NNPDF at LHC

Valerio Bertone

Albert-Ludwigs-Universität Freiburg,
Physikalisches Institut

Standard and novel QCD phenomena at hadron colliders

May 30$^{\rm th}$ - June 2$^{\rm nd}$ 2011, ECT*, Trento

On behalf of the NNPDF Collaboration:
R.D.Ball[1], V.B.[3], F.Cerutti[4], L.Del Debbio[1], S.Forte[2], A.Guffanti[3], J.I.Latorre[4], J. Rojo[2], M.Ubiali[5]

[1] PPT Group, School of Physics, University of Edinburgh
[2] Dipartimento di Fisica, Università di Milano
[3] Physikalisches Institut, Albert-Ludwigs-Universität Freiburg
[4] Departament d'Estructura i Constituents de la Matèria, Universitat de Barcelona
[5] Institut für Theoretische Teilchenphysik und Kosmologie, RWTH Aachen University

Outline

Part I

Introduction

- Extraction of a set of functions with errors from a set of data points.
- We need an error band, i.e. a **probability density** $\mathcal{P}[\mathbf{f}(\mathbf{x})]$ in the space of PDFs :

$$\langle \mathcal{O} \rangle = \int \mathcal{D}f\, \mathcal{P}[f]\mathcal{O}[f]$$

$$\sigma^2_{\mathcal{O}} = \int \mathcal{D}f\, \mathcal{P}[f]\left(\mathcal{O}[f] - \langle \mathcal{O} \rangle\right)^2$$

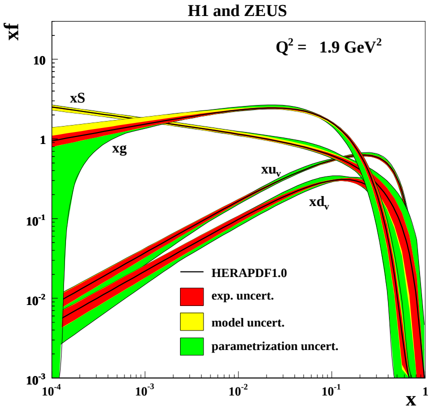
Standard approach:

1. Choose a specific functional form:
$$q_i(x, Q_0^2) = A_i x^{b_i}(1-x)^{c_i}(1+\ldots).$$
2. Determine best-fit values of parameters which define the functions.
3. Errors determined via gaussian error propagation and, in some cases, using a tolerance criterion.

But:

1. How can we know if the **parametrization** is flexible enough?
2. What is the **error** associated to any particular choice?

Reference: [arXiv:0911.0884]



- HERAPDF analysis: computation of parametrization uncertainty on top of the model (initial scale and heavy quark mass effects) and experimental uncertainty.
- In many regions the **parametrization uncertainty** is **dominating**.

Part II

The NNPDF Methodology

- Generate N **Monte Carlo replicas** of the experimental data
  (sampling of the probability density in the space of data).

- Fit a set of Parton Distribution Functions with a set of **Neural Networks** on each replica
  (sampling of the probability density in the space of PDFs).

- **Expectation values** for observables are **Monte Carlo integrals**:

$$\langle \mathcal{O}[f] \rangle = \frac{1}{N} \sum_{k=1}^{N} \mathcal{O}[f_k]$$

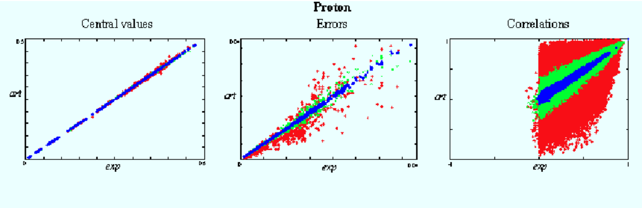
...the same is true for errors, correlations, etc.

- **Generation** of artificial Monte Carlo data according to the distribution:

$$\mathcal{O}_i^{(art)(k)} = (1 + r_{norm}^{(k)} \sigma_{norm}) \left[\mathcal{O}_i^{(exp)} + r_{stat}^{(k)} \sigma_{stat} + \sum_{p=1}^{N_{sys}} r_{sys,p}^{(k)} \sigma_{sys,p} \right]$$

where $r_i^{(k)}$ are univariate (gaussianly distributed) random numbers.

- **Validation** of the Monte Carlo replicas against experimental data .



Red points: 10 replicas
Green points: 100 replicas
Blue points: 1000 replicas

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

- **Unbiased basis** of functions parameterized by a very large and **redundant set of parameters** $\Rightarrow$ **Neural Networks**.
- Each one of the **7 independent PDFs**:

Gluon	$g(x)$
Singlet	$\Sigma(x) = \sum_q (q(x) + \overline{q}(x))$
Valence	$V(x) = \sum_q (q(x) - \overline{q}(x))$
Triplet	$T_3(x) = (u(x) + \overline{u}(x)) - (u(x) + \overline{u}(x))$
Sea asymmetry	$\Delta_S(x) = \overline{d}(x) - \overline{u}(x)$
Total Strangeness	$s^+(x) = s(x) + \overline{s}(x)$
Strange Valence	$s^-(x) = s(x) - \overline{s}(x)$

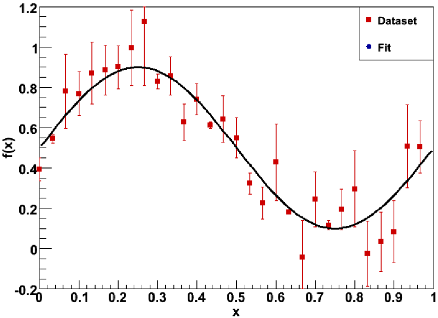
is parametrized at the initial scale $Q_0^2 = 2\ \mathrm{GeV}^2$ by an individual Neural Network having architecture 2-5-3-1 $\Rightarrow$ **37 parameters**.

