HERAFitter Open source QCD Fit Framework

Helmholtz International Summer School - HISS
Dubna International Advanced School of Theoretical Physics - DIAS TH

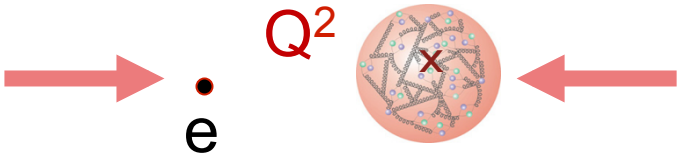
INTERNATIONAL SCHOOL-WORKSHOP
"CALCULATIONS FOR MODERN AND FUTURE COLLIDERS"

July 23 - August 2, 2012, Dubna, Russia

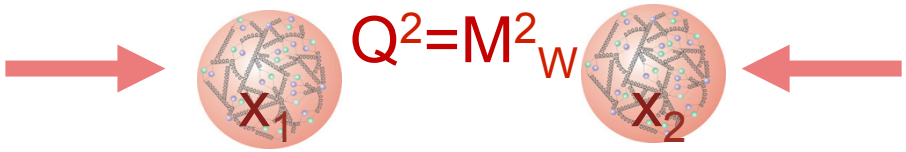
Proton Structure

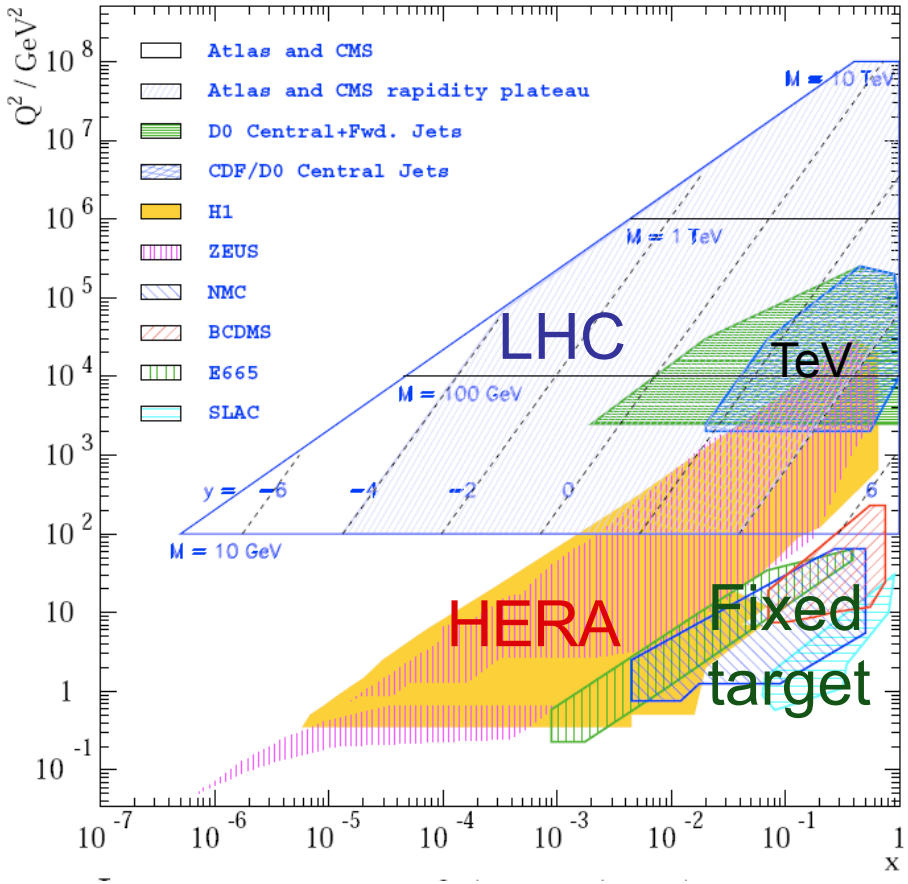
- Factorization theorem states that cross section can be calculated using universal partons x short distance calculable partonic reaction.

- 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

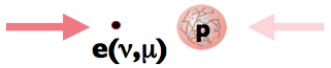




Persistent experimental effort over the last 40 years both by fixed-target and collider experiments around the world supported by the theoretical developments

PDF Constraints from LHC

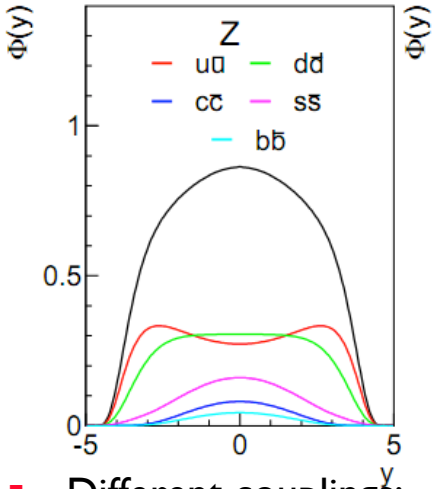
- Main information on proton structure comes from DIS data at HERA:
  - probes linear combination of quarks:
    - CC: provides constraints on valence quarks
    - NC: $F_2 \sim 0.44x(u+\bar{u}+c+\bar{c}) + 0.11x(d+\bar{d}+s+\bar{s}+b+\bar{b})$

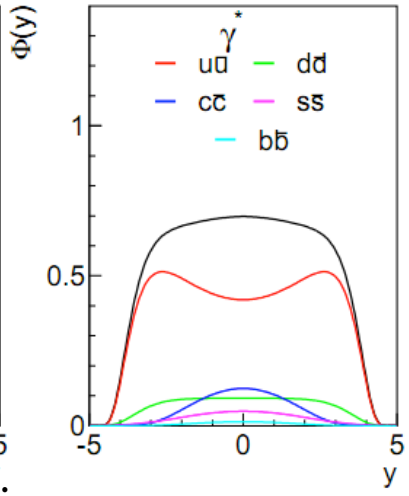


➔ No flavour decomposition of the sea distribution [S=2(ubar+dbar+sbar)]

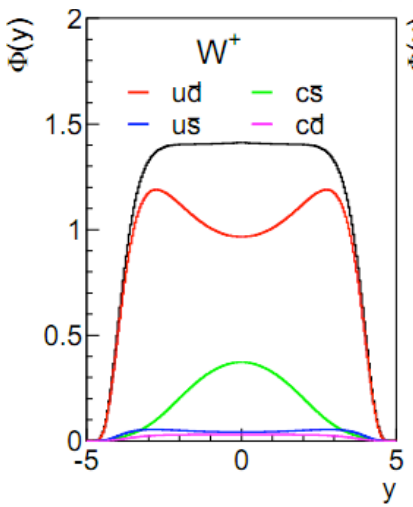
- Additional constrain come from DY and jet data at the LHC
  - probe a bi-linear combination of quarks

