NNPDFs for the LHC
&
the search for deviations from DGLAP evolution in HERA data

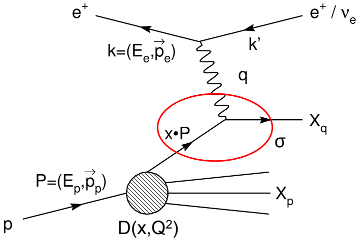
Alberto Guffanti

Niels Bohr International Academy & Discovery Center
Niels Bohr Institute - Copenhagen

INT, Seattle,
October 27, 2011

What are Parton Distribution Functions?

- Consider a process with one hadron in the initial state



- According to the **Factorization Theorem** we can write the cross section as

$$d\sigma = \sum_a \int_0^1 \frac{d\xi}{\xi} D_a(\xi,\mu^2) d\hat{\sigma}_a\left(\frac{x}{\xi},\frac{\hat{s}}{\mu^2},\alpha_s(\mu^2)\right) + \mathcal{O}\left(\frac{1}{Q^p}\right)$$

What are Parton Distribution Functions?

- The **absolute value** of PDFs at a given x and Q^2 **cannot be computed** in QCD Perturbation Theory
  (Lattice? In principle yes, but ...)

- ... but the **scale dependence** is governed by **DGLAP** evolution equations

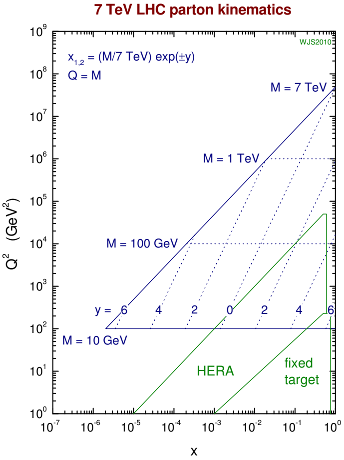
$$\frac{\partial}{\ln Q^2} q^{NS}(\xi, Q^2) = P^{NS}(\xi, \alpha_s) \otimes q^{NS}(\xi, Q^2)$$

$$\frac{\partial}{\ln Q^2} \begin{pmatrix} \Sigma \\ g \end{pmatrix} (\xi, Q^2) = \begin{pmatrix} P_{qq} & P_{qg} \\ P_{gq} & P_{gg} \end{pmatrix} (\xi, \alpha_s) \otimes \begin{pmatrix} \Sigma \\ g \end{pmatrix} (\xi, Q^2)$$

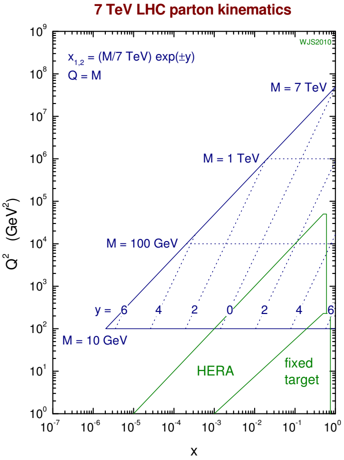
- ... and the **splitting functions** P can be computed in PT and are known up to **NNLO**

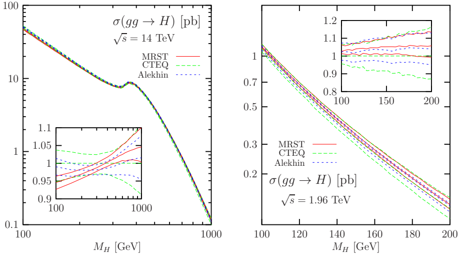
[**LO** - Dokshitzer; Gribov, Lipatov; Altarelli, Parisi; 1977]
[**NLO** - Floratos, Ross, Sachrajda; Gonzalez-Arroyo, Lopez, Yndurain; Curci, Furmanski, Petronzio, 1981]
[**NNLO** - Moch, Vermaseren, Vogt; 2004]

Why care about PDFs (and their uncertainties)?



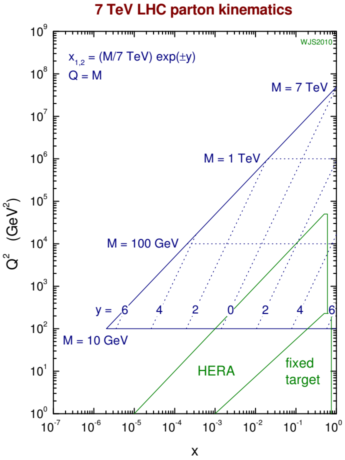
Why care about PDFs (and their uncertainties)?

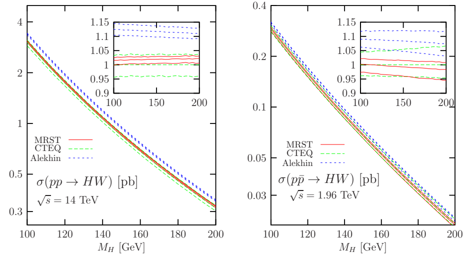




[A. Djouadi and S. Ferrag, hep-ph/0310209]

Why care about PDFs (and their uncertainties)?





[A. Djouadi and S. Ferrag, hep-ph/0310209]

Why care about PDFs (and their uncertainties)?

- Errors on PDFs are in some cases the **dominating theoretical error** on precision observables

Ex. $\sigma(Z^0)$ at the LHC: $\delta_{PDF} \sim 3\%$, $\delta_{NNLO} \sim 2\%$

[J. Campbell, J. Huston and J. Stirling, (2007)]

Why care about PDFs (and their uncertainties)?

