HELMHOLTZ GEMEINSCHAFT

Bundesministerium für Bildung und Forschung

HERAFitter Tutorial
overview and examples

Ringailė Plačakytė,
Voica Radescu

DESY

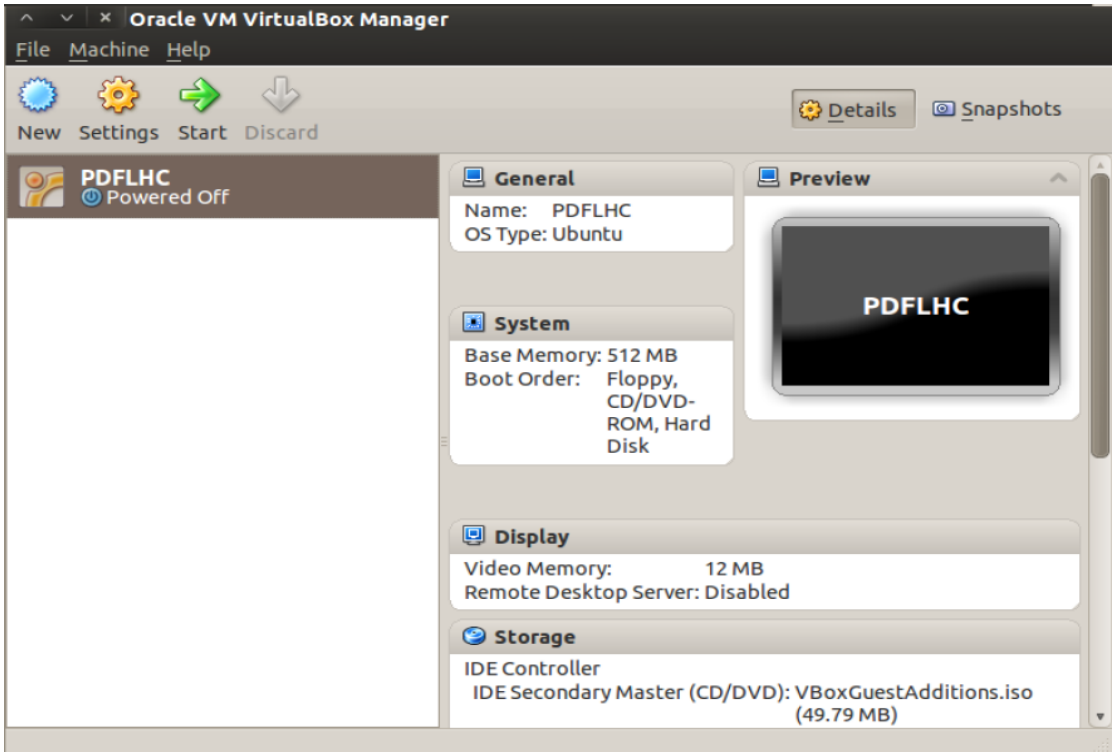
“Proton Structure in the LHC Era” school
22-24 October 2012

Before getting started

Basic requirements for school's tutorials:

- have VirtualBox installed on your laptops (≥ 3 GB space required)
  - 4.1 version or later
- have downloaded image file http://wwwiexp.desy.de/users/schorner/pdf2012/
  - PDFLHC.ova (1.7 GB)

→ open VirtualBox and import PDFLHC.ova file (File → Import Appliance)

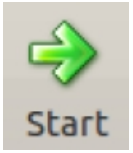


You need to activate box:

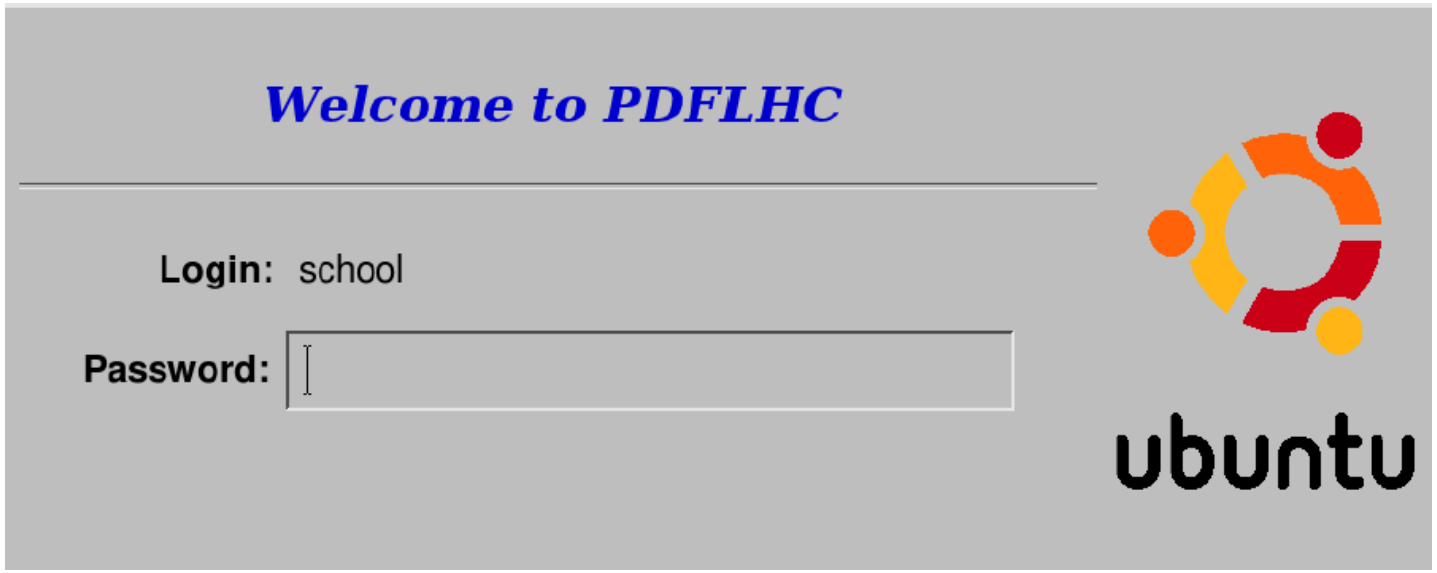
"reinitialize the MAC address of all network cards"

during the import - otherwise ALL school participants will have the same MAC address!