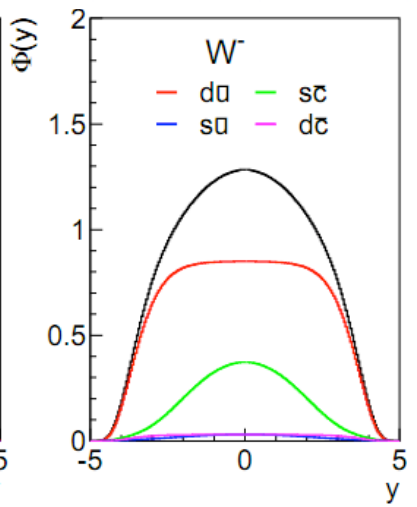
$$Z \sim 0.29(u\bar{u}+c\bar{c}) + 0.37(d\bar{d}+s\bar{s}+b\bar{b})$$
$$\gamma^* \sim 0.44(u\bar{u}+c\bar{c}) + 0.11(d\bar{d}+s\bar{s}+b\bar{b})$$





$$W^+ \sim 0.95(u\bar{d}+c\bar{s}) + 0.05(u\bar{s}+c\bar{d})$$
$$W^- \sim 0.95(d\bar{u}+s\bar{c}) + 0.05(d\bar{c}+s\bar{u})$$

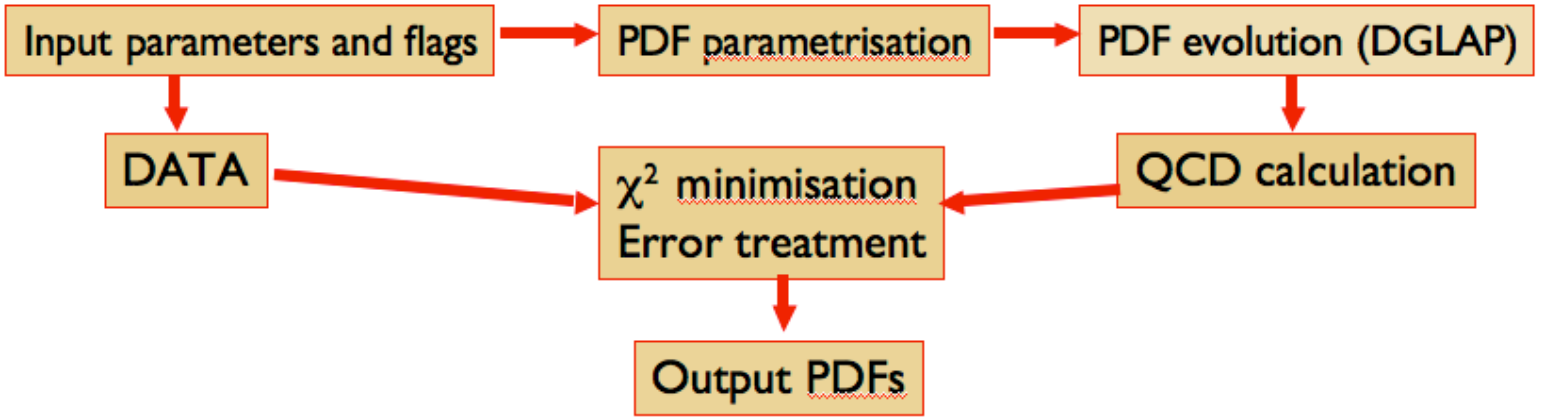




- Different couplings:
  - Z production more sensitive to **d** vs **u** quarks and more sensitive to **s** than W production

➔ LHC data can provide complementary information:
flavour decomposition of the quark sea

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

- Various data sets have constraining powers on PDFs:
  - Fixed Target experiments - high x, HERA (ep collider) - low x
  - Tevatron (ppbar collider), LHC (pp collider)
- 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

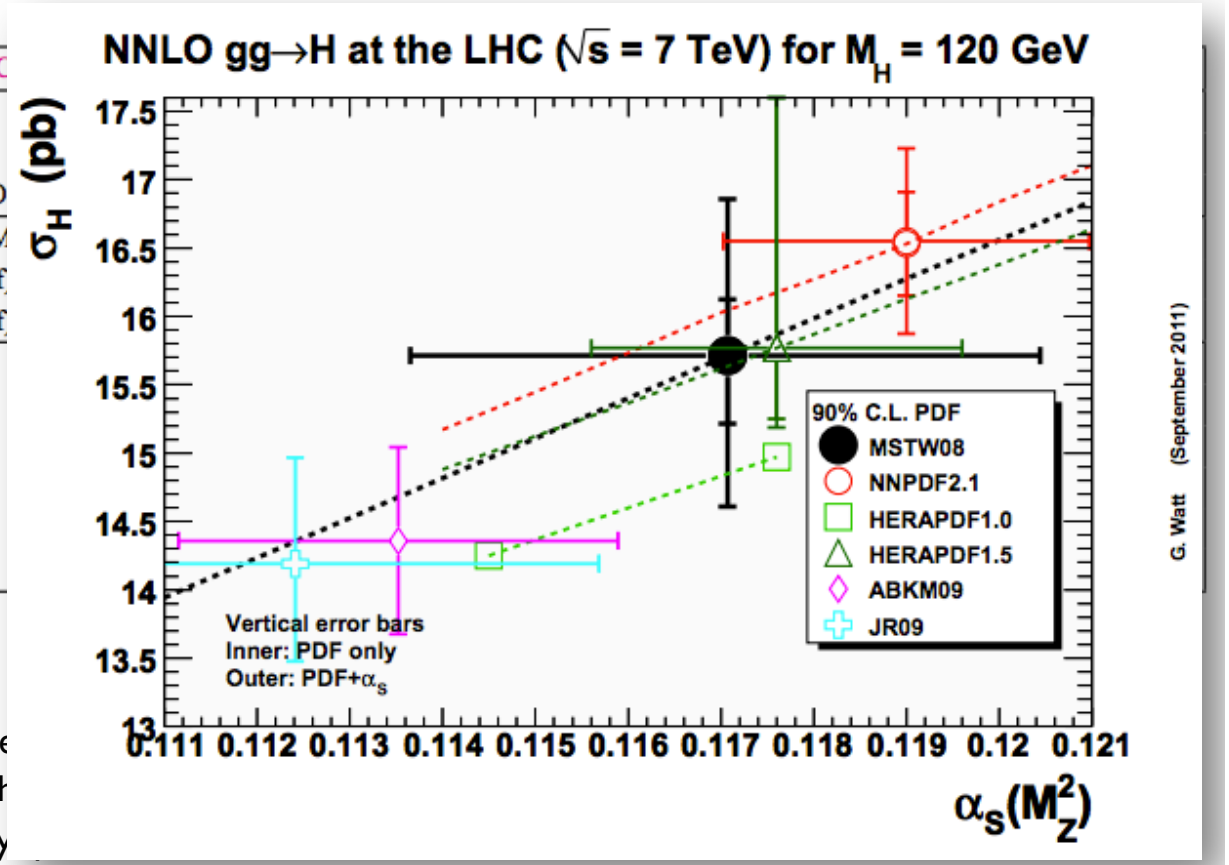
- Different data sets
- Different parametrisations
- Different arrangements of the perturbative series
- Different input values for alphas, charm masses
- Different treatment for heavy quark

There is a need for an open platform to benchmark various theory predictions under the same conditions.