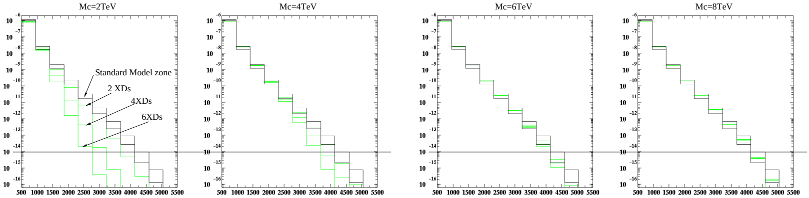
- Errors on PDFs are in some cases the **dominating theoretical error** on precision observables

Ex. $\sigma(Z^0)$ at the LHC: $\delta_{PDF} \sim 3\%$, $\delta_{NNLO} \sim 2\%$

[J. Campbell, J. Huston and J. Stirling, (2007)]

- Errors on PDFs might **reduce sensitivity to New Physics**

Ex. Extra Dimensions discovery in dijet cross section at the LHC:



[S. Ferrag (ATLAS), hep-ph/0407303]

Problem

Faithful estimation of errors on PDFs

- Single quantity: **1-σ error**
- Multiple quantities: **1-σ contours**
- Function: need an **"error band" in the space of functions**
  (*i.e.* the probability density $\mathcal{P}[f]$ in the space of functions $f(x)$)

Expectation values are **Functional integrals**

$$\langle \mathcal{F}[f(x)] \rangle = \int \mathcal{D}f \mathcal{F}[f(x)] \mathcal{P}[f(x)]$$

Problem

Faithful estimation of errors on PDFs

- Single quantity: **1-σ error**
- Multiple quantities: **1-σ contours**
- Function: need an **"error band" in the space of functions**
  (*i.e.* the probability density $\mathcal{P}[f]$ in the space of functions $f(x)$)

Expectation values are **Functional integrals**

$$\langle \mathcal{F}[f(x)] \rangle = \int \mathcal{D}f \mathcal{F}[f(x)] \mathcal{P}[f(x)]$$

Determine a function from a finite set of data points

Solution

Standard Approach

- Introduce a simple functional form with enough free parameters

$$q(x, Q^2) = x^{\alpha}(1 - x)^{\beta} P(x; \lambda_1, ..., \lambda_n).$$

- Fit parameters minimizing χ^2.

Solution

Standard Approach

- Introduce a simple functional form with enough free parameters

$$q(x, Q^2) = x^{\alpha}(1-x)^{\beta}P(x;\lambda_1, ..., \lambda_n).$$

- Fit parameters minimizing χ^2.

Open problems:

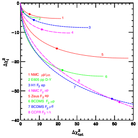
- **Error propagation** from data to parameters and from parameters to observables is **not trivial**.
- **Theoretical bias** due to the chosen **parametrization** is difficult to assess.

Shortcomings of the Standard approach

What is the meaning of a one-σ uncertainty?

- Standard $\Delta\chi^2 = 1$ criterion is **too restrictive** to account for large discrepancies among experiments in a **global fit**.

  [Collins & Pumplin, 2001]

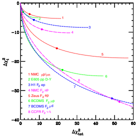


Shortcomings of the Standard approach

What is the meaning of a one-σ uncertainty?

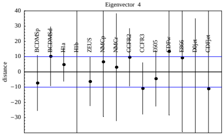
- Standard $\Delta\chi^2 = 1$ criterion is **too restrictive** to account for large discrepancies among experiments in a **global fit**.

  [Collins & Pumplin, 2001]



- Introduce a **TOLERANCE** criterion, i.e. take the envelope of uncertainties of experiments to determine the $\Delta\chi^2$ to use for the global fit (CTEQ).

  [Tung et al., 2006]

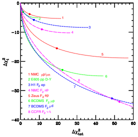


Shortcomings of the Standard approach

What is the meaning of a one-σ uncertainty?

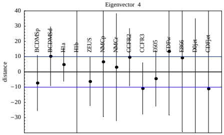
- Standard $\Delta\chi^2 = 1$ criterion is **too restrictive** to account for large discrepancies among experiments in a **global fit**.

  [Collins & Pumplin, 2001]



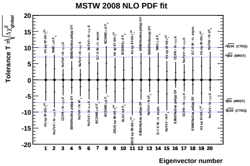
- Introduce a **TOLERANCE** criterion, i.e. take the envelope of uncertainties of experiments to determine the $\Delta\chi^2$ to use for the global fit (CTEQ).

  [Tung et al., 2006]



- Make it **DYNAMICAL**, i.e. determine $\Delta\chi^2$ separately for each hessian eigenvector (MSTW).

  [Thorne et. al, 2008]

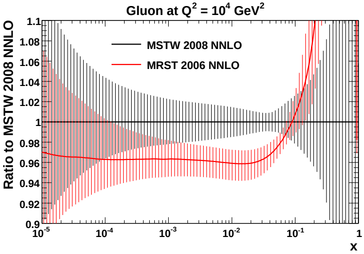


Shortcomings of the standard approach

What determines PDF uncertainties?

- **Uncertainties** in standard fits often **increase when adding data** (i.e. when adding information) even if they are compatible with the old data.
- **Reason**: need change the parametriztion in order to accomodate the new data.

Smaller high-x gluon (and slightly smaller α_S) results in larger small-x gluon – now shown at NNLO.



Larger small-x uncertainty due to extrat free parameter.