Before getting started



→ login: school
→ password: desi1234



Before getting started

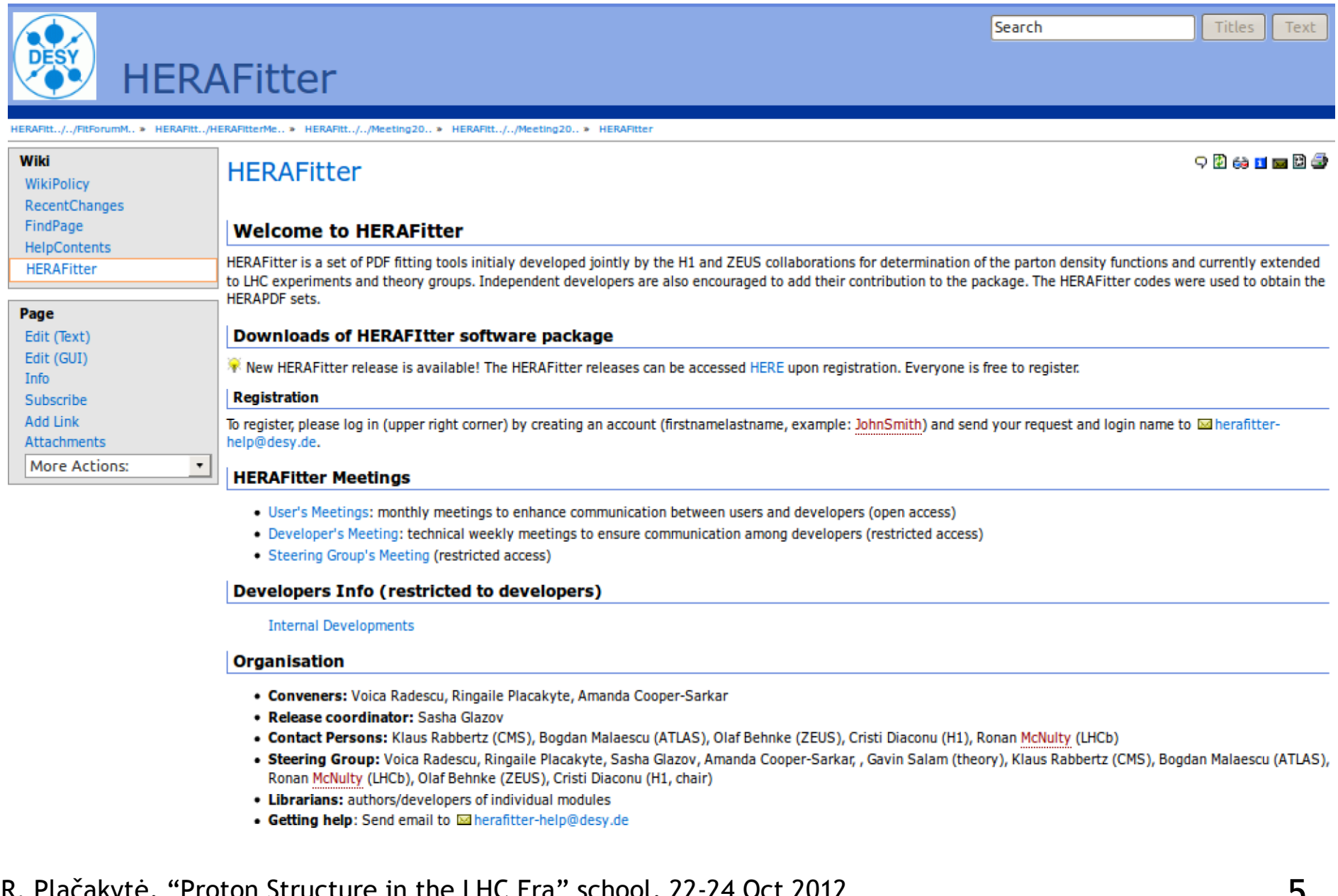
→ Installed programs:

```
applgrid-1.2.4/
applgrid-full-installation-01-02-04/
.bash_history
.bash_logout
.bashrc
bin/
.cache/
.emacs.d/
fastnlo/
fastnlo-2.1.0-0/
.fehbg
.fluxbox/
grids/
hoppet-1.1.5/
include/
largetables/
.lesshst
lhapdf-5.8.8/
lib/
libexec/
MCFM/
nlojet++-4.1.3/
.profile
.profile~
programs/
root/
```

→ available packages:

```
root_v5.32.04.source.tar.gz
qcdnum170006.tar.gz
lhapdf-5.8.8.tar.gz
applgrid-full-installation-01-02-04.tar.gz
Hathor-1.3.tar.gz
herafitter-0.2.1.tgz
hoppet-1.1.5.tar.gz
openqcdrad-1.5.zip
fn-common-lhc.tar.gz
fastnlo-2.1.0-0.tar.gz
applgrid-1.2.4.tar.gz
nlojet++-4.1.3.tar.gz
mcfm-6.3.tar.gz
```

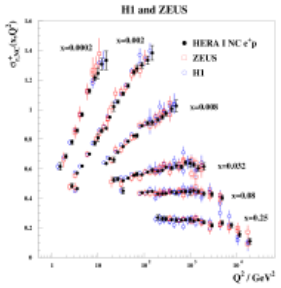
HERAFitter: Project



HERAFitter: Overview

Modular structure of HERAFitter:

experimental data



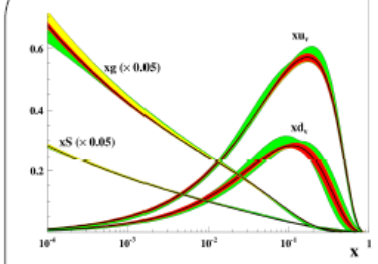
Data: HERA, Tevatron, LHC, fixed target experiments

Processes:
Inclusive DIS, Jets, Diffraction
Drell-Yan, Top production
W and Z production

theory calculations

Heavy Flavour schemes:	MSTW, CTEQ, ABM
Jets, W, Z:	FastNLO, Applgrid
Top:	Hathor
Evolution:	QCDNUM, k_T scheme
Other:	NNPDF reweighing Dipole model

HERAFitter



Parton Density Functions (PDF, DPDF, uPDFS)

$\alpha_s(M_Z)$, m_c, m_b, m_t ...

Theoretical cross sections

Comparisons to other PDFs (lhapdf)

HERAFitter description

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

- standard Fortran method used to input information: namelist files

In the central steering file input data and fitting parameters defined

- contains several namelists, e.g. data cuts and data module

Output contains basic text information to

– control consistency of the input data/fit parameters (error logging),

– report quality of the fit: χ^2, pulls, etc,

– report resulting PDFs: simple text and HERAGRID LHAPDF format.

HERAFitter tutorial: introduction

In today's tutorial we will do:

→ install HERAFitter
→ run default example (central HERAPDF1.0 fit)
→ analyse fit output, create PDFs in LHAPDF format (.LHgrid)
→ run with different heavy flavour scheme
→ learn how to apply/change cuts on data
→ add data set
→ run with external PDF set (LHAPDF)

other (HERAFitter) tutorials in this school:

→ uncertainty treatment in PDFs (S. Glazov)
→ including additional measurements (K. Nowak)

Separate tutorials/lectures on:

→ PDF reweighting (A. Guffanti)
→ FastNLO (K. Robbertz)
→ APPLgrid (P. Starovoitov)
→ OPENQCDrad (S. Alekhin)

HERAFitter: installation

Described in README

Pre-requirements
==========

-- QCDNUM version at least qcdnum-17-00/04
http://mbotje.web.cern.ch/mbotje/qcdnum/Site/QCDNUM17.html

QCDNUM: M. Botje lecture

-- CERNLIB
e.g. from CERN: /afs/cern.ch/sw/lcg/external/cernlib/

-- Optional:
link to recent Root libraries (e.g. version 5.26)
APPLGRID, HATHOR, LHAPDF, NNPDFreweighting tool
(APPLGRID and NNPDF reweighting tool require Root to be installed)

HERAFitter has been tested on various platforms:
SL4, SL5 (32 and 64 bit), Ubuntu 10.10, 11.10, MAC

HERAFitter: installation

In VirtualBox:
(after QCDNUM was installed)

-- QCDNUM version at least qcdnum-17-00/04

```
export QCDNUM_ROOT=/home/school/programs/qcdnum-17-00-06
```

-- CERNLIB

```
export CERN_ROOT=/usr/bin
```

-- ROOT needs the thisroot.sh script to be executed first

```
cd ~/root
. bin/thisroot.sh
```

(I recommend to put it in .bashrc)

HERAFitter: installation

In VirtualBox:
(after QCDNUM was installed)

-- now configure (more on configuration options on next slide) and compile:

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

You can also put that all in initialisation script (e.g. ini.sh) in order to not repeat that every time you login !

HERAFitter: configuration options

→ you can see available option by typing: configure -help

```
./configure  --help
…....

Optional Features:
  --disable-option-checking  ignore unrecognized --enable/--with options
  --disable-FEATURE                do not include FEATURE (same as --enable-FEATURE=no)
  --enable-FEATURE[=ARG]           include FEATURE [ARG=yes]
  --disable-dependency-tracking    speeds up one-time build
  --enable-dependency-tracking     do not reject slow dependency extractors
  --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)
```

Note: before each configuration 'make clean' is recommended !