259 parameters
Standard fits have $\sim$ 25 parameters in total

Drawback:

A redundant parametrization might fit not only to physical behavior but also to random statistical fluctuations of data.

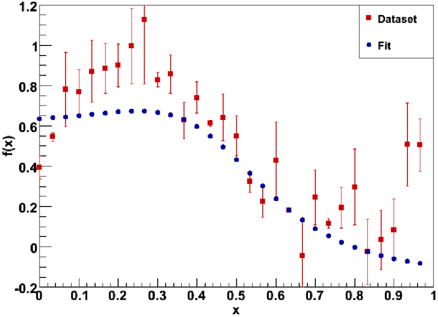
UNDERLYING PHYSICAL LAW



Drawback:

A redundant parametrization might fit not only to physical behavior but also to random statistical fluctuations of data.

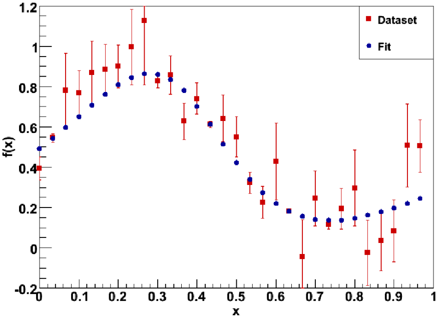
UNDERLEARNING



Drawback:

A redundant parametrization might fit not only to physical behavior but also to random statistical fluctuations of data.

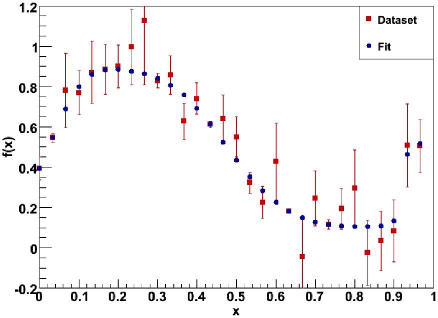
PROPER LEARNING



Drawback:

A redundant parametrization might fit not only to physical behavior but also to random statistical fluctuations of data.

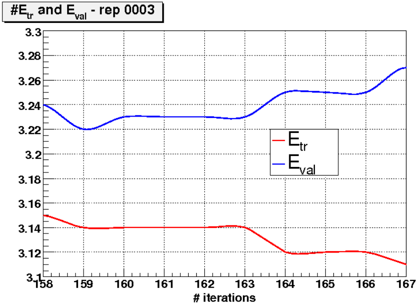
OVERLEARNING



Need for a suitable stopping criterion!

Cross-validation method:

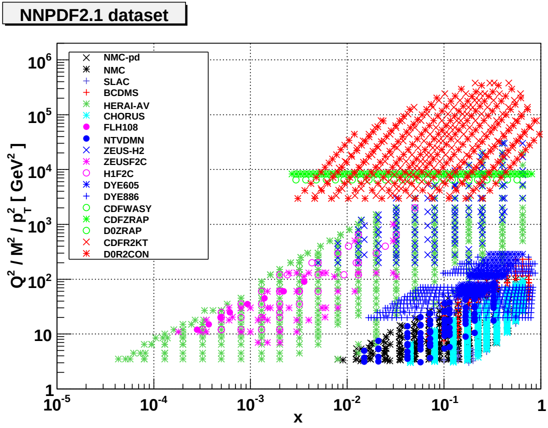
- Divide data in two sets: training and validation (for each experiment).
- Random division for each replica (tipically $f_t = f_v = 0.5$).
- Minimisation performed only on the training set. Meantime, the validation χ^2 is monitored.
- When the training χ^2 still decreases while the validation χ^2 stops decreasing → STOP.



Part III

NNPDF2.1

Reference: [Nucl. Phys. B849(2011)296]



3415 data points

For comparison, MSTW08 includes 2699 data points

OBS	Data sets
F_2^p	NMC,SLAC,BDCMS
F_2^d	SLAC,BCDMS
F_2^d / F_2^p	NMC-pd
σ_{NC}	HERA-I AV, ZEUS-H2
σ_{CC}	HERA-I AV, ZEUS-H2
F_L	H1
σ_ν , $\sigma_{\bar{\nu}}$	CHORUS
dimuon prod.	NuTeV
$d\sigma^{\mathrm{DY}} / dM^2 dy$	E605
$d\sigma^{\mathrm{DY}} / dM^2 dx_F$	E886
W asymmetry	CDF
Z rap. distr.	CDF,D0
incl. $\sigma^{(\mathrm{jet})}$	D0(cone) Run II
incl. $\sigma^{(\mathrm{jet})}$	CDF(k_T) Run II
F_2^c	**ZEUS (99,03,08,09)**
F_2^c	**H1 (01,09,10)**

- **HQ mass effects implemented**.
- $W^2 = Q^2(1-x)/x > 12.5$ GeV^2
- $\mathbf{Q^2 > 3}$ **GeV^2** + further cuts on F_2^c:
  - no data below the charm mass.
- No cuts on hadronic data.