[R. Thorne, PDF4LHC]

THE NNPDF METHODOLOGY

[R. D. Ball, V. Bertone, F. Cerutti, L. Del Debbio, S. Forte, J. I. Latorre, A. Piccione, J. Rojo, M. Ubiali and AG]

NNPDF Methodology

Main Ingredients

- **Monte Carlo** determination of errors
  - No need to rely on linear propagation of errors
  - Possibility to test for the impact of non gaussianly distributed errors
  - Possibility to test for non-gaussian behaviour in fitted PDFs ($1 - \sigma$ vs. 68% CL)
- **Neural Networks**
  - Provide an **unbiased** parametrization
- **Stopping based on Cross-Validation**
  - Ensures proper fitting avoiding overlearning

NNPDF Methodology

... in a Nutshell

- Generate N_{rep} **Monte-Carlo replicas** of the experimental data (sampling of the probability density in the space of data)
- Fit a set of Parton Distribution Functions on each replica (sampling of the probability density in the space of PDFs)
- **Expectation values** for observables are **Monte Carlo integrals**

$$\langle \mathcal{F}[f_i(x, Q^2)] \rangle = \frac{1}{N_{rep}} \sum_{k=1}^{N_{rep}} \mathcal{F}\left(f_i^{(net)(k)}(x, Q^2) \right)$$

... the same is true for errors, correlations, etc.

NNPDF Methodology

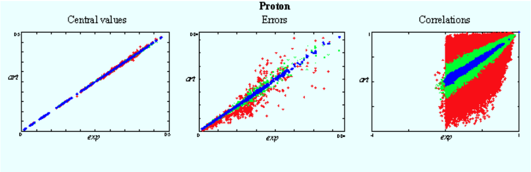
Monte Carlo replicas generation

- **Generate** artificial data according to distribution

$$O_i^{(art)\,(k)} = (1 + r_N^{(k)}\,\sigma_N)\left[O_i^{(exp)} + \sum_{p=1}^{N_{sys}} r_p^{(k)}\,\sigma_{i,p} + r_{i,s}^{(k)}\,\sigma_s^i\right]$$

where r_i are univariate (gaussianly distributed) random numbers

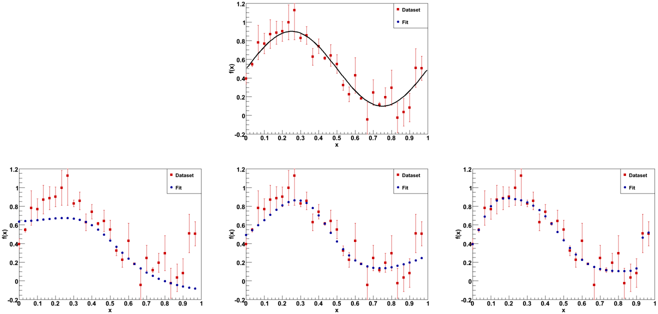
- **Validate** Monte Carlo replicas against experimental data (statistical estimators, faithful representation of errors, convergence rate increasing N_{rep})



- $\mathcal{O}(1000)$ replicas needed to reproduce correlations to percent accuracy

Proper Fitting avoiding Overlearning

Parametrization bias in a toy model



- Need a **redundant parametrization** to avoid parametrization bias.
- Need a way of **stopping the fit before overlearning** sets in to avoid fitting statistical noise.

Neural Networks

... a suitable basis of functions

- We use **Neural Networks** as **functions** to represent **PDFs at the starting scale**
- We employ **Multilayer Feed-Forward** Neural Networks trained using a **Genetic Algorithm**
- Activation determined by weights and thresholds

$$\xi_i = g\left(\sum_j \omega_{ij}\xi_j - \theta_i\right) , \qquad g(x) = \frac{1}{1 + e^{-\beta x}}$$

Neural Networks

... a suitable basis of functions

- We use **Neural Networks** as **functions** to represent **PDFs at the starting scale**
- We employ **Multilayer Feed-Forward** Neural Networks trained using a **Genetic Algorithm**
- Activation determined by weights and thresholds

$$\xi_i = g\left(\sum_j \omega_{ij}\xi_j - \theta_i\right), \qquad g(x) = \frac{1}{1+e^{-\beta x}}$$

Ex.: 1-2-1 NN:

$$\xi_1^{(3)}(\xi_1^{(1)}) = \frac{1}{1 + e^{\theta_1^{(3)} - \frac{\omega_{11}^{(2)}}{1+e^{\theta_1^{(2)} - \xi_1^{(1)}\omega_{11}^{(1)}}} - \frac{\omega_{12}^{(2)}}{1+e^{\theta_2^{(2)} - \xi_1^{(1)}\omega_{21}^{(1)}}}}}$$

Neural Networks

... a suitable basis of functions

- We use **Neural Networks** as **functions** to represent **PDFs at the starting scale**
- We employ **Multilayer Feed-Forward** Neural Networks trained using a **Genetic Algorithm**
- Activation determined by weights and thresholds

$$\xi_i = g\left(\sum_j \omega_{ij}\xi_j - \theta_i\right) , \qquad g(x) = \frac{1}{1+e^{-\beta x}}$$

Ex.: 1-2-1 NN:

$$\xi_1^{(3)}(\xi_1^{(1)}) = \frac{1}{1+e^{\theta_1^{(3)} - \frac{\omega_{11}^{(2)}}{1+e^{\theta_1^{(2)} - \xi_1^{(1)}\omega_{11}^{(1)}}} - \frac{\omega_{12}^{(2)}}{1+e^{\theta_2^{(2)} - \xi_1^{(1)}\omega_{21}^{(1)}}}}}$$

- They provide a parametrization which is **redundant** and **robust** against variations

Genetic Algorithm

1. Set network parameters randomly.
2. Make *clones* of the set of parameters.
3. Mutate each clone.
4. Evaluate χ^2 for all the clones.
5. Select the clone that has the lowest χ^2.
6. Back to 2, until stability in χ^2 is reached.