PDF Fit Analysis Group

- Various data sets have constraining powers on PDFs:
  - Fixed Target experiments - high x, HERA (ep collider) - low x
  - Tevatron (ppbar collider), LHC (pp collider)
- Following Fit groups are active:

	MSTW08	CTEQ6.6/C
Evolution	LO	LO
Order	NLO	NLO
	NNLO	NNLO
HF Scheme	RT-GMVF	ACOT-GM
α_S NLO	0.120	0.118(f
α_S NNLO	0.1171	0.118(f
HERA DIS	not up-to-date	+
Fixed target DIS	+	+
DY	+	+
Tevatron W,Z	some	some
Tevatron jets	some	+
LHC	-	-

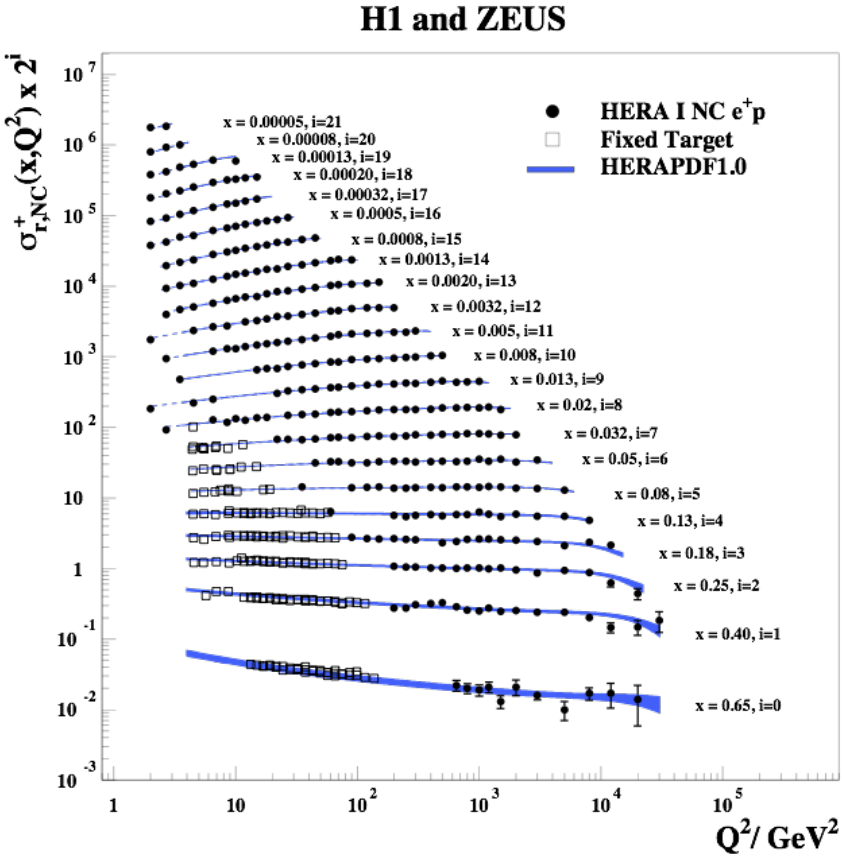


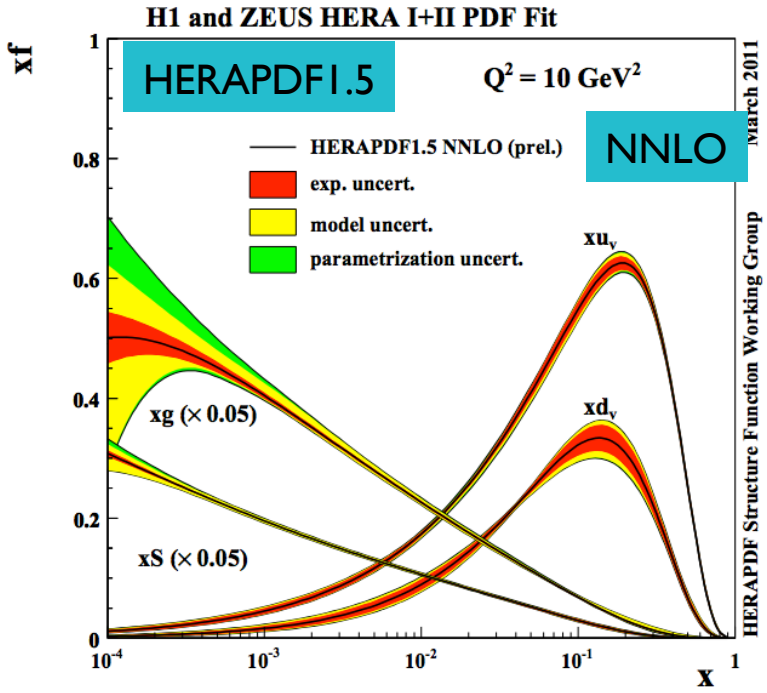
- Different data sets
- Deifferent parametrisations
- Different arrangements of the
- Different input values for alph
- Different treatment for heavy

➔ There is a need for a tool to benchmark PDF analyses

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.

- Theoretical calculations for DIS and DY processes are available to NNLO accuracy in QCD (and NLO in EW). However, calculations, e.g. FEWZ and SANC [Dubna Group] 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 talk of Pavel)

