NNPDF: results and comparisons

Luigi Del Debbio

luigi.del.debbio@ed.ac.uk

University of Edinburgh

NNPDF Collaboration, Nucl. Phys. B 809, 1 (2009), [arXiv:0808.1231], **NNPDF1.0**
arXiv:0811.2288, **NNPDF1.1**
in preparation, **NNPDF1.2**
Work in collaboration

The NNPDF Collaboration

R.D.Ball\textsuperscript{1}, L.Del Debbio\textsuperscript{1}, S.Forte\textsuperscript{2}, A.Guffanti\textsuperscript{3}, J.I.Latorre\textsuperscript{4}, A. Piccione\textsuperscript{2}, J. Rojo\textsuperscript{2}, M.Ubiali\textsuperscript{1}

\textsuperscript{1} PPT Group, School of Physics, University of Edinburgh
\textsuperscript{2} Dipartimento di Fisica, Università di Milano
\textsuperscript{3} Physikalisches Institut, Albert-Ludwigs-Universität Freiburg
\textsuperscript{4} Departament d’Estructura i Constituents de la Matèria, Universitat de Barcelona
Motivation

• robust input for analyses at LHC
• HERA-LHC, PDF4LHC workshops: problems with global fits
• determination of unbiased PDFs with faithful estimate of their error

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

[kosower & giele 99]

NNPDF approach

• redundant parametrization
• MC determination of the measure \( \mathcal{P} \)
• take the experimental data at face value
• statistical estimators to assess size of systematics
Parametrization

**NNPDF**

\[ f(x, Q^2_0) = Ax^n (1 - x)^m \text{NN}(x) \]

\( \text{NN}(x) \) depends on the weights \( w_{ij} \) (fit parameters)
architecture of the net: number of parameters, \( O(100) \)

**CTEQ/MSTW/Alekhin**

\[ f(x, Q^2_0) = Ax^n (1 - x)^\xi (1 + B \sqrt{x} + \ldots) \]

smaller number of parameters (30)
asymptotic behaviour fixed by a small number of them
The Neural Monte Carlo

- Experimental data
  - BCDMS: \( F_2, \sigma_s, \sigma_{ns}, \sigma_{N}, \rho_{q} \)
  - NMC: \( F_2, \sigma_s, \sigma_{ns}, \sigma_{N}, \rho_{q} \)

- Generation of artificial data
  - \( F_{2}^{(art)(1)} \)
  - \( F_{2}^{(art)(2)} \)
  - \( F_{2}^{(art)(k)} \):
    - Tests: \( \text{net - art} \)

- Training of neural networks
  - \( F_{2}^{(net)(1)} \)
  - \( F_{2}^{(net)(2)} \)
  - \( F_{2}^{(net)(k)} \):
    - Tests: \( \text{net - net} \)

- Evolution
  - \( q_{NS}^{(net)(1)} \)
  - \( q_{NS}^{(net)(2)} \)
  - \( q_{NS}^{(net)(k)} \):
    - Tests: \( \text{net - net} \)
    - Tests: \( \text{net - exp} \)

- Determination of the probability density
  - \( q_{NS}^{(net)} \), \( \sigma_{net} \), \( \rho_{q}^{(net)} \)
The Monte Carlo ensemble

$$\langle \mathcal{F} \left[ q^{(\text{net})}_\alpha \right] \rangle = \frac{1}{N_{\text{rep}}} \sum_{k=1}^{N_{\text{rep}}} \mathcal{F} \left[ q^{(\text{net})}_\alpha(k) \right]$$

$\mathcal{F}$ can be any (LHC) observable that involves PDFs

in particular:

$$\bar{q}_\alpha = \frac{1}{N_{\text{rep}}} \sum_{k=1}^{N_{\text{rep}}} q^{(\text{net})}_\alpha(k)$$

$$\text{Var} q_\alpha = \frac{1}{N_{\text{rep}}} \sum_{k=1}^{N_{\text{rep}}} \left( q^{(\text{net})}_\alpha(k) - \bar{q}_\alpha \right)^2$$

Very different from the tolerance: $\Delta \chi^2 \sim 50$ used by MSTW/CTEQ
PDFs uncertainties