Reference: S.Forte, E.Laenen, P.Nason,J. Rojo [ArXiv:1001.2312]

- HQ mass effects: NNPDF2.1 implements the so-called **FONLL method** which is a prescription for combining **FFN** (Massive) and **ZM-VFN** (Massless) computations, at any given order, **avoiding double counting**.

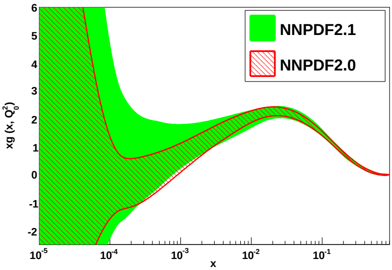
- Definition of FONLL structure function:

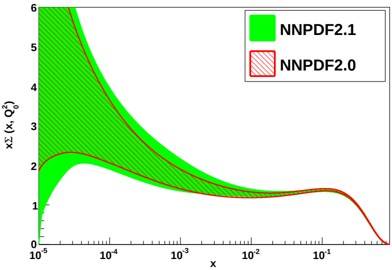
$$F^{\mathrm{FONLL}}(x,Q^2) = F^{(n_f+1)}(x,Q^2) + F^{(n_f)}(x,Q^2) - F^{(n_f,0)}(x,Q^2)$$

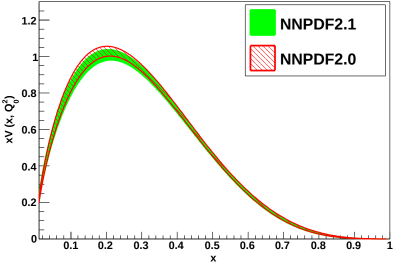
$F^{(n_f+1)}(x,Q^2)$: massless-scheme structure function
$F^{(n_f)}(x,Q^2)$: massive-scheme structure function $\Rightarrow$ inclusion of the mass suppressed terms
$F^{(n_f,0)}(x,Q^2)$: massless limit of $F^{(n_f)}(x,Q^2)$ $\Rightarrow$ subtraction of the double counting terms

- At the moment, three possibilities:
  - FONLL-A: $\mathcal{O}(\alpha_s)$ PDFs + $\mathcal{O}(\alpha_s)$ coefficient functions,
  - FONLL-B: $\mathcal{O}(\alpha_s)$ PDFs + $\mathcal{O}(\alpha_s^2)$ coefficient functions,
  - FONLL-C: $\mathcal{O}(\alpha_s^2)$ PDFs + $\mathcal{O}(\alpha_s^2)$ coefficient functions.

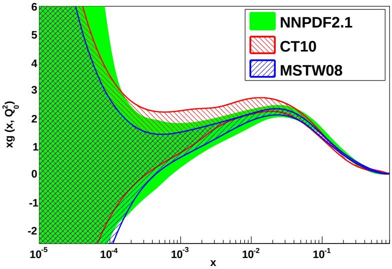
- **NNPDF2.1** is a **NLO** fit obtained using **FONLL-A**.

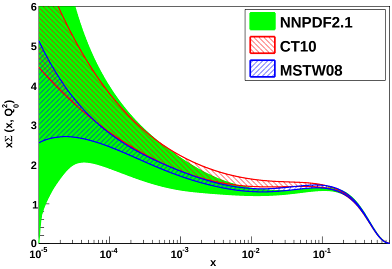


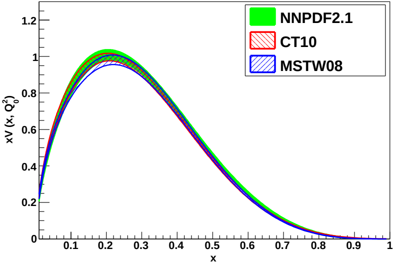




- Small and medium-x gluon and singlet are the most affected due to:
  - GM scheme,
  - new kinematic cuts.
- Effect is within 1σ.
- Other PDFs almost unaffected.

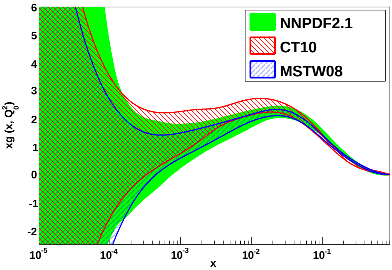


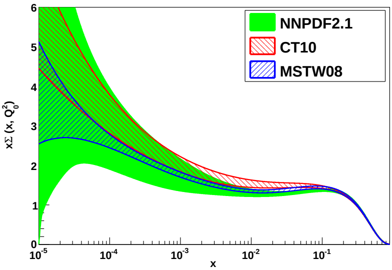


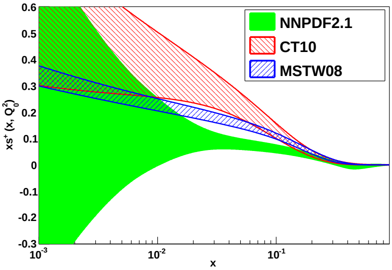


- General agreement of the gluon in the small and medium-x.
- Good agreement in the valence sector.