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

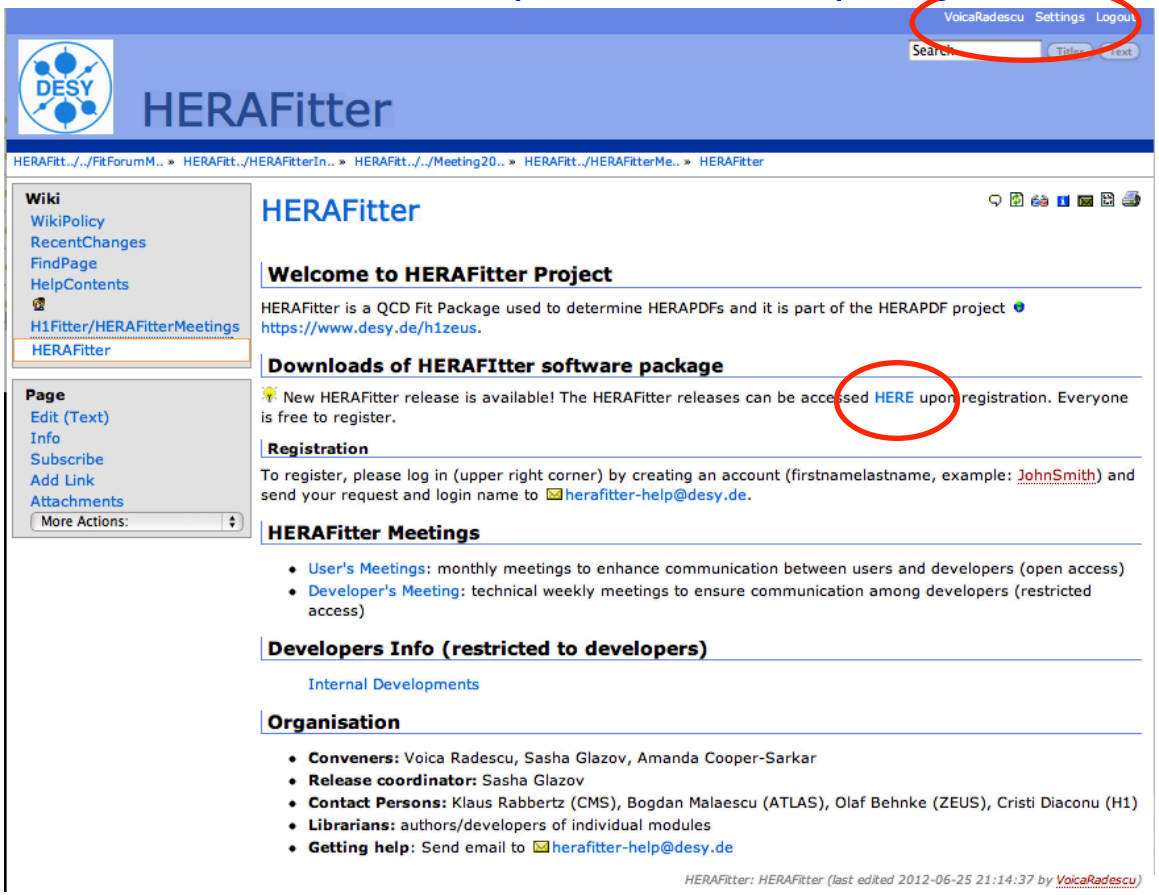
HERAFitter Package

- 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
  - References should follow citations provided with the package



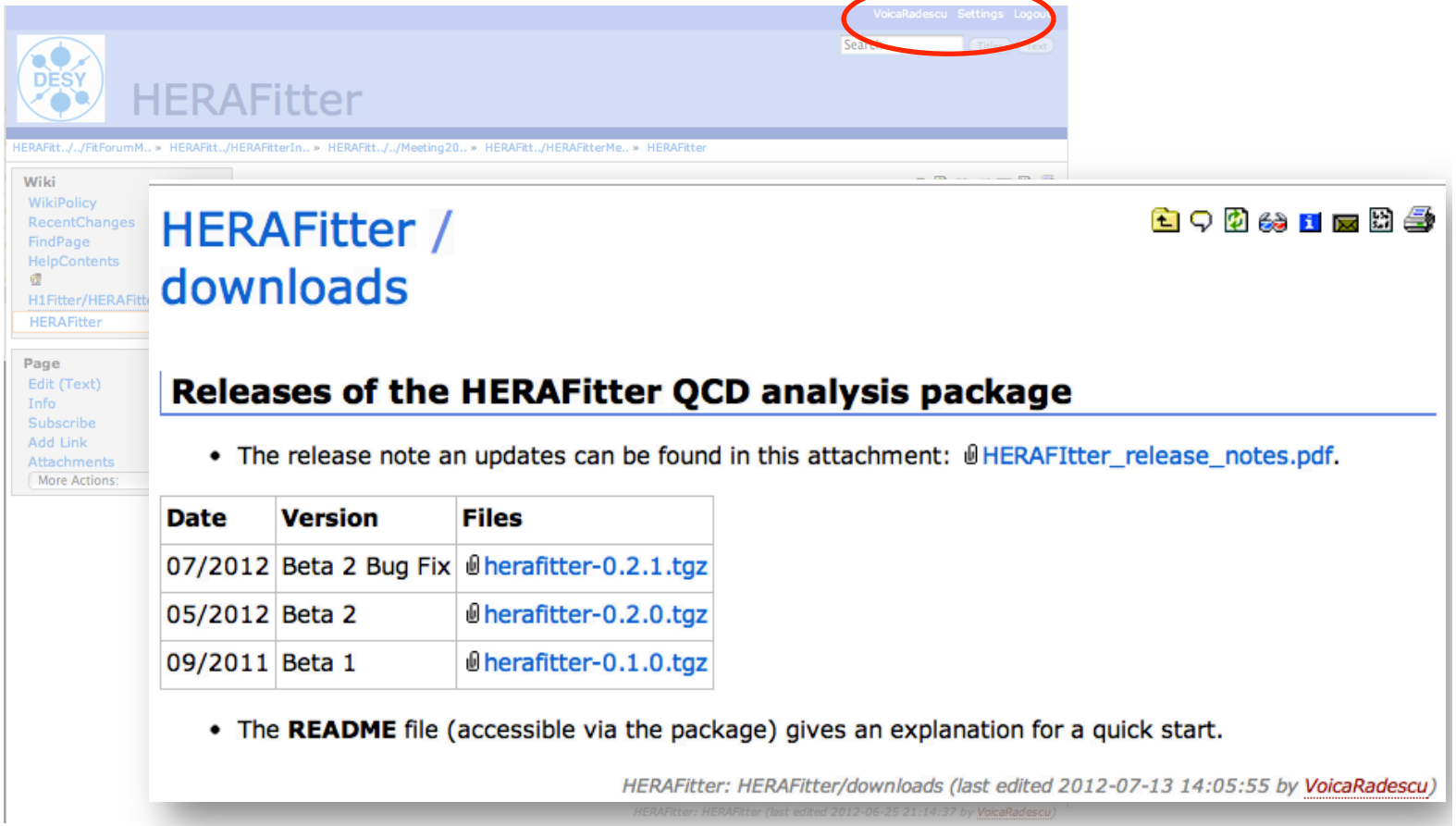
HERAFitter Package

- A ready platform to analyse new data and their impact.
- The beta releases can be accessed through the HEPFORGE site:

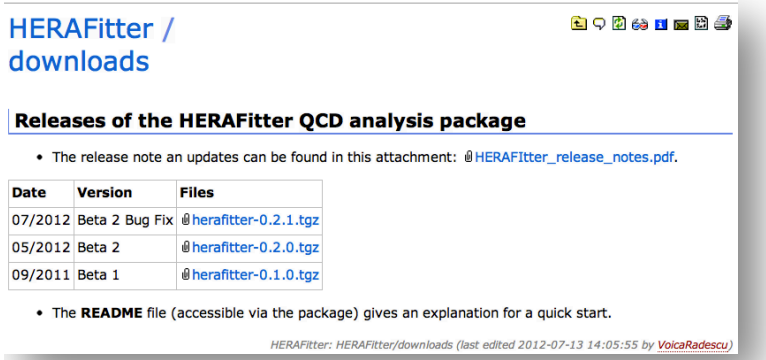
http://projects.hepforge.org/herafitter

[it requires the QCDNUM package [M. Botje] for evolution]

- Accessible to anyone for download via registration to feedback users
- References should follow citations provided with the package



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)