HERAFitter: data

Data are provided as text files with a namelist header and the main body, as a table:

```
&Data
  Name  = 'CC cross section HERA-I H1-ZEUS combined e-p.'
  NData =  34
  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','x-section','stat','uncor',110*'uncor',4*'ignore'

 …..

!   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', ……..
```

← *define data format*

← *error treatment*

HERAFitter: systematic uncertainties

Three types of systematic uncertainties are distinguished:
- statistical (’stat’)
- uncorrelated (name contains ’uncor’ sub-string)
- correlated (any other name)
- also possible to not take uncertainty into account ('ignore')

Systematic uncertainties are correlated among different data files if they have the same name;

After minimization, shifts and estimated reduction of uncertainty for the nuisance parameters corresponding to the systematic error sources are reported in output/Results.txt file:

A. Glazov lecture and tutorial

Systematic shifts	(sigma)		(error)
1 h1	0.1348	+/-	0.4709
2 h2	-0.6975	+/-	0.8086
3 h3	0.0209	+/-	0.9939
4 h4	-1.0008	+/-	0.8446
5 h5	0.3473	+/-	0.9906
6 h6	0.5485	+/-	0.9268
7 h7	-2.1711	+/-	0.8740
…..			

HERAFitter: data

```
&Data
  Name  = 'CC cross section HERA-I H1-ZEUS combined e-p.'
  NData =  34
  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','x-section','stat','uncor',110*'uncor',4*'ignore'

  NInfo    = 4
  DataInfo =  318.,       -1.,   0.,      0.
  CInfo    = 'sqrt(S)','e charge', 'reduced', 'e polarity'

  IndexDataset = 63
  Reaction = 'CC 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', ……..

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

← *define data format*

← *error treatment*

← *additional information*

← *reaction name*

← *abs or rel errors?*

Note: inclusion of new data tables for existing processes doesn't require code recompilation

HERAFitter: data relation to theory

Theory predictions for a given data file are controlled by the Reaction namelist variable

- at each iteration, a loop is run over the data files and the theory calculations are matched to the data inside GetTheoryForDataset subroutine

```
……..............
elseif (DATASETREACTION(IDataSet).eq.'NC e+-p charm') then
   Call GetNCCharmXsection(IDataSet, HFSCHEME)
elseif (DATASETREACTION(IDataSet).eq.'CC e+-p') then
   Call GetCCXsection(IDataSet, HFSCHEME)
elseif (DATASETREACTION(IDataSet).eq.'CC pp' .or.
$       DATASETREACTION(IDataSet).eq.'CC ppbar' ) then
   Call GetDYCCXsection(IDataSet)
……..........
```

To add a new theory module:

-add initialization call in init_theory.f, see subroutine init_theory_datasets
-add call to your theory module in theory_dispatcher.f, see GetTheoryForDataset

HERAFitter: data relation to theory

Theory predictions for a given data file are controlled by the Reaction namelist variable

- at each iteration, a loop is run over the data files and the theory calculations are matched to the data inside GetTheoryForDataset subroutine

```
Subroutine GetCCXsection(IDataSet, local_hfscheme)
call GetDisXsection(IDataSet, 'CCDIS', local_hfscheme)

C Get indexes for Q2, x and y bins:
      idxQ2 = GetBinIndex(IDataSet,'Q2')
      idxX  = GetBinIndex(IDataSet,'x')
      ….
      do i=1,NDATAPOINTS(IDataSet)

        idx =  DATASETIDX(IDataSet,i)

        X(i)   = AbstractBins(idxX,idx)
        Q2(i)  = AbstractBins(idxQ2,idx)
      enddo

      call CalcReducedXsectionForXYQ2(X,Y,Q2,NDATAPOINTS(IDataSet),
     $     charge,polarity,IDataSet,XSecType, local_hfscheme,XSec)
        ….
        XSec = 0.5*(yplus*F2 + yminus*xF3 – y*y*FL)
```

← *index of abstract bins determined by name*

← *get index of global data array*

← *build the cross section*

To add a new theory module:

-add initialization call in init_theory.f, see subroutine init_theory_datasets
-add call to your theory module in theory_dispatcher.f, see GetTheoryForDataset

HERAFitter: steering.txt

Central steering file to define input data and fitting parameters steering.txt:

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

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

*
* (Optional) Modify renormalisation/factorisation scales, dataset
* dependently. The numbering follows sequential numbering of input files
*
&Scales
    DataSetMuR  = 4*1.0     ! Set muR scale to 1 for all 4 datasets
    DataSetMuF  = 4*1.0     ! Set muF scale to 1 for all 4 datasets
&End
```

← *data files (HERAPDF1.0)*

← *vary μ_R and μ_F scale (currently works only for jet data)*

HERAFitter: steering.txt

```
&H1Fitter
 ITheory = 0   ! =0 use collinear factorisation with QCDNUM

 IOrder  = 2   ! For itheory =0 (collinear factorisation) : LO fit (1) or NLO (2) or NNLO (3)

 Q02     = 1.9 ! Evolution starting scale

! --- 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 scheme, usign k-factor
! ---  HF_SCHEME = 'RT OPT'        :  Thorne-Roberts VFNS (massive)
! ---  HF_SCHEME = 'RT OPT FAST'   : Fast approximate TR VFNS scheme, usign k-factor
! ---  HF_SCHEME = 'ACOT Full' : ACOT - F.Olness Version (massive), using k-factors
! ---  HF_SCHEME = 'ACOT Chi' : ACOT - F.Olness Version (massive), using k-factors
! ---  HF_SCHEME = 'ACOT ZM'   : ACOT - F.Olness Version (massless), using k-factors
! ---  HF_SCHEME = 'FF'        : Fixed Flavour Number Scheme (qcdnum)
! ---  HF_SCHEME = 'FF ABM'    : Fixed Flavour Number Scheme (ABM)

 HF_SCHEME = 'RT'
```

← *choose order*

← *starting scale*

← *treatment for heavy flavour (more on next slide)*

HERAFitter: steering.txt

S. Alekhin and
A. Cooper-Sarkar lectures

```
HF_SCHEME = 'ZMVFNS'          : ZM-VFNS (massless)

HF_SCHEME = 'RT'              : Thorne-Roberts VFNS (massive)
HF_SCHEME = 'RT OPT'          :  Thorne-Roberts VFNS (massive)
HF_SCHEME = 'RT OPT FAST'     : Fast approximate TR VFNS scheme, usign k-factor
HF_SCHEME = 'RT FAST'         : Fast approximate TR VFNS scheme, using k-factor

HF_SCHEME = 'ACOT Full'       : ACOT - F.Olness Version (massive), using k-factors
HF_SCHEME = 'ACOT Chi'        : ACOT - F.Olness Version (massive), using k-factors
HF_SCHEME = 'ACOT ZM'         : ACOT - F.Olness Version (massless), using k-factors

HF_SCHEME = 'FF'              :  Fixed Flavour Number Scheme (qcdnum)
HF_SCHEME = 'FF ABM'          :  Fixed Flavour Number Scheme (ABM)
```

used by MSTW08

used by CTEQ/CT

ABM (OPENQCDRAD, incl running mass)

Note:

RT FAST : kfactors are ratio of SF(RT) to SF(QCDNUM), speed up factor ~15;
ACOT : kfactors are ratio of SF(NLO) to SF(LO)

FF scheme (n_f=3): requires smaller $\alpha_s(M_Z)$, proper scale choice and only NC DIS data can be fitted (NLO coefficients for CC, W→c available in latest OPENQCDRAD version)

HERAFitter: steering.txt

```
! PDF parameterisation style. Possible styles are currently available:
!  '10p HERAPDF' -- HERAPDF-like with extra assumption Buv = Bdv
!  '13p HERAPDF' -- HERAPDF-like with Buv and Bdv floated independently
!  '10p H12000'  -- H12000-like (D,U,Dbar,Ubar+g)
!  'CTEQ'        -- CTEQ-like parameterisation
!  'CHEB'        -- CHEBYSHEV parameterisation based on glu,sea, uval,dval evolved pdfs
!  'LHAPDFQ0'    -- use lhapdf library to define pdfs at starting scale and evolve with local qcdnum parameters
!  'LHAPDF'      -- use lhapdf library to define pdfs at all scales
!  'DDIS'        -- use Diffractive DIS

PDFStyle = '10p HERAPDF'
```