Parton Ditribution Functions: NNPDF2.1 vs. MSTW08 and CT10

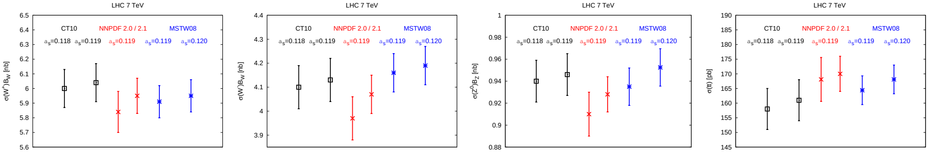


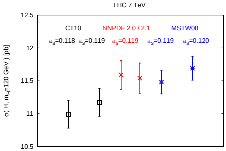




- General agreement of the gluon in the small and medium-x improved due to:
- Big difference in the strange distribution:
  - less flexible parametrization adopted by CT10 and MSTW08,
  - solution of the NuTeV anomaly (see below).

LHC standard candles:



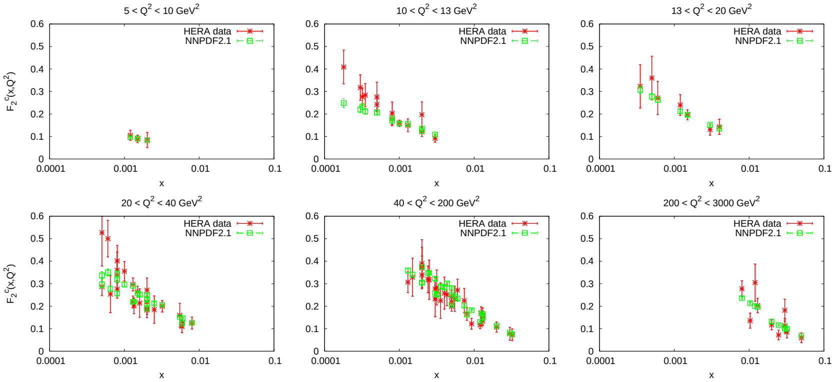


- The use of a common value for α_s leads to better agreement between collaborations.
- NNPDF2.1 and MSTW08 always in good agreement.
- Slightly worse agreement between NNPDF2.1 and CT10, especially for observables sensitive to the gluon distribution, like $t\bar{t}$ and Higgs.
- Differences between NNPDF2.1 and NNPDF2.0 mostly due to the heavy quark effects:
  - $W^{\pm}$ and Z predictions within 1-σ,
  - $t\bar{t}$ and Higgs predictions unchanged.

Data in the smallest x and Q^2 bins affected by **large $\mathcal{O}(\alpha_s^2)$ corrections** $\Rightarrow$ further kinematic cuts applied to HERA F_2^c:

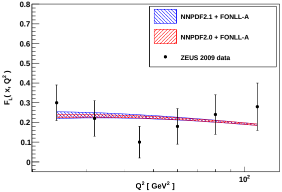
- $Q^2 > 4\ \text{GeV}^2$,
- $Q^2 > 10\ \text{GeV}^2$ for $x \leq 10^{-3}$

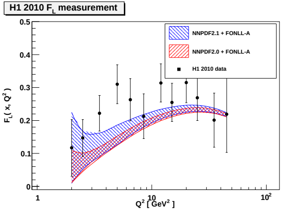
Comparison of NNPDF2.1 predictions for F_2^c with present data



Rather good agreement!

Comparison of NNPDF2.1 predictions for F_L with present data

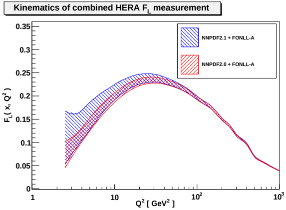


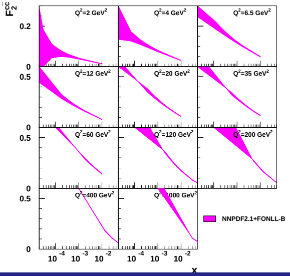


NNPDF2.1/2.0 vs. ZEUS and H1:

- General agreement of NNPDF2.0/2.1 with data.
- Heavy quark effects larger in the low-Q^2 region (larger gluon).

NNPDF2.1 predictions in view of the upcoming combined HERA data



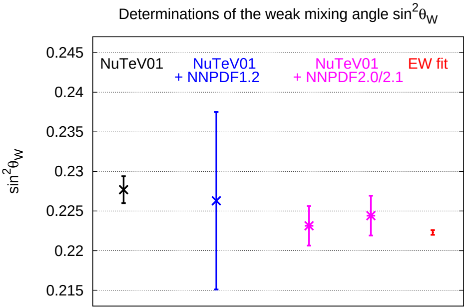


- Large PDF uncertainties at small-x and Q^2 for F_2^c $\Rightarrow$ severe constraints on the small-x gluon.

NuTeV anomaly solved!