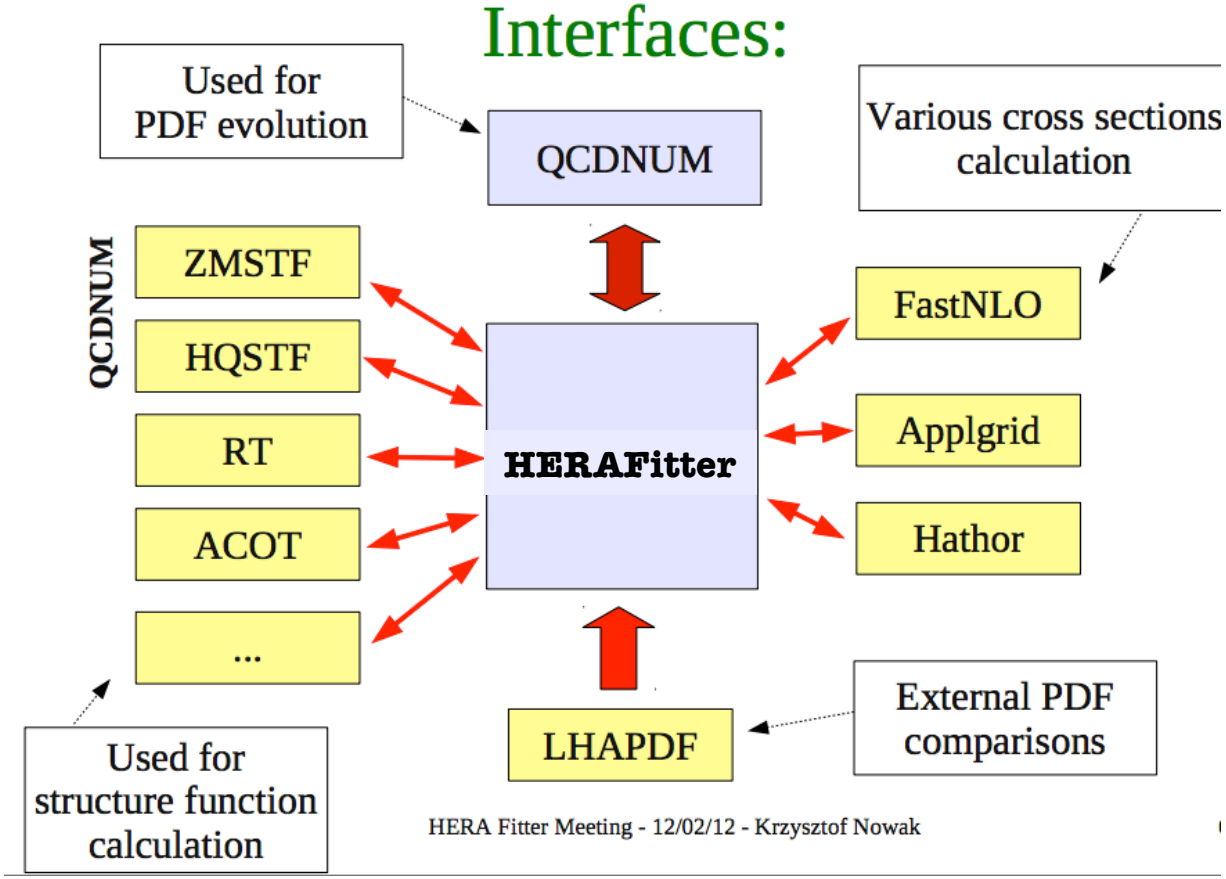
```
--enable-trapFPE        Stop of floating point errors (default=no)
--enable-checkBounds    add -fbounds-check flag for compilation (default=no)
--enable-nnpdfWeight    use NNPDF weighting (default=no)
--enable-lhapdf         use lhapdf (default=no)
--enable-applgrid       use applgrid for fast pdf 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
  - Works on different platforms from MacOS, Ubuntu

HERAFitter Structure

Modular Structure with reduced external dependencies:

- new developments can be added in a modular way



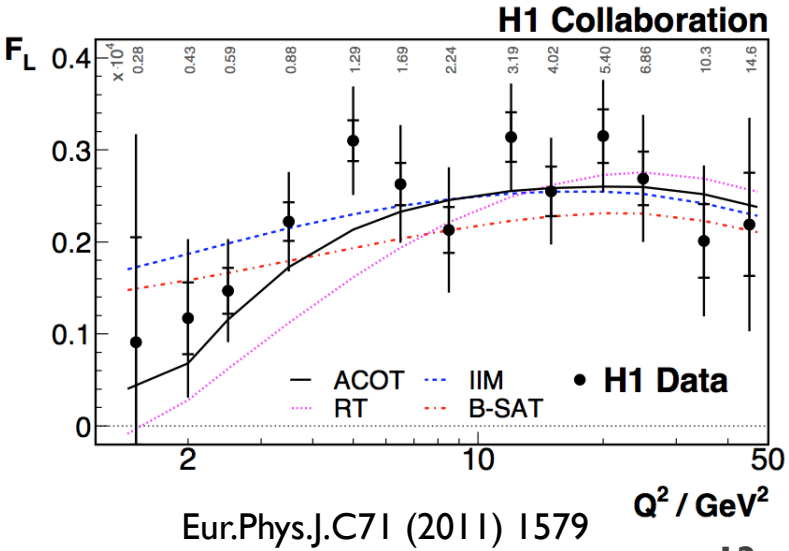
- Beta releases contain a set of tools for its use at the LHC experiments
  - It can produce out of the box HERAPDF1.0
  - It contains additional data from HERA, Fixed target, Tevatron, LHC

Functionality

- Interfaces to DIS, DY (*Andrey Sapronov), Applgrids and FASTNLO modules
- Heavy flavour schemes:
  - RT standard and optimal as in MSTW
  - ACOT as in CTEQ
  - FFNS and BMSN as in ABM
  - Developments in the top area: ttbar cross section using HATHOR
- Possibility to link to LHAPDF and draw/compare various predictions
- Access to the NNPDF reweighting tool
- Diffractive fits

- Additions to HERAFitter package: HERAaverager
  - Used for combining the measurements

- Others developments for cross model benchmarking:
  - DIPOLE Models
  - Various evolutions
  - Kt-evolution for unintegrated PDFs



Eur.Phys.J.C71 (2011) 1579

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

```
bcdms  hera  lhc  tevatron
```

```
bcdms/BCDMS_F2p.100gev.dat
bcdms/BCDMS_F2p.120gev.dat
bcdms/BCDMS_F2p.200gev.dat
bcdms/BCDMS_F2p.280gev.dat

bcdms/oldfiles:
hepdata.f2pbcdms.dat
```

```
hera/H1ZEUS_CC_e+p_HERA1.0.dat
hera/H1ZEUS_CC_e-p_HERA1.0.dat
hera/H1ZEUS_NC_e+p_HERA1.0.dat
hera/H1ZEUS_NC_e-p_HERA1.0.dat
hera/H1_LowEp_460_575.dat
```

```
hera/H1_NormInclJets_HighQ2_99-07.dat
hera/ZEUS_LPS_98-00.dat
hera/ZEUS_LRG_98-00.dat
```