- Monte Carlo prescription ([NNPDF])

\[
\sigma_{\mathcal{F}} = \left( \frac{N_{\text{set}}}{N_{\text{set}} - 1} \left( \langle \mathcal{F}[\{f\}] \rangle^2 - \langle \mathcal{F}[\{f\}] \rangle^2 \right) \right)^{1/2}
\]

- HEPDATA prescription ([CTEQ and MRST/MSTW])

\[
\sigma_{\mathcal{F}} = \frac{1}{2C_{90}} \left( \sum_{k=1}^{N_{\text{set}}/2} \left( \mathcal{F}[\{f^{(2k-1)}\}] - \mathcal{F}[\{f^{(2k)}\}] \right)^2 \right)^{1/2}, \quad C_{90} = 1.64485
\]

\(C_{90}\) accounts for the fact that the upper and lower parton sets correspond to 90% confidence levels rather than to one-\(\sigma\) uncertainties.

- HEPDATA* prescription ([Alekhin])

\[
\sigma_{\mathcal{F}} = \left( \sum_{k=1}^{N_{\text{set}}} \left( \mathcal{F}[\{f^{(k)}\}] - \mathcal{F}[\{f^{(0)}\}] \right)^2 \right)^{1/2}.
\]
Datasets

**NNPDF1.0**: full DIS fit + ZM-VFN + flavour assumptions, \(Q^2 > 2 \text{ GeV}^2, W^2 > 12 \text{ GeV}^2\)

**NNPDF1.1**: \(s\) and \(\bar{s}\) distributions parametrized

**NNPDF1.2**: dimuon data added
Some more details

$$\chi^2(k)[\omega] = \sum_{i,j}^{N_{\text{dat}}} (F_i^{(\text{art})(k)} - F_i^{(\text{net})(k)}) \left( \left( \text{COV}(k) \right)^{-1} \right)_{ij} (F_j^{(\text{art})(k)} - F_j^{(\text{net})(k)})$$

- fully–correlated $\chi^2$ is used
- $\text{COV}(k)$ defined from an experimental covariance matrix which does not include normalization errors [D'Agostini 2003]

$$\text{COV}_{ij}^{(k)} = \text{COV}^{(\text{exp})}_{ij} S_{i,N}^{(k)} S_{j,N}^{(k)}$$

- $F_i^{(\text{net})}$ computed from PDFs using NLO, ZM-VFN scheme
- $\alpha_s$ is not fitted - kept fixed
- $N_{\text{rep}} \simeq 100 \div 1000$ to obtain an accurate description of data
Results – Singlet
- NNPDF1.0 uncertainties correspond to a genuine 68% CL
- PDF error larger than other PDF sets in some regions (extrapolation), smaller in others (not artificially inflated by large $\Delta \chi^2 \sim 50/100$)
- in general close to CTEQ6.5 in data region
Monte Carlo sampling of PDFs
Results – Valence
### W & Z cross-sections

#### W⁺ Cross Section at the LHC [MCFM]

<table>
<thead>
<tr>
<th>PDF Set</th>
<th>(\sigma_{W^+}) [nb]</th>
<th>(\Delta \sigma_{W^+}/\sigma_{W^+})</th>
</tr>
</thead>
<tbody>
<tr>
<td>NNPDF1.0</td>
<td>11.83 ± 0.26</td>
<td>2.2%</td>
</tr>
<tr>
<td>CTEQ6.1</td>
<td>11.65 ± 0.34</td>
<td>2.9%</td>
</tr>
<tr>
<td>MRST01</td>
<td>11.71 ± 0.14</td>
<td>1.2%</td>
</tr>
<tr>
<td>CTEQ6.5</td>
<td>12.54 ± 0.29</td>
<td>2.3%</td>
</tr>
</tbody>
</table>