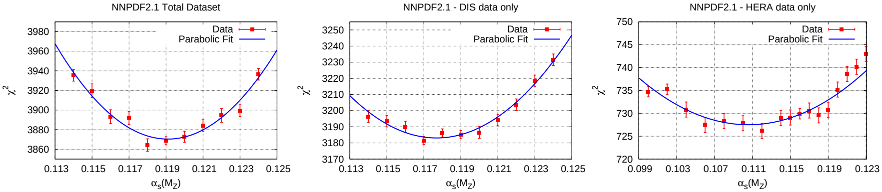
- Discrepancy at 3-σ level between **indirect determination** from NuTeV measurement (black point obtained assuming $s = \bar{s}$ and isospin symmetry) and **direct determination** (red point: EW fit).
- Uncertainty reduced (with respect of NNPDF1.2) by addition of DY data.
- Striking agreement with the EW fit relaxing $s = \bar{s}$ and isospin symm etry.

  Catani et al. [hep-ph/0404240]



Reference: [arXiv:1103.2369]

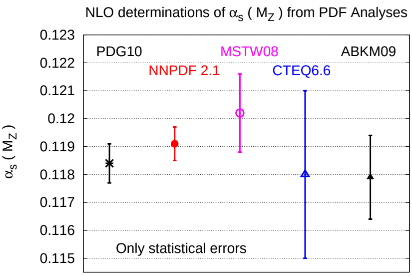
- Large dataset $\Rightarrow$ **Small statistical errors**
- α_s and PDFs correlated $\Rightarrow$ **Parametrization bias**
- Need to tame statistical fluctuations in χ^2 $\Rightarrow$ **Large replica samples**



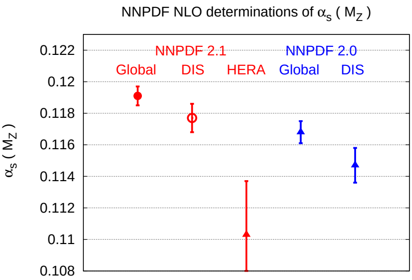
Parabolic Fit:

	$\alpha_s(M_Z)$	$\chi^2_{\mathrm{par}}/N_{\mathrm{dof}}$
NNPDF2.1	0.1191 $\pm 0.0006^{\mathrm{exp}} \pm 0.0001^{\mathrm{proc}}$	1.6
NNPDF2.1 DIS–only	$0.1177 \pm 0.0009^{\mathrm{exp}} \pm 0.0002^{\mathrm{proc}}$	0.8
NNPDF2.1 HERA–only	$0.1103 \pm 0.0033^{\mathrm{exp}} \pm 0.0003^{\mathrm{proc}}$	1.1
NNPDF2.0	0.1168 $\pm 0.0007^{\mathrm{exp}} \pm 0.0001^{\mathrm{proc}}$	0.4
NNPDF2.0 DIS–only	$0.1145 \pm 0.0010^{\mathrm{exp}} \pm 0.0003^{\mathrm{proc}}$	1.4

HQ impact rather large ($\Delta\alpha_s^{\mathrm{HQ}} \sim 0.002$, much bigger than the experimental uncertainties).

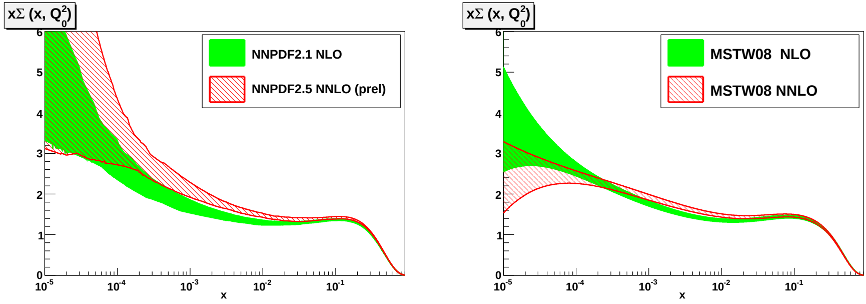


- Our best-fit in good agreement with the PDG value.
- Surprisingly small experimental uncertainty.
- Theoretical uncertainties likely dominant.

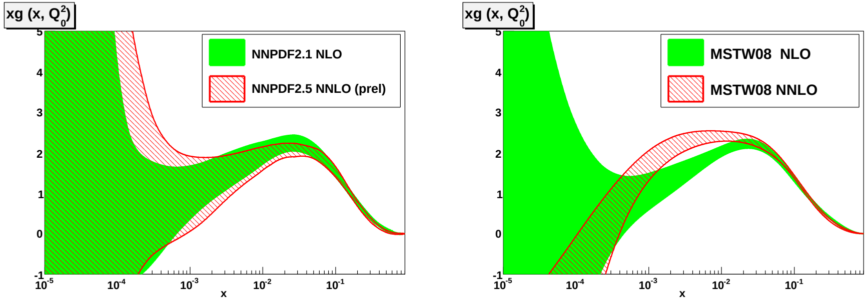


- No evidence that DIS data prefer a significantly lower value of α_s.
- Conversely, HERA data do prefer a lower value of α_s:
  - deviation between HERA data and NLO DGLAP evolution.
  - possible evidence of small-x resummation or saturation effects.