```
lhc/atlas:
Jets2010  WZ2010  oldfiles

lhc/cms:
CMS-TOP-11-024_prelim.dat  CMS_Z_boson_Rapidity.dat
```

```
tevatron/CDF-TOP-CONF-NOTE-9913_prelim.dat
tevatron/CDF_JETS2008.dat
tevatron/D0_JETS.dat
tevatron/D0_Z_Boson_Rapidity.dat
```

- 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.

Data format

```
&Data
   Name  = 'NC cross section'

 ! Data table definitions:
   NData   =  145   ! 145 rows, corresponding to 145 data points
   NColumn =  120   ! 3 bins, sigma and 116 errors

 ! Layout of the data table columns: 3 bins, cross-section and 116 errors
 ! The following types are predefined: Bin, Sigma, Error and Dummy (case sensitive!)
  ColumnType = 3*'Bin','Sigma',116*'Error'

!   To treat error uncorrelately, then: first  is uncor, then the sys_i(i=1,114) -> 115 sources
            !     Bins       x-sec                            Errors
  ColumnName = 'x','Q2','y','reduced x-section','stat','uncor',110*'uncor',4*'ignore'

 ! options for bins: depends on Reaction type, for NC e+-p these are 'x', 'Q2' and 'y'
 ! options for the errors: 'stat', 'uncor', 'ignore' or different names for correlated uncertainties


 ! Extra information for the x-section calculation:
   NInfo    = 3
   DataInfo =  318.,        -1.,        0.
   CInfo    = 'sqrt(S)','e charge','e polarity'

   IndexDataset = 61

   Reaction = 'NC e+-p'


!   To take into account the correlations then set SystScales to 1. and uncomment below:
!   ColumnName = 'x', 'Q2', 'y', 'Sigma',
!                'stat', 'uncor',  'h1', 'h2', 'h3', 'h4', 'h5', 'h6', 'h7', 'h8', 'h9', 'h10',
!                'h11', 'h12', 'h13', 'h14', 'h15', 'h16', 'h17', 'h18', 'h19', 'h20',
!                'h21', 'h22', 'h23', 'h24', 'h25', 'h26', 'h27', 'h28', 'h29', 'h30',
!                'h31', 'h32', 'h33', 'h34', 'h35', 'h36', 'h37', 'h38', 'h39', 'h40',
!                'h41', 'h42', 'h43', 'h44', 'h45', 'h46', 'h47', 'h48', 'h49', 'h50',
!                'h51', 'h52', 'h53', 'h54', 'h55', 'h56', 'h57', 'h58', 'h59', 'h60',
!                'h61', 'h62', 'h63', 'h64', 'h65', 'h66', 'h67', 'h68', 'h69', 'h70',
!                'h71', 'h72', 'h73', 'h74', 'h75', 'h76', 'h77', 'h78', 'h79', 'h80',
!                'h81', 'h82', 'h83', 'h84', 'h85', 'h86', 'h87', 'h88', 'h89', 'h90',
!                'h91', 'h92', 'h93', 'h94', 'h95', 'h96', 'h97', 'h98', 'h99', 'h100',
!                'h101', 'h102', 'h103', 'h104', 'h105', 'h106', 'h107', 'h108', 'h109', 'h110',
!                'hproc1', 'hproc2', 'hproc3',   ! procedural errors
!                'hcor_lum',                     ! common lumi error