← *choose parametrisation style*

HERAFitter: steering.txt

```
! PDF parameterisation style. Possible styles are currently available:
!  '10p HERAPDF' -- HERAPDF-like with extra assumption Buv = Bdv
!  '13p HERAPDF' -- HERAPDF-like with Buv and Bdv floated independently
!  '10p H12000'  -- H12000-like (D,U,Dbar,Ubar+g)
!  'CTEQ'        -- CTEQ-like parameterisation
!  'CHEB'        -- CHEBYSHEV parameterisation based on glu,sea, uval,dval evolved pdfs
!  'LHAPDFQ0'    -- use lhapdf library to define pdfs at starting scale and evolve with local qcdnum parameters
!  'LHAPDF'      -- use lhapdf library to define pdfs at all scales
!  'DDIS'        -- use Diffractive DIS

PDFStyle = '10p HERAPDF'
```

← *choose parametrisation style*

10p HERAPDF:

A. Cooper-Sarkar lecture

$$
\begin{aligned}
xg(x) &= A_g x^{B_g}(1-x)^{C_g}, \\
xu_v(x) &= A_{u_v} x^{B_{uv}}(1-x)^{C_{uv}}\left(1+E_{u_v}x^2\right), \\
xd_v(x) &= A_{d_v} x^{B_{dv}}(1-x)^{C_{dv}}, \\
x\bar{U}(x) &= A_{\bar{U}} x^{B_{\bar{U}}}(1-x)^{C_{\bar{U}}}, \\
x\bar{D}(x) &= A_{\bar{D}} x^{B_{\bar{D}}}(1-x)^{C_{\bar{D}}}.
\end{aligned}
$$

A: overall normalisation
B: small x behavior
C: $x \to 1$ shape

xg, xu_v, xd_v, $x\bar{U}$, $x\bar{D}$
where $x\bar{U}=x\bar{u}$ and $x\bar{D}=x\bar{d}+x\bar{s}$ at the starting scale ($x\bar{s}=f_s x\bar{D}$)

13p HERAPDF: Buv≠Bdv, $xg(x)=A_g x^{Bg}(1-x)^{Cg} - A'_g x^{B'g}(1-x)^{25}$

HERAFitter: steering.txt

```
! PDF parameterisation style. Possible styles are currently available:
!  '10p HERAPDF' -- HERAPDF-like with extra assumption Buv = Bdv
!  '13p HERAPDF' -- HERAPDF-like with Buv and Bdv floated independently
!  '10p H12000'  -- H12000-like (D,U,Dbar,Ubar+g)
!  'CTEQ'        -- CTEQ-like parameterisation
!  'CHEB'        -- CHEBYSHEV parameterisation based on glu,sea, uval,dval evolved pdfs
!  'LHAPDFQ0'    -- use lhapdf library to define pdfs at starting scale and evolve with local qcdnum parameters
!  'LHAPDF'      -- use lhapdf library to define pdfs at all scales
!  'DDIS'        -- use Diffractive DIS

PDFStyle = '10p HERAPDF'
```

← *choose parametrisation style*

Note:

The χ^2 minimization in HERAFitter is based on MINUIT package which uses minuit steering cards included in minuit.in.txt file

HESSE reports improved error estimation and correlations among PDF parameters

HERAFitter: minuit.in.txt

```
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
  ....
  23   'Cdv'             4.291377    0.500000
  .....
  33   'CUbar'           2.582025    0.100000
  41   'ADbar'           0.162609    0.001000
  42   'BDbar'          -0.165110    0.001000
  43   'CDbar'           2.404802    0.100000
 ......
```

other parametrisation styles available in *input_steering/*

PDF parameters for HERAPDF1.0

```
*set print 3
call fcn 3
*migrad 200000
*hesse

return
```

1 iteration

run minimisation with error estimation (HESSE)

HERAFitter: steering.txt

```
! PDF parameterisation style. Possible styles are currently available:
!  '10p HERAPDF' -- HERAPDF-like with extra assumption Buv = Bdv
!  '13p HERAPDF' -- HERAPDF-like with Buv and Bdv floated independently
!  '10p H12000'  -- H12000-like (D,U,Dbar,Ubar+g)
!  'CTEQ'        -- CTEQ-like parameterisation
!  'CHEB'        -- CHEBYSHEV parameterisation based on glu,sea, uval,dval evolved pdfs
!  'LHAPDFQ0'    -- use lhapdf library to define pdfs at starting scale and evolve with local qcdnum parameters
!  'LHAPDF'      -- use lhapdf library to define pdfs at all scales
!  'DDIS'        -- use Diffractive DIS

 PDFStyle = '10p HERAPDF'

! -- Choice of the chi2 function
!    'H12000'  : Pascaud-like, systematic shifts to theory, no scaling of stat, uncor errors.
!    'HERAPDF' : Pascaud-like + "mixed error scaling"
!    'HERAPDF Sqrt'   : Pascaud-like + "sqrt error scaling"
!    'HERAPDF Linear' : Pascaud-like + "linear error scaling"

 CHI2Style = 'HERAPDF'

! Debug flag
 LDEBUG     = False
```

← choose parametrisation style (has to be consistent with minuit.in.txt)

← choose χ^2 style

HERAFitter: steering.txt

```
* Add extra to minuit parameters. These MUST include alpha_S and fs
*
&ExtraMinimisationParameters
  name  = 'alphas',   'fs',   'fcharm'
  value =  0.1176 ,   0.31,      0.
  step  =  0.0    ,   0.0 ,      0.     ! set to 0 to avoid minimisation
&End


* Output steering cards
&Output
 ! -- Error bands on parton distributions
 DoBands = True

 ! -- Q2 values at which the pdfs & errors are done (up to 20)
 Q2VAL = 1.9, 4., 10., 100., 6464, 8317
 ……
&End

 * Process dependent cuts
 &Cuts
  !---------------------  CC ep  ------------------

   ProcessName(3)     = 'CC e+-p'
   Variable(3)        = 'Q2'
   CutValueMin(3)     = 3.5
   CutValueMax(3)     = 1000000.0

   ProcessName(4)     = 'CC e+-p'
   Variable(4)        = 'x'
   CutValueMin(4)     = 0.000001
   CutValueMax(4)     = 1.0
```

← *extra parameters*

← *uncertainty calculation (alternative is MC method)*

A. Glazov lecture and tutorial

← *apply cuts on data*

To run: bin/FitPDF < steering.txt

HERAFitter: output

contains basic text information to
– control consistency of the input data/fit parameters (error logging)
– report quality of the fit: χ^2, pulls, etc.
– report resulting PDFs: simple text and HERAGRID in LHAPDF format

```
 After minimisation     575.04   582     0.988

Dataset   1    106.61   145  NC cross section HERA-I H1-ZEUS combined e-p.
Dataset   2    419.26   379  NC cross section HERA-I H1-ZEUS combined e+p.
Dataset   3     19.82    34  CC cross section HERA-I H1-ZEUS combined e-p.
Dataset   4     29.35    34  CC cross section HERA-I H1-ZEUS combined e+p.
 Correlated Chi2    0.0000000000000000
 --------- in store-pdfs -------
cpu_time     39.86     40.17      0.31


          **********************************
          ***       Error Summary        ***
          **********************************

 Total number of logged errors:          710
 Total number of errors not recorded:      0
```