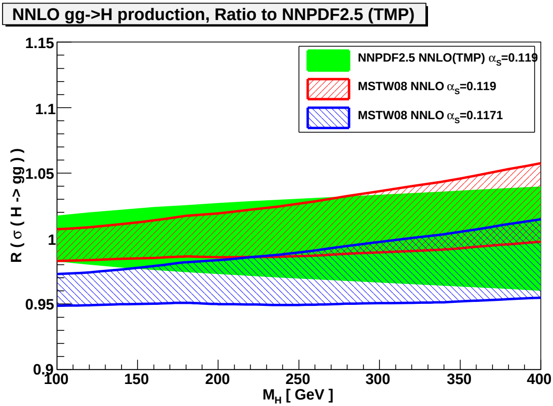
- NNLO global analysis.
- Same dataset as in NNPDF2.1, FONLL-C for DIS structure functions.
- Impact of NNLO: Larger small-x sea quarks ans 2-σ effect at medium-x



- NNLO global analysis.
- Same dataset as in NNPDF2.1, FONLL-C for DIS structure functions.
- Impact of NNLO: Stable small-x gluons (opposite to MSTW08)



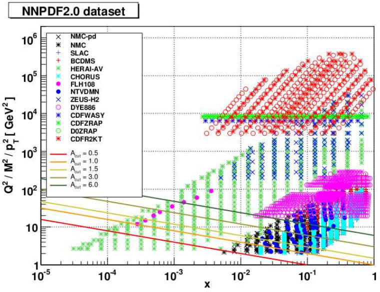
Higgs cross section at NNLO



NNPDF2.5/MSTW08 agree for the NNLO Higgs at ±5%

Part IV

Deviations from NLO DGLAP Evolution

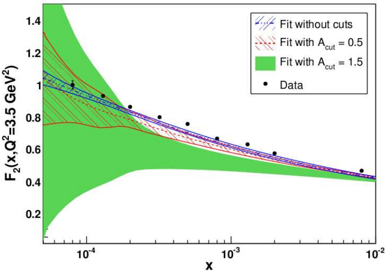


The idea (F.Caola et al. [arXiv:1007.5405]):

1. global fit only in the large-x and Q^2 region ($Q^2 \geq A_{\rm cut} x^\lambda$), where NLO DGLAP is reliable $\Rightarrow$ **safe PDFs**, free of non-DGLAP effects,
2. backward evolution and extrapolation of PDFs to potentially **unsafe** low-x and Q^2 region,
3. comparison of theoretical predictions with data.

NNPDF vs. data:

- the deviation grows bigger as $A_{\rm cut}$ is raised:
  - tension even if no cut is applied,
- clear systematic downward trend of the theoretical prediction $\Rightarrow$ **TOO MUCH EVOLUTION!**



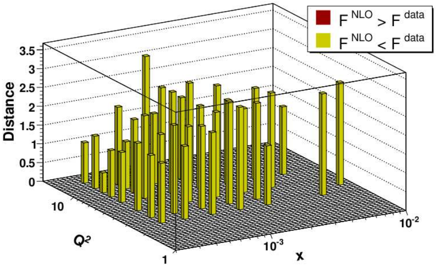
Quantification of the deviation:

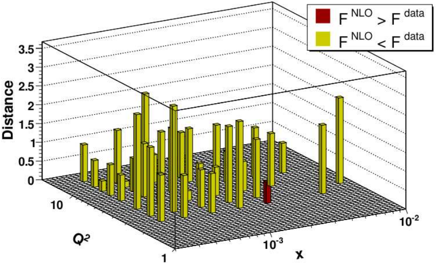
$$d(x, Q^2) = \frac{F_{\rm data} - F_{\rm fit}}{\sqrt{\sigma^2_{\rm data} + \sigma^2_{\rm fit}}}$$

Fit without any cut ($A_{\rm cut} = 0$):

$$\langle d \rangle = 1.1 \pm 0.7$$

$\Rightarrow$ systematic tension at 1-σ level.





Fit with $A_{\rm cut} = 1.5$:

$$\langle d \rangle = 2.0 \pm 0.7$$

$\Rightarrow$ **systematic deviation at 3-σ level.**

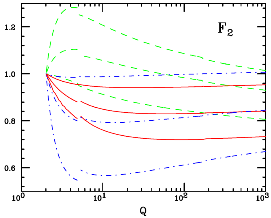
There are small but significant deviations from NLO DGLAP at small-x and Q^2

Can we understand it?

Evidence: NLO DGLAP systematically **overestimates** the evolution.

Hypothesis:

- charm mass effects ruled out by preliminary studies with NNPDF2.1 $\Rightarrow$ too small,
- NNLO effects ruled out $\Rightarrow$ even stronger evolution,
- effect compatible with small-x resummation,
- effect compatible with saturation (recombination).



Conclusions:

- need for a fully resummed PDF fit,
- with current PDFs still impossible to disentangle resummation from saturation,
- very small impact on the LHC phenomenology.