!   To treat the uncertainties as absolute use "false"
   Percent     = 116*true
```

User's choice

Provide number of points, columns

Structure of the data table

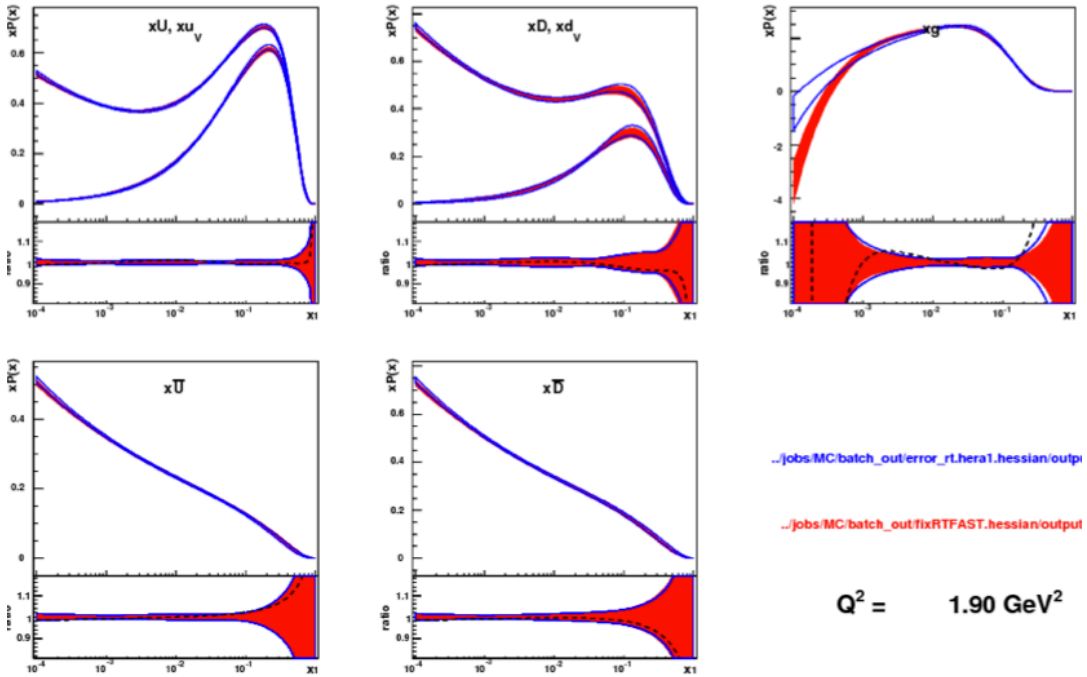
Meaning of columns (predefined names)

Additional useful info in a file

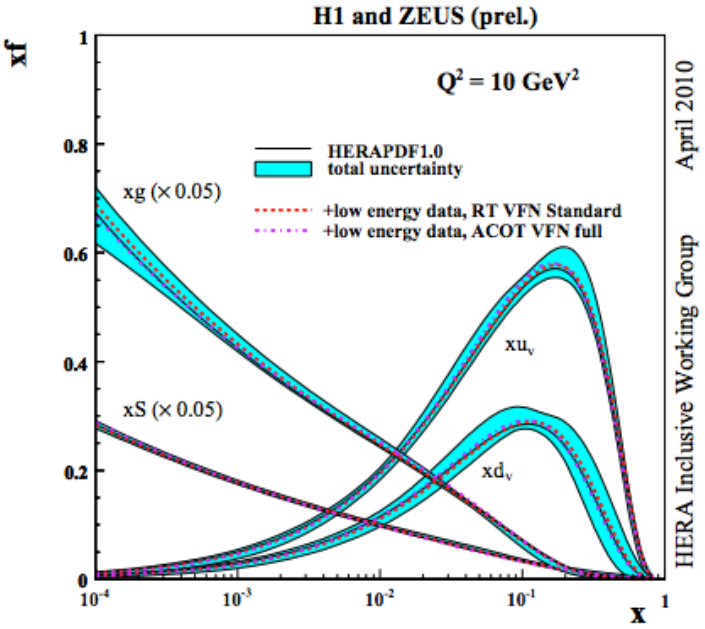
Process identifier

Performance

- HERAFitter Package is designed to be able to provide a fast feedback to the user:
  - Full NLO (or NNLO for DIS) calculations vs FAST options based on kfactors
    - RT scheme (NLO): ~3h for ~1500 iterations
    - RT FAST:
      - {NLO(RT)/NLO(ZMVFNS)}_kfactor x NLO(ZMVNS) ~10min for ~2500 iterations
      - basically it takes the speed of QCDNUM (few miliseconds per iterations)
    - ACOT Full: :
      - {NLO(ACOT Full)/LOmassive(ACOT Full)}_kfactor x LOmassive(ACOT Full)~30min



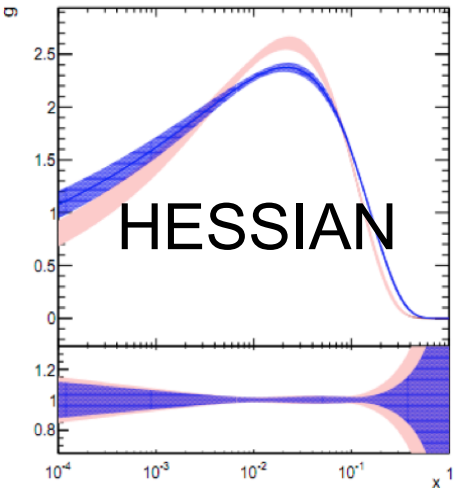
HERAFitter Outputs

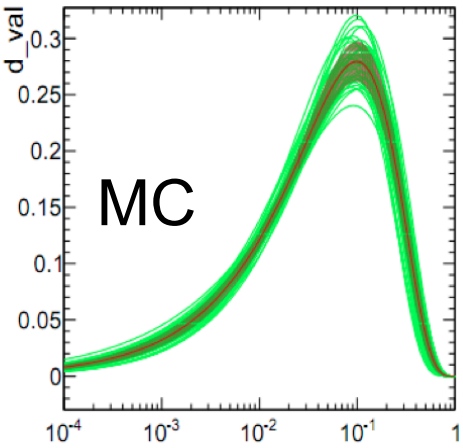


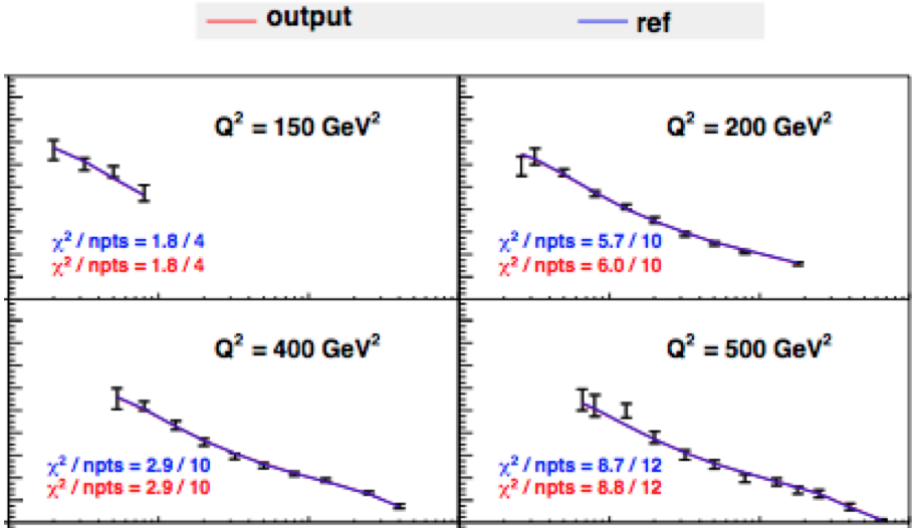
Output contains basic text (and graphic) information on:

- Error logging controlling consistency between input data/fit parameters
- 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]

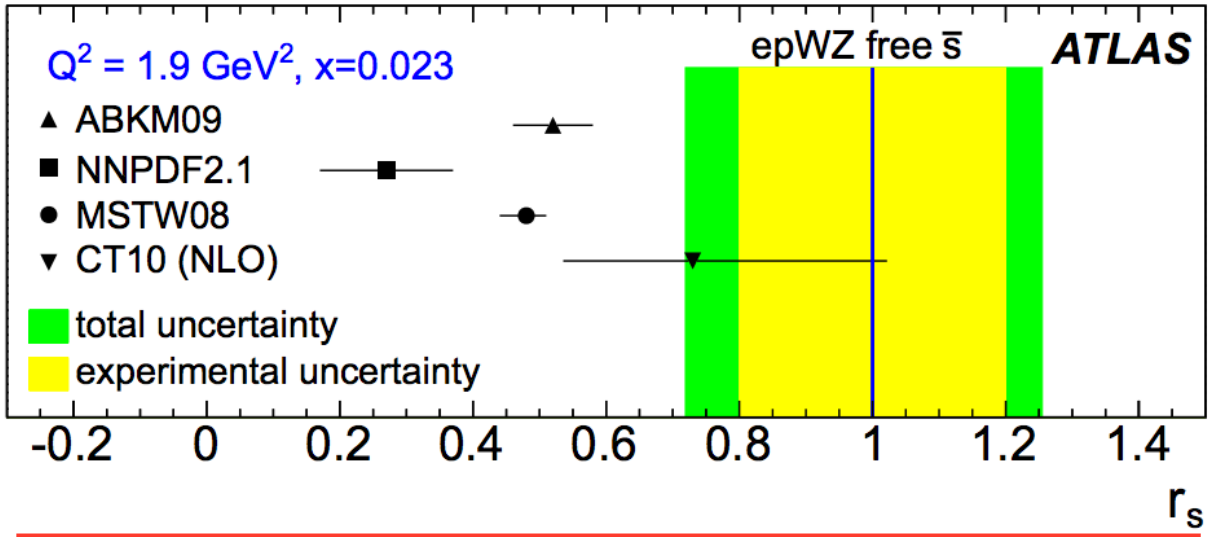






ATLAS determination of the s/d

- 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



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

http://arxiv.org/pdf/1203.4051v1.pdf

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

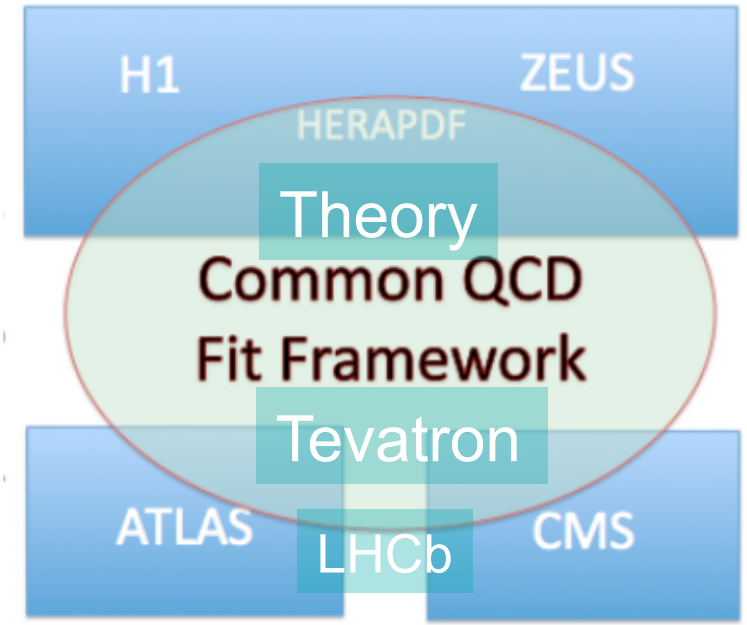
HERAFitter Physics Cases

- 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
  - 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 → Renat Sadykov, Vladimir Kolesnikov
  - Benchmarking of theories
  - Fits using kt evolution
  - Nuclear PDFs

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



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
  - 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 can also add their contribution to the package

- HERAFitter User's interaction
  - Weekly developer's meeting https://znwiki3.ifh.de/HERAFitter/HERAFitter/HERAFitterInternal/FitForumMeetings
  - Monthly users's meeting (https://znwiki3.ifh.de/HERAFitter/HERAFitter/HERAFitterMeetings)

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]
```