HERAFitter: output

Fit parameters after minimistion are stored in minuit.out.txt

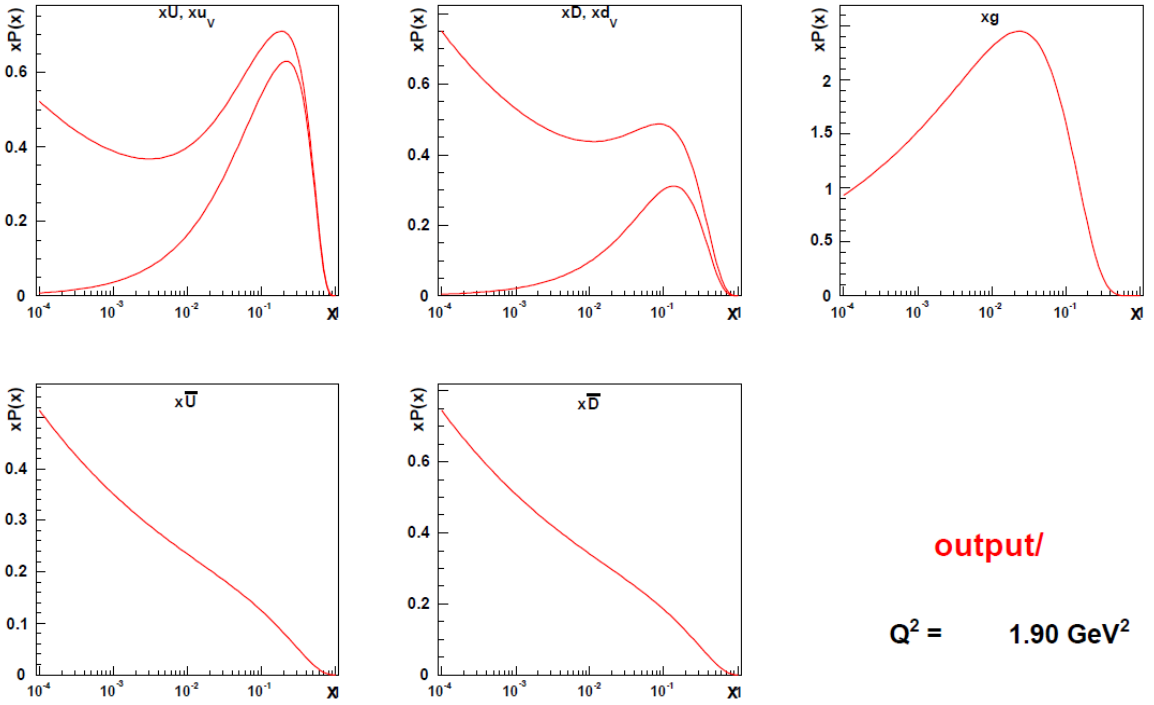
```
FCN=  575.0359    FROM MIGRAD    STATUS=CONVERGED    615 CALLS      618 TOTAL
                EDM=  0.15E-05    STRATEGY= 1      ERROR MATRIX ACCURATE

 EXT PARAMETER                                   STEP         FIRST
 NO.   NAME      VALUE            ERROR          SIZE      DERIVATIVE
  1   Ag          0.0000       constant
  2   Bg          0.21632       0.29456E-01   0.22902E-03   0.25760
  3   Cg          9.0533        0.67835       0.36572E-02  -0.15650E-01
 ....
 12   Buv         0.66922       0.25666E-01   0.13553E-03   0.55788
 13   Cuv         4.6363        0.16280       0.11213E-02  -0.62138E-01
 15   Euv         9.5510        1.8416        0.98395E-02   0.72997E-02
 23   Cdv         4.2411        0.40275       0.42453E-02  -0.17455E-01
 33   CUbar       2.6620        0.42256       0.43560E-02  -0.20266E-01
 ....
 41   ADbar       0.16364       0.66138E-02   0.63512E-04  -0.67733
 42   BDbar      -0.16460       0.54047E-02   0.43168E-04   1.6647
 43   CDbar       2.4644        0.52313       0.28737E-02  -0.23417E-01
```

... and grid file in LHAPDF format
(use tools/tolhapdf.cmd to create PDF.LHgrid file)

To plot: bin/DrawResults output

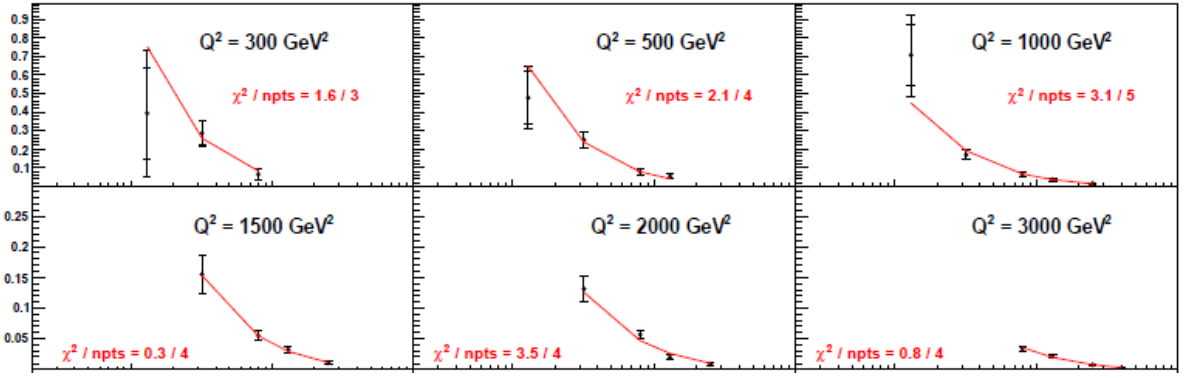
HERAFitter: output

- PDFs at specified scales



- Theory predictions for each data point

(also in output/fittedresults.txt)



HERAFitter: error handling

HERAFitter has a simple error logging facility, based on an H1 tool. The code allows for 4 levels of errors (I,W,S,F), with a steerable printout and sorted summary of the messages at the end of the run:

```
*********************************
***      Error Summary       ***
*********************************

Total number of logged errors:          709
Total number of errors not recorded:      0

List of errors sorted by severity level:

*-------------------------------------------------------------------------
* Module |   Error |   Error |
*   Name |    Type |   Count |   Error Description
*-------------------------------------------------------------------------

 Informational messages:
*----------------------
 H1Fitter  12020501        1 I: steering.txt has been read successfully
 H1Fitter  12020502        1 I: data tables have been read successfully
 H1Fitter  12020503        1 I: theory modules initialised successfully
 H1Fitter  12020504        1 I: read minuit input params from file minuit.in.txt
 H1Fitter  12020515      705 I: FCN is called

*-------------------------------------------------------------------------
 End of Error Summary
```

HERAFitter: changing cuts in data

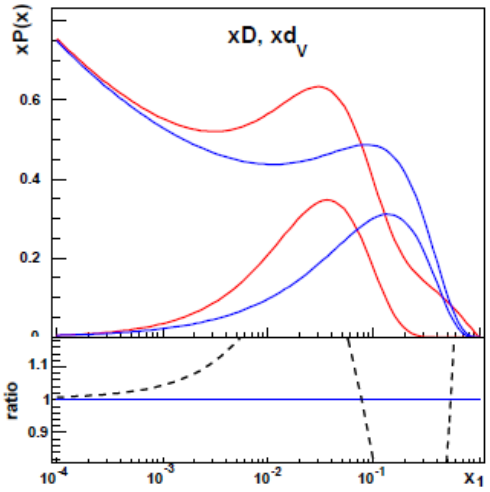
```
* Process dependent cuts
&Cuts
 !--------------------  CC ep  ------------------

  ProcessName(3)     = 'CC e+-p'
  Variable(3)        = 'Q2'
  CutValueMin(3)     = 3.5
  CutValueMax(3)     = 500.0

  ProcessName(4)     = 'CC e+-p'
  Variable(4)        = 'x'
  CutValueMin(4)     = 0.000001
  CutValueMax(4)     = 1.0
```

← *apply cuts on data*

Output with $Q^2 < 500$ GeV^2 cut in CC data set:

```
After minimisation     521.46   521     1.001

Dataset    1    106.92   145  NC cross section HERA-I H1-ZEUS combined e-p.
Dataset    2    411.22   379  NC cross section HERA-I H1-ZEUS combined e+p.
Dataset    3      1.75     3  CC cross section HERA-I H1-ZEUS combined e-p.
Dataset    4      1.57     4  CC cross section HERA-I H1-ZEUS combined e+p.
 Correlated Chi2    0.0000000000000000
 --------- in store-pdfs -------
```

