See A. Glazov's lecture]



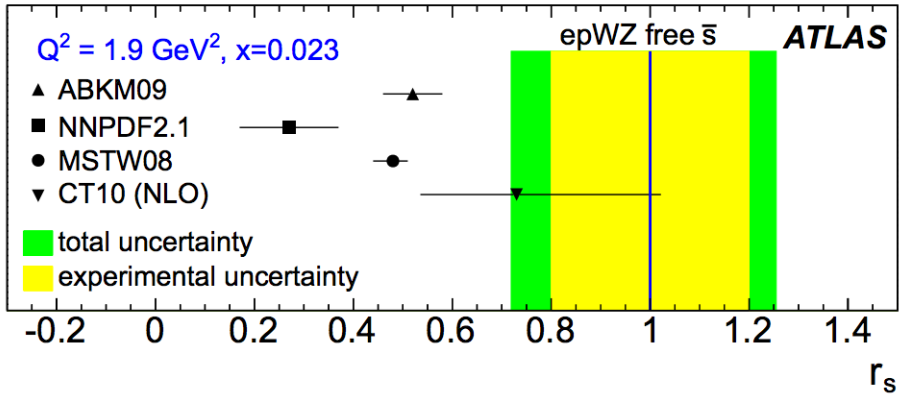




Results using HERAFitter framework

- The differential $W^{\pm}$, Z cross section data of ATLAS (2010, 35/pb) were jointly analysed with $e^{\pm}p$ cross sections from HERA using the HERAFitter framework

Phys.Rev.Lett. 109 (2012) 012001

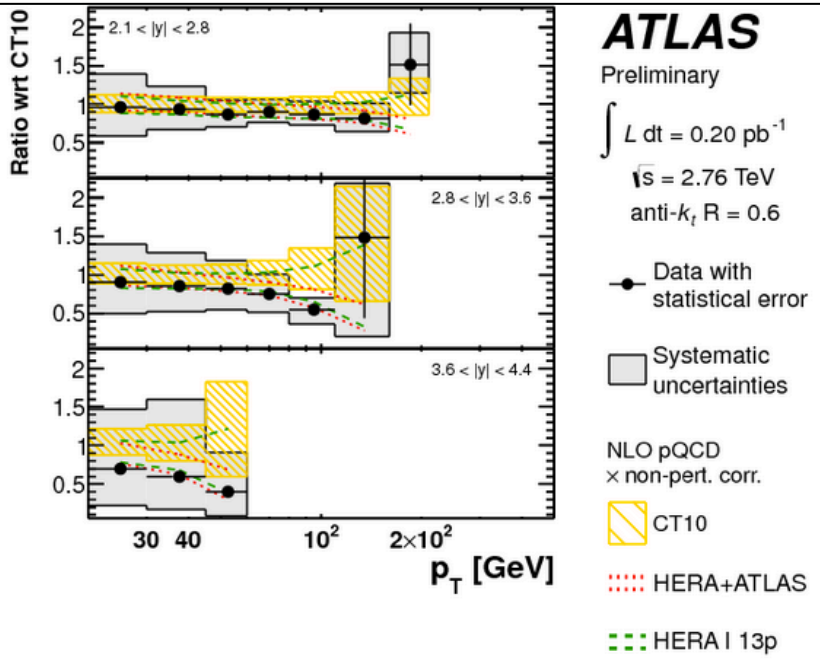


At LHC, ratio of W/Z cross sections together with yZ shape provide a constraint on **s**-quark density.

$$r_s = 1.00 \pm 0.20_{\text{exp}} \pm 0.07_{\text{mod}}{}^{+0.10}_{-0.15}{}_{\text{par}}{}^{+0.06}_{-0.07}\alpha_S \pm 0.08_{\text{th.}}$$

- Measurement of the ATLAS inclusive jet cross section in pp collisions at sqrt(s) = 2.76 TeV and comparison to the inclusive jet cross section at sqrt(s) = 7 TeV taking fully into account the correlations of the measurements [ATLAS-CONF-2012-128]

See J.Guimaraes lecture



HERAFitter organisation

- Timescale:
  - September 2011 — First Beta Release
  - Septemeber and October — Package presented to the LHC community (ATLAS and CMS)
  - October 2011 — First HERAFitter User's Meeting
  - November 2011 — First presentation of the HERAFitter at a workshop
  - February 2012 — HERAFitter Workshop in Marseille
  - May 2012 — Second Beta Release
  - October 2012 — PDF School at DESY
  - Winter 2012 — Next release
- Package is supported by
  - a group of developers originally from H1 and ZEUS collaborations and extended to LHC experiments and theory groups (Independent developers are also encouraged to add their contribution to the package)
  - a steering group committee with contact persons to HERA, LHC and theory
    - C. Diaconu (H1), O. Behnke (ZEUS), B. Malaescu (ATLAS), K. Rabbertz (CMS), R. McNaulty (LHCb), G. Salam (theory), V. Radescu, R. Placakyte, A.Cooper-Sarkar, A. Glazov
- HERAFitter User's interaction
  - Weekly developer's meeting https://herafitter.org/HERAFitter/HERAFitter/HERAFitterInternal/FitForumMeetings
  - Monthly users's meeting (https://herafitter.org/HERAFitter/HERAFitter/HERAFitterMeetings)
- NEXT User's meeting 29.10.2012

New in the next release

- Bayesian reweighting technique also for PDF predictions based on central values with errors determined using the Hessian Eigenvector Method [proposed by G. Watt at R. Thorne implemented by K. Lohwasser, A. Guffanti].