Steering card

- Input files (number and names)

```
&InFiles
  ! Number of intput files
    NInputFiles = 4

  ! Input files:
    InputFileNames(1) = 'datafiles/H1ZEUS_NC_e-p_HERA1.0.dat'
    InputFileNames(2) = 'datafiles/H1ZEUS_NC_e+p_HERA1.0.dat'
    InputFileNames(3) = 'datafiles/H1ZEUS_CC_e-p_HERA1.0.dat'
    InputFileNames(4) = 'datafiles/H1ZEUS_CC_e+p_HERA1.0.dat'
&End
```

- Starting scale
- HF scheme:

```
! --- Scheme for heavy flavors :
! ---   HF_SCHEME = 'ZMVFNS' : ZM-VFNS (massless),
! ---   HF_SCHEME = 'RT' :  Thorne-Roberts VFNS (massive)
! ---   HF_SCHEME = 'RT FAST' : Fast approximate TR VFNS

 HF_SCHEME = 'RT'
```

```
! PDF parameterisation style. Pos
!   '10p HERAPDF' -- HERAPDF-like
!   '13p HERAPDF' -- HERAPDF-like
!   '10p H12000'  -- H12000-like (
!   'CTEQ'        -- CTEQ-like par
!   'CHEB'        -- CHEBYSHEV par

 PDFStyle = '10p HERAPDF'
```

- Chisquare style

```
! -- Choice of the chi2 function
!    'H12000'  : Pascaud-like, systematic shifts to theory, n
!    'HERAPDF' : Pascaud-like + "mixed error scaling"
!    'HERAPDF Sqrt'   : Pascaud-like + "sqrt error scaling"
!    'HERAPDF Linear' : Pascaud-like + "linear error scaling"

 CHI2Style = 'HERAPDF'
```

  - NC, CC, DY

```
!--------------------- NC ep  ------

! Rule #1: Q2 cuts
 ProcessName(1)      = 'NC e+-p'
 Variable(1)         = 'Q2'
 CutValueMin(1)      = 3.5
 CutValueMax(1)      = 1000000.0

! Rule #2: x cuts
 ProcessName(2)      = 'NC e+-p'
 Variable(2)         = 'x'
 CutValueMin(2)      = 0.000001
 CutValueMax(2)      = 1.0
```

  - Hessian (Pumplin)-DOBANDS
  - MC errors

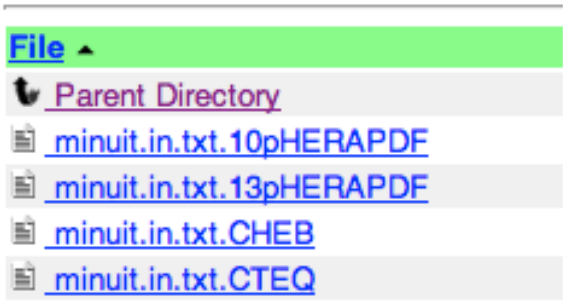
Flexible input, new data tables can be added without the need to recompile.

Minuit card

- One general minuit card for standard parametrisation form:

  $Ax^B(1-x)^C(1+Dx+Ex^2+Fx^3)-Apx^{Bp}(1-x)^{Cp}$

- In the input_steerings:



alphas, fs, fcharm
are treated as fit parameters

```
set title
new 10p HERAPDF
parameters
   1    'Ag'                0.0000      0.
   2    'Bg'                0.213846    0.010000
   3    'Cg'                9.013846    0.500000
   4    'Dg'                0.0000      0.
   5    'Eg'                0.0000      0.
   7    'Aprig'             0.0000      0.
   8    'Bprig'             0.0000      0.
   9    'Cprig'             0.0000      0.
  11    'Auv'               0.0000      0.
  12    'Buv'               0.665589    0.010000
  13    'Cuv'               4.652237    0.500000
  14    'Duv'               0.0000      0.
  15    'Euv'               9.693753    0.500000
  16    'Fuv'               0.0000      0.
  21    'Adv'               0.0000      0.
  22    'Bdv'               0.0000      0.
  23    'Cdv'               4.291377    0.500000
  24    'Ddv'               0.0000      0.
  25    'Edv'               0.0000      0.
  26    'Fdv'               0.0000      0.
  31    'AUbar'             0.0000      0.
  32    'BUbar'             0.0000      0.
  33    'CUbar'             2.582025    0.100000
  34    'DUbar'             0.0000      0.
  35    'EUbar'             0.0000      0.
  41    'ADbar'             0.162609    0.001000
  42    'BDbar'            -0.165110    0.001000
  43    'CDbar'             2.404802    0.100000
  44    'DDbar'             0.0000      0.
  45    'EDbar'             0.0000      0.
  51    'AU'                0.0000      0.
  52    'BU'                0.0000      0.
  53    'CU'                0.0000      0.
  54    'DU'                0.0000      0.
  55    'EU'                0.0000      0.
  61    'AD'                0.0000      0.
  62    'BD'                0.0000      0.
  63    'CD'                0.0000      0.
  64    'DD'                0.0000      0.
  65    'ED'                0.0000      0.
  71    'Asea'              0.0000      0.
  72    'Bsea'              0.0000      0.
  73    'Csea'              0.0000      0.
  74    'Dsea'              0.0000      0.
  75    'Esea'              0.0000      0.
  81    'Adel'              0.0000      0.
  82    'Bdel'              0.0000      0.
  83    'Cdel'              0.0000      0.
  84    'Ddel'              0.0000      0.
  85    'Edel'              0.0000      0.
  95    'alphas'            0.11760     0.
  96    'fs'                0.310       0.
  97    'fch'               0.00        0.
set eps 1.0e-11
```

g

uv

dv

Ub

Db

U

D

sea

del

other