Genetic Algorithm

1. Set network parameters randomly.
2. Make *clones* of the set of parameters.
3. Mutate each clone.
4. Evaluate χ^2 for all the clones.
5. Select the clone that has the lowest χ^2.
6. Back to 2, until stability in χ^2 is reached.

Pros:

- Allows to minimize the fully correlated χ^2
- Explores the full parameter space, reducing the risk of being trapped in a local minimum

Cons:

- Slow convergence
- χ^2 decreases monotonically - need to find a suitable stopping criterion

Neural Networks

Stopping criterion

Stopping criterion based on Training-Validation separation

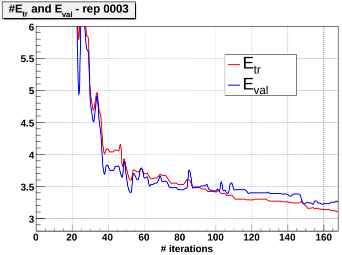
- Divide the data in two sets: Training and Validation
- Minimize the χ^2 of the data in the Training set
- Compute the χ^2 for the data in the Validation set
- When validation χ^2 stops decreasing, **STOP** the fit

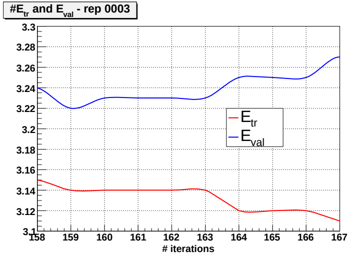
Neural Networks

Stopping criterion

Stopping criterion based on Training-Validation separation

- Divide the data in two sets: Training and Validation
- Minimize the χ^2 of the data in the Training set
- Compute the χ^2 for the data in the Validation set
- When validation χ^2 stops decreasing, **STOP** the fit

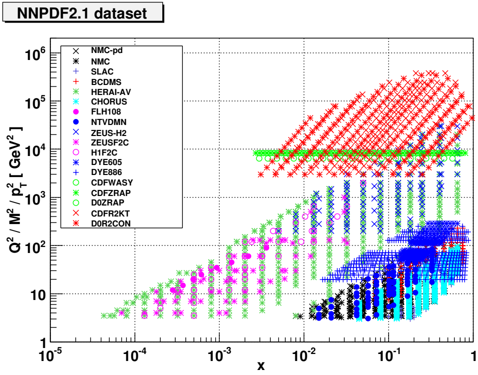




RESULTS

NNPDF 2.1

A family of global fits ...



3415 data points (NLO fit)

(3408 - LO and 3473 - NNLO)

[R. D. Ball et. al, arXiv:1101.1300] - **NLO**
[R. D. Ball et. al, arXiv:1107.2652] - **LO/NNLO**

OBS	***Data set***
Deep Inelastic Scattering	
F_2^d / F_2^p	NMC-pd
F_2^p	NMC, SLAC, BCDMS
F_2^d	SLAC, BCDMS
$\sigma_{NC}^{\pm}$	HERA-I, ZEUS (HERA-II)
$\sigma_{CC}^{\pm}$	HERA-I, ZEUS (HERA-II)
F_L	H1
$\sigma_\nu, \sigma_{\bar{\nu}}$	CHORUS
dimuon prod.	NuTeV
F_2^c	ZEUS, H1
Drell-Yan & Vector Boson prod.	
$d\sigma^{\rm DY}/dM^2 dy$	E605
$d\sigma^{\rm DY}/dM^2 dx_F$	E866
W asymm.	CDF
Z rap. distr.	D0/CDF
Inclusive jet prod.	
Incl. $\sigma^{\rm (jet)}$	CDF (k_T) - Run II
Incl. $\sigma^{\rm (jet)}$	D0 (cone) - Run II

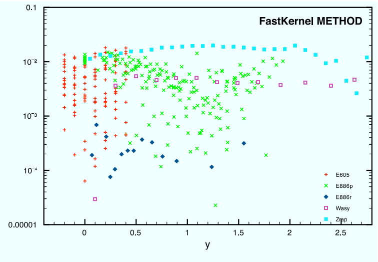
NNPDF 2.x

Inclusion of Higher Order corrections - *FastKernel*

- NLO computation of hadronic observables too slow for parton global fits.
- MSTW08 and CTEQ include Drell-Yan NLO as (local) K factors rescaling the LO cross section
- K-factor depends on PDFs and it is not always a good approximation.

* NNPDF2.0 includes full NLO calculation of hadronic observables.
* Use available fastNLO interface for jet inclusive cross-sections.[hep-ph/0609285]
* Built up our own **FastKernel** computation of DY observables.

$$\int_{x_{0,1}}^{1} dx_1 \int_{x_{0,2}}^{1} dx_2\, f_a(x_1) f_b(x_2) C^{ab}(x_1, x_2) \rightarrow \sum_{\alpha,\beta=1}^{N_x} f_a(x_{1,\alpha}) f_b(x_{2,\beta}) \int_{x_{0,1}}^{1} dx_1 \int_{x_{0,2}}^{1} dx_2\, \mathcal{I}^{(\alpha,\beta)}(x_1, x_2) C^{ab}(x_1, x_2)$$



- Both PDFs evolution and double convolution sped up by
  - Use high-orders polynomial interpolation
  - Precompute all Green Functions

A truly NLO analysis

NNPDF 2.1

Heavy Flavour treatment - FONLL

- We adopt the **FONLL** General Mass-Variable Flavour Number Scheme

  [M. Cacciari, M. Greco and P. Nason, (1998)]

  [S. Forte, P. Nason E. Laenen and J. Rojo, (2010)]

- FONLL gives a prescription to **combine FFN** (Massive) and **ZM-VFN** (Massless) computations, at any given order, **avoiding double counting**.