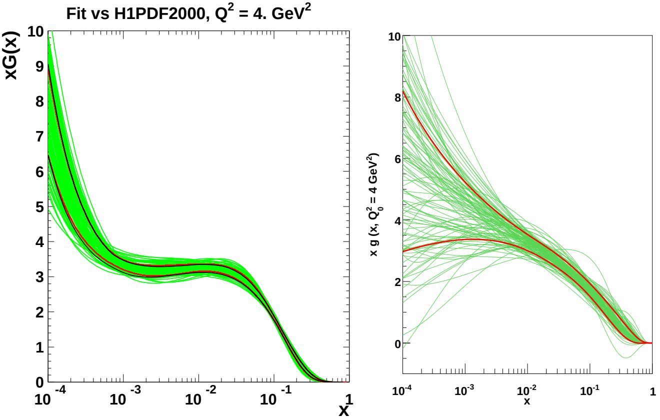
Conclusions and Outlook

- NNPDF Method: Monte Carlo ensemble + Neural Network parametrization.
  - Statistical property of PDFs reproduced using standard statistical methods.
  - Very flexible functional form able to avoid parametrization bias.
- The NNPDF2.1 analysis:
  - Enlarged data set: F_2^c data from HERA added.
  - Inclusion of heavy quark mass effects (FONLL)
  - NuTeV anomaly sorted out.
  - Precise determination of α_s.
- NLO DGLAP Studies:
  - Evidence for deviation from NLO DGLAP evolution at small-x and Q^2,
  - Resummation or saturation effects? Hard to disentangle at the moment.
  - Very small effect on the LHC phenomenology.
- The NNPDF2.1 at NNLO: upcoming.

Part VI

Backup Slides

Simple Functional Froms vs. Neural Networks



- Simple functional forms $q(x) = Ax^b(1-x)^c P(x)$ (CT, MSTW, ABKM, HERAPDF) $\rightarrow$ **systematic underestimation of uncertainties**
- Artificial Neural Networks as universal interpolants (NNPDF) $\rightarrow$ avoid theoretical bias from choice of **PDF functional form**

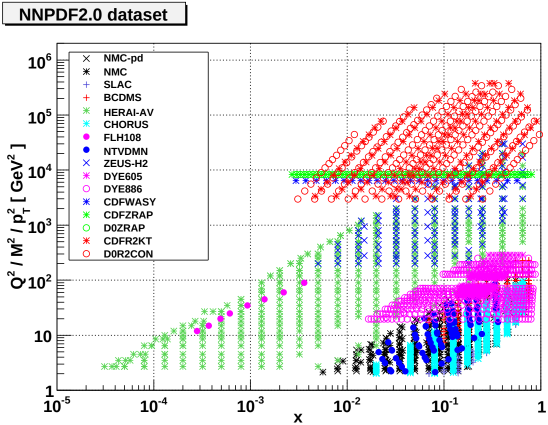
- We use **Neural Networks** as functions to represent PDFs at the starting scale.
- We employ Multilayer Feed-Forward Neural Networks trained using a **Genetic Algorithm**
- Activation determined by weights and thresholds:

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

For instance, a (1-2-1) NN is:

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

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



3477 data points

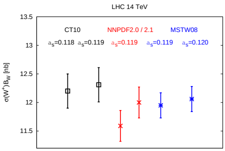
For comparison MSTW08 includes 2699 data points

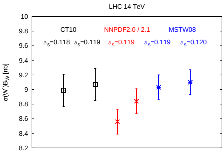
OBS	Data sets
F_2^p	NMC,SLAC,BDCMS
F_2^d	SLAC,BCDMS
F_2^d / F_2^p	NMC-pd
σ_{NC}	HERA-I AV, ZEUS-H2
σ_{CC}	HERA-I AV, ZEUS-H2
F_L	H1
σ_ν , $\sigma_{\bar{\nu}}$	CHORUS
dimuon prod.	NuTeV
$d\sigma^{\rm DY}/dM^2dy$	E605
$d\sigma^{\rm DY}/dM^2dx_F$	E886
W asymmetry	CDF
Z rap. distr.	CDF,D0
incl. $\sigma^{\rm (jet)}$	D0(cone) Run II
incl. $\sigma^{\rm (jet)}$	CDF(k_T) Run II

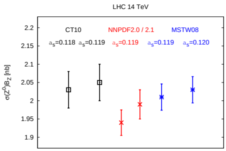
- Kinematical cuts on DIS data
  $Q^2 > 2$ GeV^2
  $W^2 = Q^2(1-x)/x > 12.5$ GeV^2
- No cuts on hadronic data
- Improved treatment of normalizations (t_0 method)[arXiv:0912.2276]
- FastKernel (exact NLO analysis)

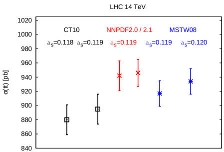
NNPDF2.1

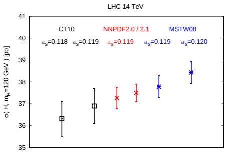
LHC Benchmark Cross Sections at $\sqrt{s} = 14$ TeV









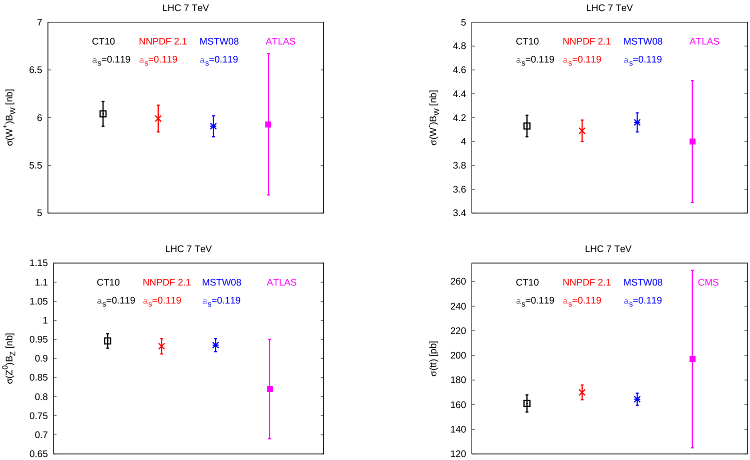


- Differences between NNPDF2.1 and NNPDF2.0 enhanced at hight energy:
  - $W^{\pm}$ and Z predictions at or just above 1-σ,
  - $t\bar{t}$ and Higgs predictions stay unchanged.
- Comparison of NNPDF2.1 with MSTW08 and CT10 is similar as before but:
  - slightly better agreement for the Higgs,
  - slightly worse agreement for $t\bar{t}$.