Removing charged current DIS data we lose information on valence quarks →



HERAFitter: use different heavy flavour scheme

steering.txt:

```
HF_SCHEME = 'ZMVFNS'          : ZM-VFNS (massless)

HF_SCHEME = 'RT'              : Thorne-Roberts VFNS (massive)
HF_SCHEME = 'RT OPT'          :  Thorne-Roberts VFNS (massive)
HF_SCHEME = 'RT OPT FAST'     : Fast approximate TR VFNS scheme, usign k-factor
HF_SCHEME = 'RT FAST'         : Fast approximate TR VFNS scheme, using k-factor

HF_SCHEME = 'ACOT Full'       : ACOT - F.Olness Version (massive), using k-factors
HF_SCHEME = 'ACOT Chi'        : ACOT - F.Olness Version (massive), using k-factors
HF_SCHEME = 'ACOT ZM'         : ACOT - F.Olness Version (massless), using k-factors

HF_SCHEME = 'FF'              :  Fixed Flavour Number Scheme (qcdnum)
HF_SCHEME = 'FF ABM'          :  Fixed Flavour Number Scheme (ABM)
```

used by MSTW08

used by CTEQ/CT

ABM (OPENQCDRAD, incl running mass)

```
HF_SCHEME = 'ACOT Chi'
```

Output with ACOT-chi scheme:

```
After minimisation     579.59   582     0.988

 Dataset   1    108.58   145  NC cross section HERA-I H1-ZEUS combined e-p.
 Dataset   2    421.17   379  NC cross section HERA-I H1-ZEUS combined e+p.
 Dataset   3     19.69    34  CC cross section HERA-I H1-ZEUS combined e-p.
 Dataset   4     30.15    34  CC cross section HERA-I H1-ZEUS combined e+p.
  Correlated Chi2    0.0000000000000000
```

HERAFitter: adding additional data

Including e.g. CMS Z rapidity data (2010) arXiv:1110.4973

```
&InFiles
  ! Number of intput files
    NInputFiles = 5
 ! Input files:
    InputFileNames(1) = 'datafiles/hera/H1ZEUS_NC_e-p_HERA1.0.dat'
    InputFileNames(2) = 'datafiles/hera/H1ZEUS_NC_e+p_HERA1.0.dat'
    InputFileNames(3) = 'datafiles/hera/H1ZEUS_CC_e-p_HERA1.0.dat'
    InputFileNames(4) = 'datafiles/hera/H1ZEUS_CC_e+p_HERA1.0.dat'
    InputFileNames(5) = 'datafiles/lhc/cms/CMS_Z_boson+Rapidity.dat'
&End


&Data
  Name = 'CMS Boson rapidity'
  ….

  NInfo   = 6
  DataInfo = 7000., 0., 0.,60.,120.,1.
  CInfo    = 'sqrt(S)', 'pte cut','ptnu cut','Minv min','Minv max','Normalised'
  IndexDataset = 32

  TheoryInfoFile = 'theoryfiles/CMS_20111022_KFACTORSv1.dat'
  TheoryType     = 'kfactor'
  NKFactor       = 1
  KFactorNames = 'Z0'

  Percent = False, False
!y1   y2    Combined Normalized CS    d(Combined Normalized CS)
```

kfactors (LO to NLO)

Please add this missing line!!

HERAFitter: adding additional data

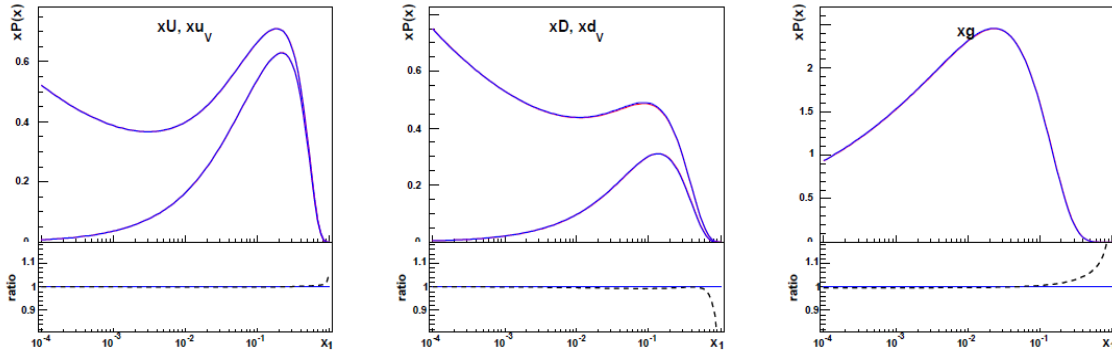
Including e.g. CMS Z rapidity data (2010) arXiv:1110.4973

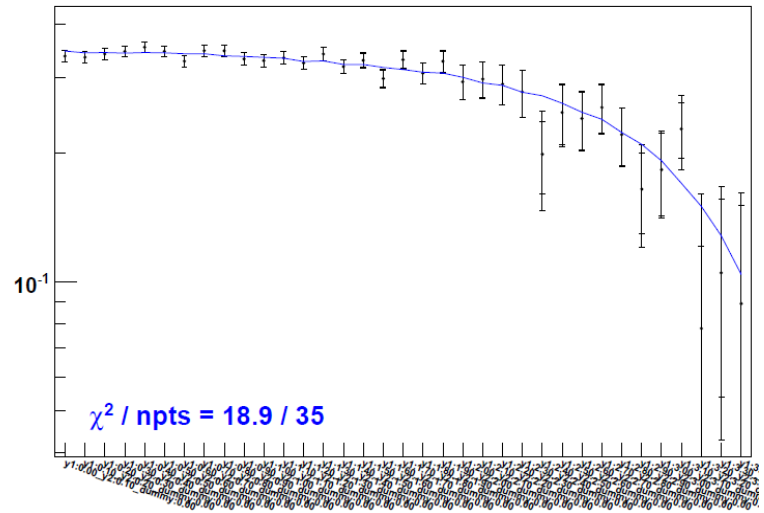
First iteration 594.16372572273519 617 0.96298821024754488
After minimisation 593.95 617 0.963

Partial chi2s
Dataset 1 106.83 145 NC cross section HERA-I H1-ZEUS combined e-p.
Dataset 2 419.07 379 NC cross section HERA-I H1-ZEUS combined e+p.
Dataset 3 19.76 34 CC cross section HERA-I H1-ZEUS combined e-p.
Dataset 4 29.43 34 CC cross section HERA-I H1-ZEUS combined e+p.
Dataset 5 18.85 35 CMS Boson rapidity

Correlated Chi2 0.0000000000000000

(generic plotting tool)





No effect to PDF due to not accounted data correlations
→ see Sasha's and Krzys tutorials for data inclusion with correlations

HERAFitter: running with external PDFs (LHAPDF)

→ useful for various cross checks and data to different PDF comparisons

INSTALLATION with LHAPDF

==========

a) include libLHAPDF.so library in your LD_LIBRARY_PATH variable

export LD_LIBRARY_PATH=/home/school/lhapdf-5.8.8/lib:$LD_LIBRARY_PATH

b) define the place where lhapdf grid files are stored:

export LHAPATH=/usr/local/share/lhapdf/PDFsets

c) enable the lhapdf in the configure: ./configure --enable-lhapdf

Note: before each configuration 'make clean' is recommended !