- With results available three implementations of FONLL are possibile:
  - **FONLL-A**: $\mathcal{O}(\alpha_s)$ Massless + $\mathcal{O}(\alpha_s)$ Massive
  - **FONLL-B**: $\mathcal{O}(\alpha_s)$ Massless + $\mathcal{O}(\alpha_s^2)$ Massive
  - **FONLL-C**: $\mathcal{O}(\alpha_s^2)$ Massless + $\mathcal{O}(\alpha_s^2)$ Massive

- **Fixed Flavour Number Scheme** (3-, 4-, 5-) fits **available**.

NNPDF 2.1

Parametrization

Parton Distributions Combination		NN architechture
Singlet ($\Sigma(x)$)	$\Longrightarrow$	2-5-3-1 (37 pars)
Gluon ($g(x)$)	$\Longrightarrow$	2-5-3-1 (37 pars)
Total valence ($V(x) \equiv u_V(x) + d_V(x)$)	$\Longrightarrow$	2-5-3-1 (37 pars)
Non-singlet triplet ($T_3(x)$)	$\Longrightarrow$	2-5-3-1 (37 pars)
Sea asymmetry ($\Delta_S(x) \equiv \bar{d}(x) - \bar{u}(x)$)	$\Longrightarrow$	2-5-3-1 (37 pars)
Total Strangeness ($s^+(x) \equiv (s(x) + \bar{s}(x))/2$)	$\Longrightarrow$	2-5-3-1 (37 pars)
Strange valence ($s^-(x) \equiv (s(x) - \bar{s}(x))/2$)	$\Longrightarrow$	2-5-3-1 (37 pars)

259 parameters

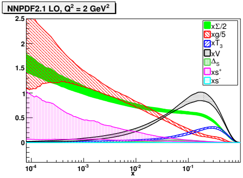
Standard fits have $\sim$ 25 parameters in total

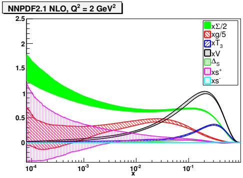
No change in the parametrization since NNPDF1.2 ... despite substantial **enlargement of the dataset**.

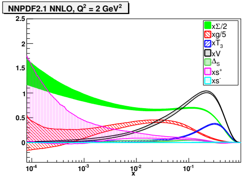
The present status of PDF fits

PDFs ... a family portrait

- At the starting scale (2 GeV^2) ...



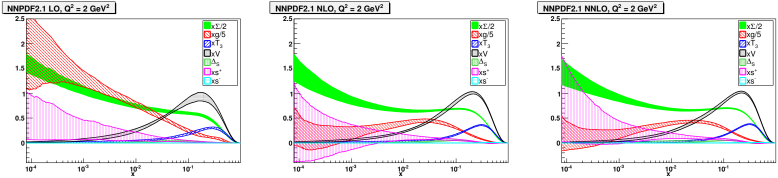




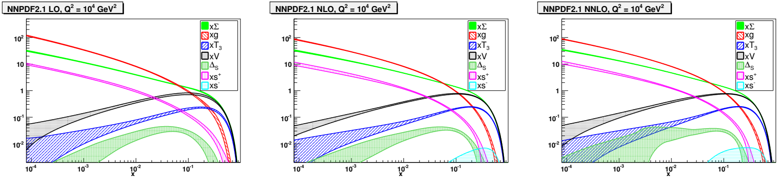
The present status of PDF fits

PDFs ... a family portrait

- At the starting scale (2 GeV^2) ...



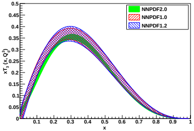
- ... and at the typical EW scale (100 GeV^2)



NNPDF 2.1

Partons - A couple of upshots

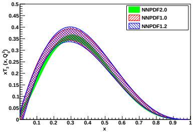
- **Reduction of uncertainties** with respect to older NNPDF sets due to **inclusion of new data**.



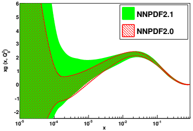
NNPDF 2.1

Partons - A couple of upshots

- **Reduction of uncertainties** with respect to older NNPDF sets due to **inclusion of new data**.



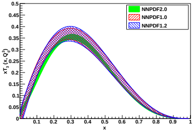
- When **uncertainties increase** we know it is **not a parametrzation effect**.



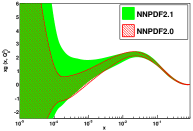
NNPDF 2.1

Partons - A couple of upshots

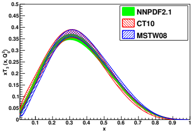
- **Reduction of uncertainties** with respect to older NNPDF sets due to **inclusion of new data**.



- When **uncertainties increase** we know it is **not a parametrzation effect**.



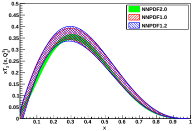
- **Uncertainties** on PDFs have **size comparable** to those obtained by other groups in kinematic regions where there are significant **contraints from data** ...