*All observables computed at NLO QCD using MCFM.

NNPDF2.1

Comparison with First Data



...too large uncertainty, dominated by the luminosity uncertaity ($\mathcal{O}(11\%)$):

they cannot yet provide any constraint on PDFs

Reference: R. D. Ball et al., [arXiv:1012.0836]

- The N replicas of a NNPDF fit give the probability density in the space of PDFs.
- Expectation values for observables are Monte Carlo integrals:

$$\langle \mathcal{O}[f] \rangle = \frac{1}{N} \sum_{k=1}^{N} \mathcal{O}[f_k]$$

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

- We can assess the impact of including a new dataset $\{y_1, \ldots, y_n\}$ in the fit, updating the probability density distribution and without refitting.
- According to Bayes Theorem we have:

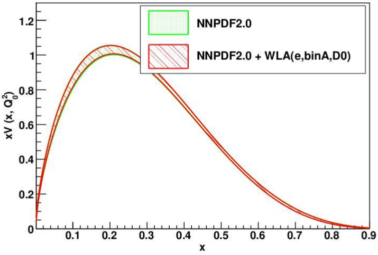
$$\langle \mathcal{O}[f] \rangle_{new} = \frac{1}{N} \sum_{k=1}^{N} w_k \mathcal{O}[f_k]$$

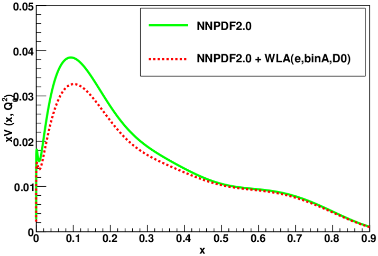
with:

$$w_k \propto \left(\chi_k^2\right)^{n/2-1} e^{-\chi_k^2/2}, \quad \sum_k w_k = 1, \quad \chi_k^2 = \sum_{i,j=1}^{n} (y_i - y_i[f_k]) \sigma_{ij}^{-1} (y_j - y_j[f_k])$$

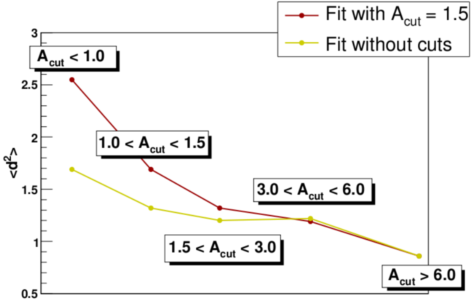
Reweighting
Reweighting data: W lepton asymmetry

- In the NNPDF2.0 fit we only included CDF W asymmetry data
- We evaluated W electron asymmetry with NNPDF20 1000 replicas set
- . . . and included D0 W electron asymmetry data points through reweighting.
- Main impact on reduction of middle-x Valence uncertainty.
- **No** need of **refitting**

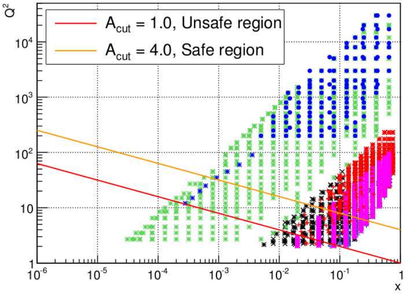




Analisys of $\langle d^2 \rangle$ in different kinematic slices:

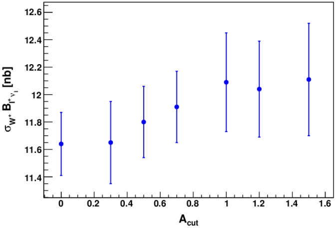


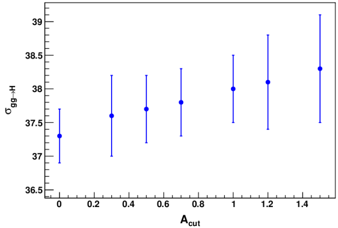
The "safe" region: No effects for $Q^2 \gtrsim 10 - 15$ GeV^2

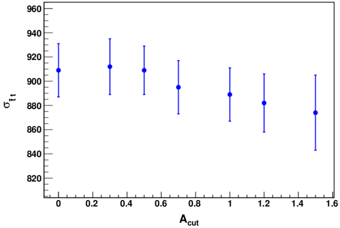


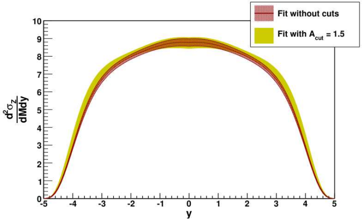
- Data and theory increasingly diverge as one moves towards small-x and Q^2.
- Deviation already present when all data are fitted, but significantly stronger when the cut is applied.

Should we worry fo the LHC phenomenology? Not Yet!









Agreement within 1-σ for all cuts