HERAFitter: running with external PDFs (LHAPDF)

steering.txt:

```
! PDF parameterisation style. Possible styles are currently available:
!  '10p HERAPDF' -- HERAPDF-like with extra assumption Buv = Bdv
…..
!  'LHAPDFQ0'    -- use lhapdf library to define pdfs at starting scale and evolve with local qcdnum parameters
!  'LHAPDF'      -- use lhapdf library to define pdfs at all scales
!  'DDIS'        -- use Diffractive DIS

 PDFStyle = 'LHAPDF'

 …..
 * (Optional) LHAPDF sttering card
 *
 &lhapdf
   LHAPDFSET  = 'CT10.LHgrid' ! LHAPDF grid file
  ILHAPDFSET = 0              ! Set number withing PdfSet
 &End
```

Note: we run with one iteration fit without minimisation → need to modify minuit.in.txt

Output using CT10.LHgrid PDFs:

```
After minimisation     1316.67   617     2.134

  Partial chi2s
Dataset    1    126.44   145  NC cross section HERA-I H1-ZEUS combined e-p.
Dataset    2   1102.13   379  NC cross section HERA-I H1-ZEUS combined e+p.
Dataset    3     20.32    34   CC cross section HERA-I H1-ZEUS combined e-p.
Dataset    4     44.77    34   CC cross section HERA-I H1-ZEUS combined e+p.
Dataset    5     23.01    35   CMS Boson rapidity

 Correlated Chi2    0.0000000000000000
```

HERAFitter: other (here not covered) tools

NNPDF reweighting in HERAFitter:

→ cross check to HERA fits

→ check floating of data normalizations / errors

→ compare results from the reweighting tool with real fit for a data set

NNPDF reweighting:
A. Guffanti lecture

Requires: ./configure --enable_nnpdfWeight

Interface to HATHOR (Hadronic Top and Heavy quarks crOss Section calculatoR) in HERAFitter is available

http://www-zeuthen.desy.de/~moch/hathor/

- useful to study the impact of PDFs on ttbar cross sections,
- goal: include ttbar cross sections in PDF fit performing an m_t scan

Requires: ./configure --enable-hathor

HERAFitter: other (here not covered) tools

Diffractive PDF fit to DIS

Diffractive DIS data are fitted within the 'proton vertex factorisation' approach where the diffractive DIS is mediated by the exchange of hard Pomeron and a secondary Reggeon. The model was used in previous HERA fits:

1. ZEUS Collaboration, S. Chekanov, et al., Nucl. Phys. B 831 (2010) 1.
2. H1 Collaboration, A. Aktas, et al., Eur. Phys. J. C 48 (2006) 715.

- input files (minuit.in.txt.DIFFRACTION, steerig.txt.DIFFRACTION and ewparam.txt.DIFFRACTION) have be copied to minuit.in.txt, steering.txt and ewparam.txt, respectively before running the program.

(for more details see DiffDIS/README)

Fitting with DIPOLE models

DIPOLE (GBW, IIM) or DGLAP+DIPOLE models are available in HERAFitter,
- activation via steering (input_steering/steering.txt.DIPOLE)

Summary

HERAFitter is a publicly available generic tool for PDF determination

→ flexible (with modular structure)

→ has many different options and tools for fast data studies

→ under continues development

In todays tutorial we covered only basic options, more in next days tutorials

HERAFitter mail-support: herafitter-help@desy.de

Weekly developers meetings:
https://www.herafitter.org/HERAFitter/HERAFitter/HERAFitterInternal/FitForumMeetings

Monthly users' meetings:
https://www.herafitter.org/HERAFitter/HERAFitter/HERAFitterMeetings

NEXT user's meeting: 29th October, 15:00 CET

Additional material

HERAFitter: PDF uncertainties

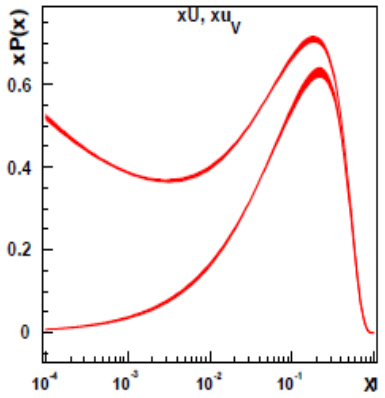
HERAFitter package includes extension of the MINUIT package from J.Pumplin which allows to perform detailed error analysis of PDFs and report full errors in terms of eigenvectors
→ steerable by DoBands namelist parameter

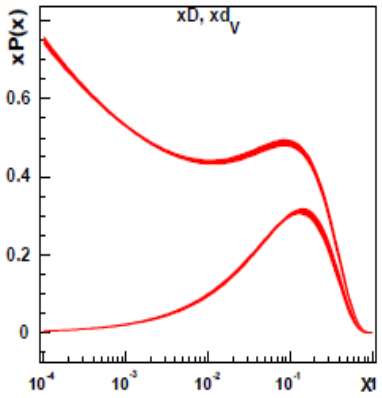
In addition, HERAFitter has code to perform error analysis using toy MC method
→ method uses uncertainties as reported in the data tables and can generate replica around the actual data values, to study error propagation, and around the predicted theoretical values, to study biases from the χ^2 definition.

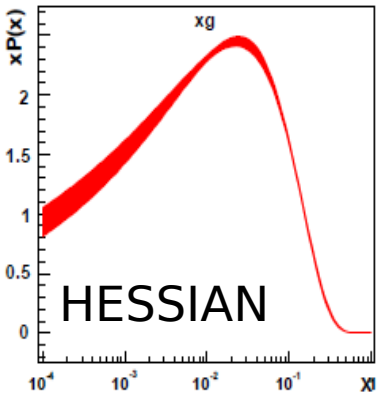
→ Sasha's lecture and tutorial

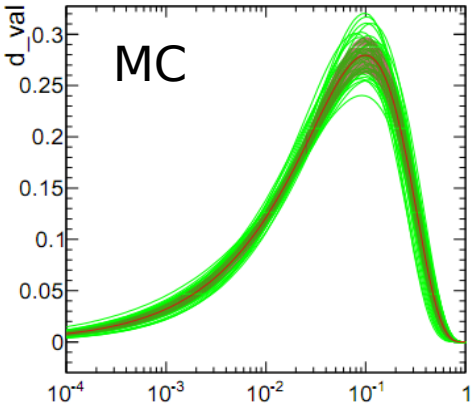
HERAFitter: output

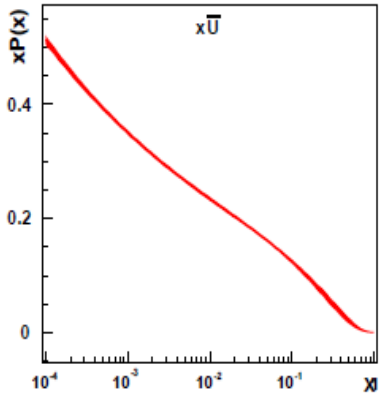
- Hessian vs MC error estimation

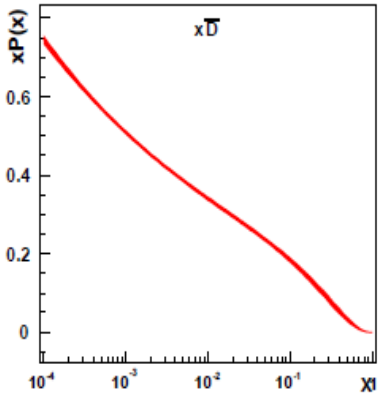












output/

Q^2 = 1.90 GeV^2

HERAFitter: data model

Several χ^2 models are available, steerable by the CHI2Style parameter.
For example, CHI2Style = 'HERAPDF' turns on default HERAPDF treatment:

$$\chi^2 = \sum_i \frac{(D_i - T_i^*)^2}{(\delta_i^{unc})^2}$$

Uncorrelated error

$$T_i^* = T_i + \sum_j \xi_j \delta_i^{cor,j}$$

Nuisance parameter

Correlated error

$$\delta_i^{cor,j} = \beta_{ij} T_i$$

Relative corr. error

Also covariance matrix approach available (NOT IN BETA)
(defined in CHI2Style and correlations provided in data file):

Full covariance matrix approach (new)

$$\chi^2 = \sum_{i,j} (D_i - T_i) Cov_{i,j}^{-1} (D_j - T_j)$$

statistical uncorrelated correlated

$$Cov = C^{stat} + C^{uncor} + C^{corr}$$

$$C_{i,j}^{stat} = Corr^{stat}\, \delta_i^{stat} \delta_j^{stat}$$

Statistical correlations between bins

$$C_{i,j}^{uncor} = \delta_{ij} \delta_i^{unc} \delta_j^{unc}$$

Kronecker delta

$$C_{i,j}^{corr} = \sum_k \delta_i^{cor,k} \delta_j^{cor,k}$$

Sum over all correlated systematics

Advantage: statistical correlations can be taken into account
Disadvantage: All errors in one bin, no nuisance parameters

HERAFitter: NNPDF reweighting tool

NNPDF reweighting in HERAFitter:

→ cross check to HERA fits

→ check floating of data normalizations / errors

→ compare results from the reweighting tool with real fit for a data set

Requires: ./configure --enable_nnpdfWeight

```
* (Optional) NNPDF steering cards
*
&nnpdf
  FLAGNNPDF    = True              ! Should reweighting be done?
  NNPDFSET     = 'NNPDF21_100.LHgrid' ! LHAPDF grid file
  NNPDFRWDATA  = 'test'            ! arbitrary name for new datasample to be put in
  NNPDFREWEIGHTMETHOD = 1          ! either 1=chi2 or 2=data
  DONNPDFONLY  = True              ! do / do not run usual HERA fit
  NNPDFOUTREPLICAS = 10            ! how many output replica of the NNPDF should be kept?
&End
```

calculated theory predictions and χ^2s provided to NNPDF code

→ C++ NNPDF reweighting code determines which replicas to keep (based on HERA chi2 or NNPDF chi2 from data/theory)

→ Output new LHAPDF reweighted NNPDF set

(more in NNPDF/README)

HERAFitter: HATHOR

Interface to HATHOR (Hadronic Top and Heavy quarks crOss Section calculatoR) in HERAFitter is available

1C) INSTALLATION with Hathor
==========

a) Download Hathor from http://www-zeuthen.desy.de/~moch/hathor/
and install it according to the instructions given there
(requires LHAPDF library -- create a symbolic link to lhapdf:
ln -s your_lhapdf_path lhapdf)

b) Define a variable HATHOR_ROOT such that $HATHOR_ROOT points to the
directory of your Hathor installation

c) Install the HERAFitter as described above but configuring it
with the option "--enable-hathor" before building it

- *for each fit iteration PDFs and α_s are read via corresponding QCDNUM routines and fed into Hathor for calculating a new ttbar prediction*

- useful to study the impact of PDFs on ttbar cross sections,
- goal: include ttbar cross sections in PDF fit performing an m_t scan

PDF determination in HERAPDF 1.0

DGLAP at NLO → QCD predictions

PDFs parametrised (at starting scale Q^2_0) using standard parametrisation form:

$$
\begin{aligned}
xg(x) &= A_g x^{B_g}(1-x)^{C_g}, \\
xu_v(x) &= A_{u_v} x^{B_{u_v}}(1-x)^{C_{u_v}}\left(1+E_{u_v}x^2\right), \\
xd_v(x) &= A_{d_v} x^{B_{d_v}}(1-x)^{C_{d_v}}, \\
x\bar{U}(x) &= A_{\bar{U}} x^{B_{\bar{U}}}(1-x)^{C_{\bar{U}}}, \\
x\bar{D}(x) &= A_{\bar{D}} x^{B_{\bar{D}}}(1-x)^{C_{\bar{D}}}.
\end{aligned}
$$

A: overall normalisation

B: small x behavior

C: x→1 shape

The optimal number of parameters chosen by saturation of the χ^2
- central fit with:
  10 free parameters for HERA I data
  14 for HERA I and II data

xg, xu_v, xd_v, $x\bar{U}$, $x\bar{D}$
where $x\bar{U}=x\bar{u}$ and $x\bar{D}=x\bar{d}+x\bar{s}$ at the starting scale ($x\bar{s}=f_s x\bar{D}$ with $f_s=0.31$)

A_g, A_{uv}, A_{dv} are fixed by sum rules

extra constrains for small x behavior of d- and u-type quarks:
$B_{uv}=B_{dv}$, $B_{\bar{U}}=B_{\bar{D}}$, $A_{\bar{U}}=A_{\bar{D}}(1-f_s)$ for $\bar{u}=\bar{d}$ as $x\to 0$

HERA PDFs

Uncertainties:

experimental

small experimental uncertainties

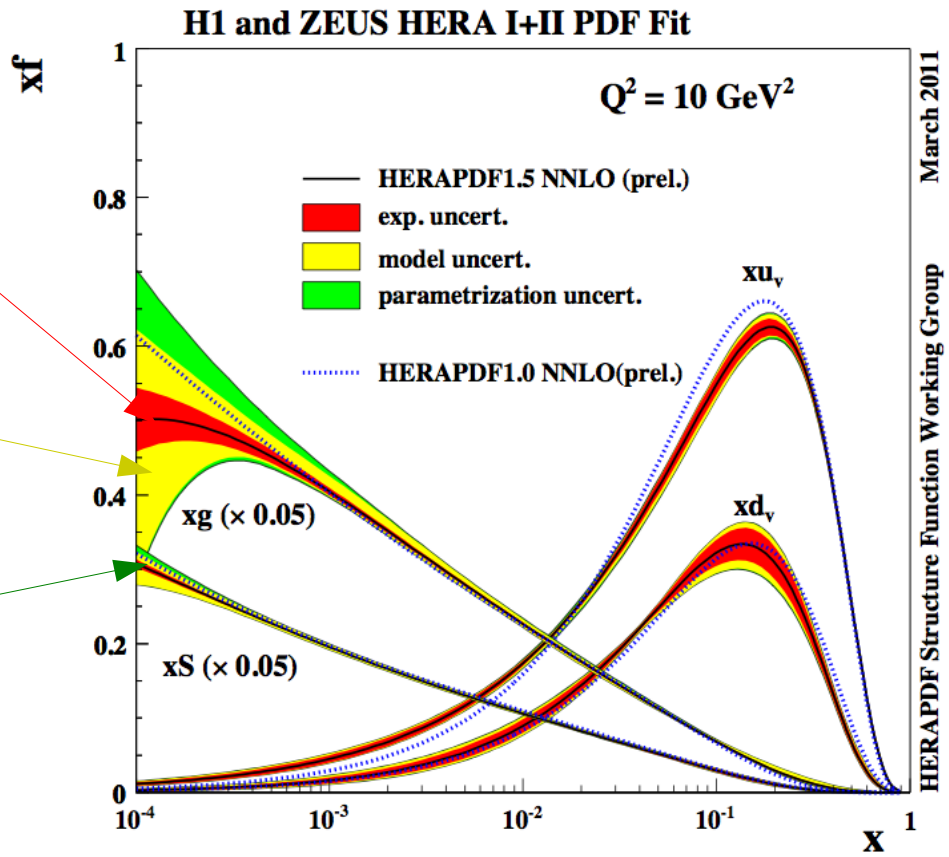
model

model uncertainties from: Q^2_{min}, f_s, m_C, m_B

parametrisation

from different parametrisation assumptions

HERAPDF1.5f NNLO
(based on combined HERA I+II data)



HERAFitter: pp, ppbar data: DY code

DY integration code:

Simple LO cross section formulae: DY NC: $pp \to Z/\gamma \to e^+e^-$

$$\frac{d\sigma^2_\gamma}{dMdydcos\theta^*} = N_c C^2_{q\bar{q}} \frac{8\alpha^2}{3M^3}\tau \times \sum_q e_q^2 f_q(x_1, M) f_{\bar{q}}(x_2, M) F_{q\bar{q}}(1+\cos^2\theta^*, \cos\theta^*)$$

DY CC: $pp \to W^\pm \to e^\pm\nu$

$$\frac{d\sigma^3_{W^\pm}}{dMdydcos\theta^*} = \frac{\pi\alpha^2}{48 s_W^4} M\tau \frac{(1-cos\theta^*)^2}{(M^2-M_W^2)^2+\Gamma_W^2 M_W^2} \times \sum_{qq'} V_{qq'} f_q(x_1, M) f_{q'}(x_2, M)$$

where $\tau = \frac{M^2}{S_0}$, S_0 - beam energy.
$F_{q\bar{q}}(1+\cos^2\theta^*, \cos\theta^*)$ is a linear homogenious dependence on $1+\cos^2\theta^*$ and $\cos\theta^*$.

Sapronov Andrey () | DY integrator PDF fitting. | June 28, 2011 | 2 / 5

- Kfactors are determined from MCFM
- Cross checks of results:
  - between LO x kfactors and NLO using Applgrid
  - between LO x kfactors and DY code from J.Stirling (ZEUSFitter)