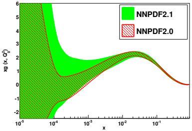


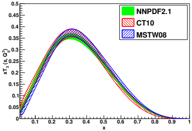
NNPDF 2.1

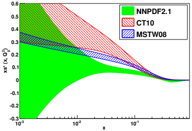
Partons - A couple of upshots

- **Reduction of uncertainties** with respect to older NNPDF sets due to **inclusion of new data**.
- When **uncertainties increase** we know it is **not a parametrzation effect**.
- **Uncertainties** on PDFs have **size comparable** to those obtained by other groups in kinematic regions where there are significant **contraints from data** ...
- ... but still retain **unbiasedness** in kinematic regions where there are little or **no experimental constraints**.









The present status of PDF fits

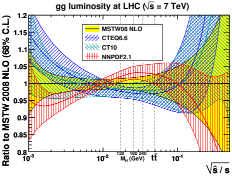
The PDF sets Matrix

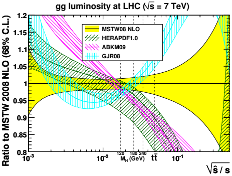
	Dataset	Pert. Order	Heavy Flavours	α_S	Param.	Uncert.
ABKM09	DIS (FT, HERA) Drell-Yan (FT)	NLO NNLO	FFN (BMSN)	fitted	6 indep. PDF Polynomial (25 param.)	Hessian ($\Delta\chi^2 = 1$)
CT10	DIS (FT, HERA) Drell-Yan (FT, Tev) Jets (Tevatron)	LO NLO	GM-VFNS (S-ACOT)	external var. avail.	6 indep. PDF Polynomial (26 param.)	Hessian ($\Delta\chi^2 = 100$)
JR09	DIS (FT, HERA) Drell-Yan (FT) Jets (Tevatron)	NLO NNLO	FFN VFN	fitted	5 indep. PDF Polyinom. (15 param.)	Hessian ($\Delta\chi^2 = 1$)
HERAPDF1.5	DIS (HERA)	NLO NNLO	GM-VFNS (TR)	external var. avail.	5 indep. PDF Polnom. (14 param.)	Hessian ($\Delta\chi^2 = 1$)
MSTW08	DIS (FT, HERA) Drell-Yan (FT, Tev) Jets (HERA, Tev)	LO NLO NNLO	GM-VFNS (TR)	fitted	7 indep. PDF Polynom. (20 param.)	Hessian ($\Delta\chi^2 \sim 25$)
NNPDF2.1	DIS (FT, HERA) Drell-Yan (FT, Tev) Jets (Tevatron)	LO NLO NNLO	GM-VFNS (FONLL)	external var. avail	7 indep. PDF Neural Netw. (259 param.)	Monte Carlo

- When trying to understand **differences between PDF** sets it is useful to look at **parton luminosities**

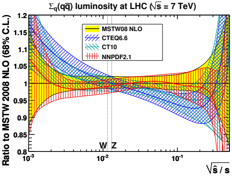
$$\Phi_{ij}(M_X^2) = \frac{1}{s}\int_\tau^1 \frac{x_1}{x_1} f_i(x_1, M_X^2) f_j(\tau/x_1, M_X^2)$$

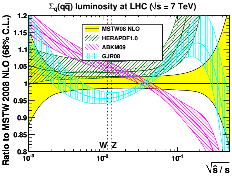
gg:





$q\bar{q}$:

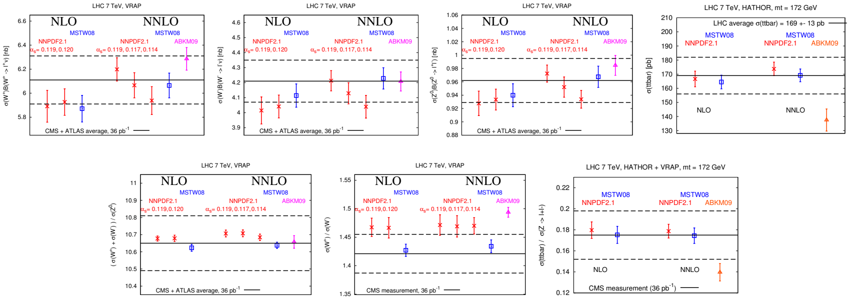




The present status of PDF fits

Comparisons to LHC data

- Predictions for **LHC Standard Candles** compared to **LHC data**



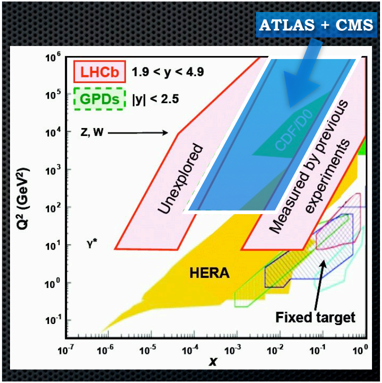
- LHC data will soon be precise enough to distinguish between different predictions.