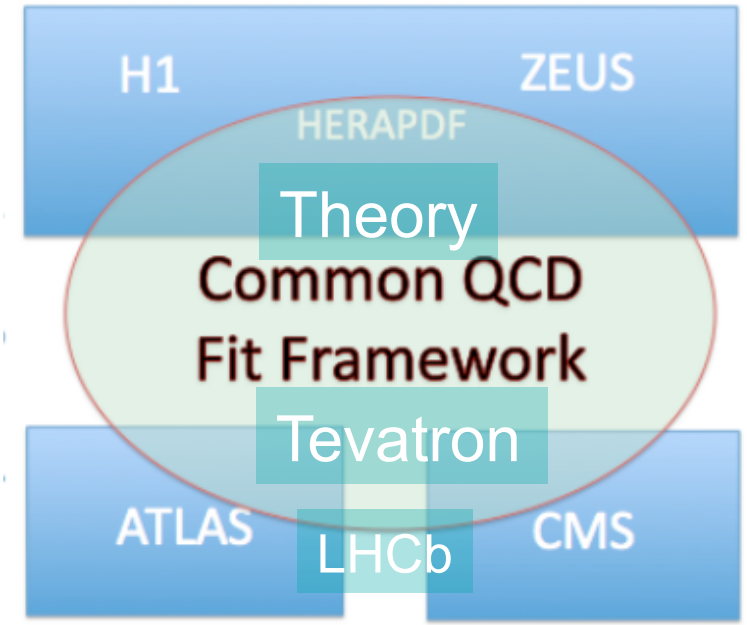
- Integrating uPDF in the HERAFitter (H. Jung) using kt-factorisation
- New dipole model implementation BGK (H. Kowalski, A. Luszczak)
- Redundant parametrisation and regularisation technique (S. Glazov)
- Covariance matrix representation for chisquare

HERAverager: A new package containing (rewritten) code to combine measurements

Summary

- Successful beta-releases of the HERAFitter package so far
  - Multi-platform usage of the package: ATLAS, CMS, theory groups
- Further development of the package towards the stable release:
  - Modular addition of the heavy flavour schemes with the support of Theory groups

- HERAFitter infrastructure has the potential to increase the scientific output of the LHC data and to provide a flexible environment for theory benchmarking



License and References

- LICENSE: under GNU GPL v3
- REFERENCES: Citation depending on the usage

```
If you use the HERAFITTER package in a scientific publication, please consider adding the
following references. The main citations list contains the papers which should be cited
for any use of the HERAFITTER program. In addition, some citations are required depending
on the modules, data and theory tables used in the program.

==================
  Main citations
==================

HERAFitter
---------
1) "Combined Measurement and QCD Analysis of the Inclusive e+- p Scattering Cross Sections at HERA."
By H1 and ZEUS Collaboration (F.D. Aaron et al.). DESY-09-158, Oct 2009. 61pp.
Published in JHEP 1001:109,2010.
e-Print: arXiv:0911.0884 [hep-ex]

2) "A Precision Measurement of the Inclusive ep Scattering Cross Section at HERA."
By H1 Collaboration (F.D. Aaron et al.). DESY-09-005, 2009. 35pp.
Published in Eur.Phys.J.C64:561-587,2009.
e-Print: arXiv:0904.3513 [hep-ex]

QCDNUM  ( evolution code )
-------
"Fast QCD Evolution and Convolution", M. Botje,
NIKHEF-10-002, May 2010. 74pp.
Published in Comput.Phys.Commun.182:490-532,2011.
e-Print: arXiv:1005.1481 [hep-ph]
```

HERAFitter Functionality (I)

- Can perform fits to DIS, DIS diffractive, jets, DY data:
  - FastNLO and APPLGRID interfaces implemented
- QCD evolution based on QCDNUM package:
  - NLO, NNLO DGLAP equations
- Access to various heavy flavour schemes:
  - RT standard and optimal as in MSTW (NLO, NNLO)
  - ACOT as in CTEQ (LO, NLO)
  - FFNS and BMSN as in ABM (NLO, NNLO)
  - Developments in the top area: ttbar cross section using HATHOR package
- Includes various methods for:
  - Error propagation:
    - Hessian vs Monte Carlo: benchmark with NNPDF [PDF4LHC Report arXiv:1101:0536
  - Experimental Error treatment: Correlated, Uncorrelated, Offset
  - Parametrisation:
    - Standard Functional form (a la MSTW, CTEQ, ABM)
    - Chebyshev polynomials [PLB27193]
  - Chisquare definitions
- Possibility to link to LHAPDF and draw/compare various predictions
- Access to the NNPDF reweighting tool
- kt-evolution for the unintegrated PDFs
- Various DIPOLE Models