LHC4PDFs

W lepton asymmetry data at the LHC

$$A_W^l = \frac{\sigma(pp \to W^+ \to l^+\nu_l) - \sigma(pp \to W^- \to l^-\bar{\nu}_l)}{\sigma(pp \to W^+ \to l^+\nu_l) + \sigma(pp \to W^+ \to l^-\bar{\nu}_l)}$$

- ATLAS: muon charge asymmetry ($31 pb^{-1}$) [ArXiv:1103:2929]
- CMS: muon charge asymmetry ($36 pb^{-1}$) [ArXiv:1103:3470]

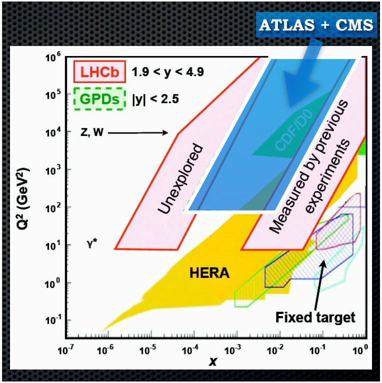


LHC4PDFs

W lepton asymmetry data at the LHC

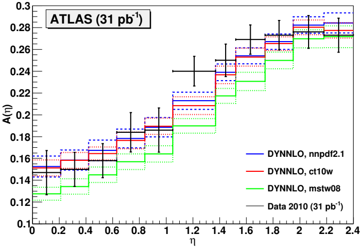
$$A_W^l \sim \frac{u(x_1,M_W^2)\bar{d}(x_2,M_W^2) - d(x_1,M_W^2)\bar{u}(x_2,M_W^2)}{u(x_1,M_W^2)\bar{d}(x_2,M_W^2) + d(x_1,M_W^2)\bar{u}(x_2,M_W^2)}$$

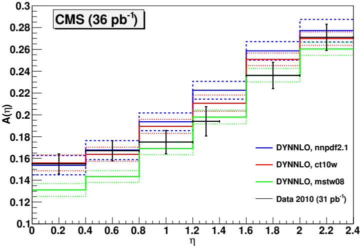
- ATLAS: muon charge asymmetry (31pb^{-1}) [ArXiv:1103:2929]
- CMS: muon charge asymmetry (36pb^{-1}) [ArXiv:1103:3470]



LHC4PDFs

The W lepton asymmetry data at LHC



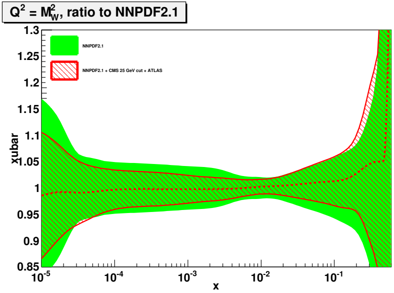


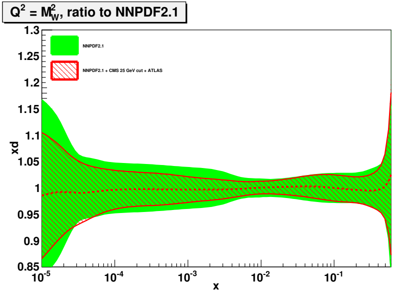
χ^2/d.o.f.	**NNPDF2.1**	**CT10w**	**MSTW08**
ATLAS	0.7	0.8	3.2
CMS e^- $p_T > 25$ GeV	1.9	0.8	2.4
CMS e^- $p_T > 30$ GeV	1.7	1.2	2.5
CMS μ $p_T > 25$ GeV	1.3	0.5	1.1
CMS μ $p_T > 30$ GeV	0.8	0.6	1.3

Theory predictions computed using DYNNLO at NLO

[ArXiv:0903.2120]

Inclusion of the LHC W lepton asymmetry data





- ATLAS and CMS data compatible with data included in global analysis
- The provide important constraint to PDFs in the small medium-x region
- Significant uncertainty reduction

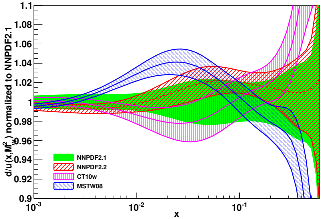
ATLAS
$N_{\rm eff} = 928$, $\chi^2_{\rm d.o.f.}$: $0.69 \to 0.65$
CMS ($p_T^l > 25$GeV)
$N_{\rm eff} = 554$, $\chi^2_{\rm d.o.f.}$: $1.41 \to 0.74$
CMS ($p_T^l > 30$GeV)
$N_{\rm eff} = 717$, $\chi^2_{\rm d.o.f.}$: $0.98 \to 0.72$

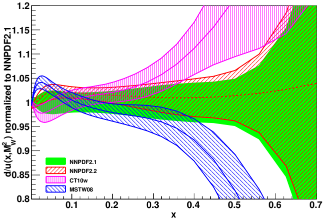
LHC4PDFs

Combining Tevatron and LHC data

LOG scale

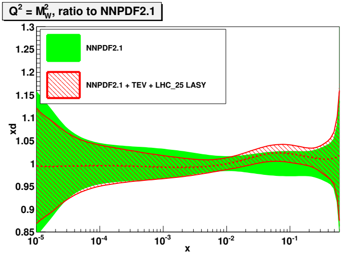
LIN scale





ATLAS+CMS(25)+D0$_\mu$+D0$_e$(20)

$N_{\rm eff}$ = 196, $\chi^2_{\rm d.o.f.}$: 2.18 $\rightarrow$ 0.86



- Uncertainty reduction medium-small x and shift in central value driven by the LHC data
- Uncertainty reduction at medium-large x driven by Tevatron data

LHC4PDFs

... the data we would love to have from the LHC

- Medium- and large-x **gluon**
  - Prompt photons
  - Inclusive Jets
  - t-quark distributions ($p_{\perp}$, y) (?)
- **Light flavour separation** at medium- & small-x
  - Low-mass Drell-Yan
  - High-mass W prduction
  - Z rapidity distribution
  - W(+jets) asymmetry
- **Strangeness & Heavy Flavours**
  - $W + c$
  - $Z + c$, $\gamma + c$
  - $Z + b$

LOOKING FOR DEVIATIONS FROM NLO DGLAP EVOLUTION

[F. Caola, S. Forte & J. Rojo, arXiv:0910.3143]

Looking for deviations from DGLAP evolution

Deviations from (NLO) DGLAP are **expected at at small-x and Q^2**

- Several **possible sources**: NNLO, resummation, saturation ...
- Possible **deviations are small**: NLO fits work well!
- **Difficult to single out** deviations: might be absorbed by deformation of fitted PDFs

Testing such an hypotesis requires

- Very precise, extensive **dataset**: HERA-I combined dataset
- A set of PDFs with a reliable error estimation: NNPDF

Looking for deviations from DGLAP evolution

The General Strategy

- Perform a **DGLAP fit** to determine PDFs in a **"safe region"**
- **Back-evolve** them using DGLAP to the **"would-be-unsafe" region** and compute observables
- **Compare** predictions **to data** $\Longrightarrow$ **systematic discrepancy** indicates **deviations from DGLAP**

Looking for deviations from DGLAP evolution

The General Strategy

- Perform a **DGLAP fit** to determine PDFs in a **"safe region"**
- **Back-evolve** them using DGLAP to the **"would-be-unsafe" region** and compute observables
- **Compare** predictions **to data** $\Longrightarrow$ **systematic discrepancy** indicates **deviations from DGLAP**

Caution

Deviations from DGLAP, if there, are small
(NLO DGLAP fits give good description of available data)

- **Reliable PDF uncertainties** determination is crucial
- **Refined statistical tools**: establish statistical significance of the discrepancy

Looking for deviations from DGLAP evolution

Define the "safe region"

The Goal

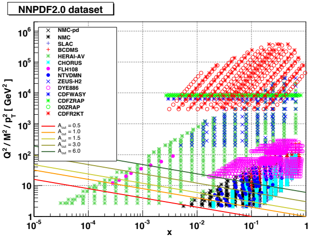
- Remove low-x, low-Q^2 region from the fit
- Retain enough experimental information to constrain PDFs

The Suggestion

- Keep data in the region

$$Q^2 > A_{\rm cut} x^{-0.3}$$

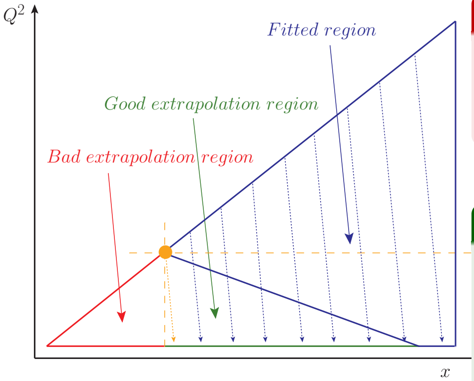
- Theoretically appealing (saturation inspired ...)
- Keep large-x data



Looking for deviations from DGLAP evolution

Define the "would-be-unsafe region"

Think of the causal structure of DGLAP evolution



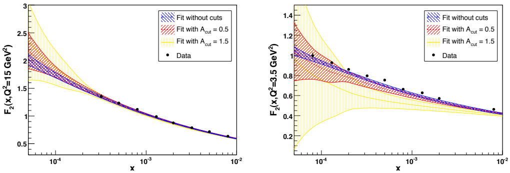
Bad extrapolation region

Evolution affected by information in the unfitted region → DGLAP prediction meaningless

Good extrapolation region

Evolution only depends on information in the fitted region
↓
Meaningful DGLAP prediction

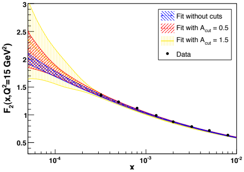
- We can have a look at the structure functions F_2 itself

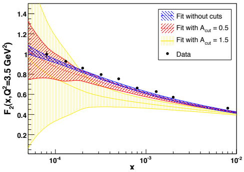


Looking for deviations from DGLAP evolution

Results

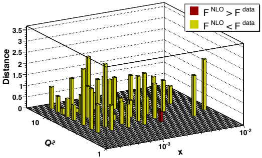
- We can have a look at the structure functions F_2 itself

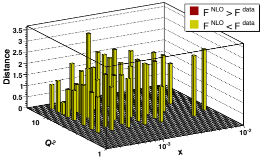




- Or, for a more quantitative insght, at the distances between the predictions and the data

$$d(x, Q^2) = \frac{F_2^{\rm dat} - F_2^{\rm fit}}{\sqrt{\sigma^2_{\rm dat} + \sigma^2_{\rm fit}}}$$





Looking for deviations from DGLAP evolution

Comments

- This analysis provides a **model-independent indication** for deviations from (NLO) DGLAP evolution
- Possible explanations:
  - Higher Order effects (NNLO)
  - Heavy Flavour Mass contributions
  - Perturbative resummation
  - Parton saturation
  - ....
- **Establishing the source** of the deviations requires a specific, **hypotesis-dependent study** (e.g. inclusion of resummation effect in fits)
- We have the tools to test some of these explanations, now!

Conclusions

- A **reliable** estimation of **PDF uncertainties** is crucial in order to exploit the full physics potential of the LHC experiments.
- The **NNPDF2.1** family of PDF sets fulfills the requirement of an ideal parton densities set for precision phenomenology at the LHC
  - it is based on a comprehensive global dataset,
  - it is (almost) free of parametrization bias,
  - it is provided with a reliable, statistically meaningful estimation of uncertainties,
  - it includes higher order corrections (almost) without resorting to K-factor approximations,
  - it includes a consistent treatment of heavy quark effects,
  - it is available for a variety of values of α_s and quark masses.
- The **NNPDF parton sets** provide an **ideal tool** to systematically **test for deviations from DGLAP** evolution at small-x and Q^2