#### Z⁰ Cross Section at the LHC [MCFM]

<table>
<thead>
<tr>
<th>PDF Set</th>
<th>(\sigma_{Z^0}) [nb]</th>
<th>(\Delta \sigma_{Z^0}/\sigma_{Z^0})</th>
</tr>
</thead>
<tbody>
<tr>
<td>NNPDF1.0</td>
<td>11.83 ± 0.26</td>
<td>2.2%</td>
</tr>
<tr>
<td>CTEQ6.1</td>
<td>11.65 ± 0.34</td>
<td>2.9%</td>
</tr>
<tr>
<td>MRST01</td>
<td>11.71 ± 0.14</td>
<td>1.2%</td>
</tr>
<tr>
<td>CTEQ6.5</td>
<td>12.54 ± 0.29</td>
<td>2.3%</td>
</tr>
</tbody>
</table>
Comparing PDFs

HERA-LHC benchmark

Benchmark PDF fit to a reduced consistent set of DIS data  [hep-ph/0511119]

3163 data → 773 data
Comparing PDFs

**HERA-LHC benchmark**

Comparison between collaborations and between benchmark/global partons.

\( u(x, Q^2 = 2\text{GeV}^2) \): MRST data region
Comparing PDFs

HERA-LHC benchmark

Comparison between collaborations and between benchmark/global partons. 
\( u(x, Q^2 = 2\text{GeV}^2) \): MRST extrapolation region
Comparing PDFs

HERA-LHC benchmark

• benchmark partons and global partons do not agree within error!
• note that PDFs input parametrization, flavor assumptions and statistical treatment ($\Delta \chi^2_{\text{global}} = 50$, $\Delta \chi^2_{\text{bench}} = 1$) are tuned to data.
• not satisfactory especially to predict the behaviour of PDFs in the extrapolation region (LHC)
Comparing PDFs

**HERA-LHC benchmark**

Comparison between collaborations and between benchmark/global partons.

\( u(x, Q^2 = 2\text{GeV}^2) \): data region
Comparing PDFs

HERA-LHC benchmark

Comparison between collaborations and between benchmark/global partons. 

\( u(x, Q^2 = 2 \text{GeV}^2) \): extrapolation region

![Graph showing comparisons between different PDF models](image-url)
Comparing PDFs

HERA-LHC benchmark

- NNPDF1.0 is consistent with MRST global fit.
- NNPDFbench is consistent with NNPDF1.0 and MRST.
- Same parametrization and flavour assumption.
- Same statistical treatment.
- Underestimation of the error in the standard approach.
- New MSTW fit has improved, but still tension btw global and benchmark fits
Conclusions

• Standard approaches with fixed parametrization tend to underestimate uncertainties unless experimental errors are inflated.

• **Monte Carlo** ensemble
  - Any statistical property of PDFs can be calculated using standard statistical methods.
  - No need of any tolerance criterion.

• The **Neural Network** parametrization
  - Small uncertainties come from an underlying physical law, not from parametrization bias.
  - Inconsistent data or underestimated uncertainties do not require a separate treatment and are automatically signalled by a larger value of the $\chi^2$.

• The first NNPDF parton set [arXiv:0808.1231] is available on the common LHAPDF interface [http://projects.hepforge.org/lhapdf].
Conclusions

Unbiased PDFs with statistically meaningful error bars are important for LHC analyses

- **NNPDF1.0**: NLO DIS fit, available from LHAPDF
- **NNPDF1.1**: available from http://sophia.ecm.ub.es/nnpdf
- **NNPDF1.2**: dimuon data added, strange content fitted - to appear soon

Forthcoming developments:

- inclusion of hadronic data
- heavy quarks effects

towards **NNPDF2.0**: global fit with faithful errors
Some details about Neural Networks

Multilayer feed-forward networks

- Each neuron receives input from neurons in preceding layer and feeds output to neurons in subsequent layer
- Activation determined by weights and thresholds
  \[ \xi_i = g \left( \sum_j w_{ij} \xi_j - \theta_i \right) \]
- Sigmoid activation function
  \[ g(x) = \frac{1}{1+e^{-\beta x}} \]

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

Thm: any function can be represented by a sufficiently big neural network

... just another set of basis functions!
MC sampling of exp data

\[ F_{I,p}^{(\text{art})(k)} = S_{p,N}^{(k)} F_{I,p}^{(\text{exp})} \left( 1 + \sum_{l=1}^{N_c} r_{p,l}^{(k)} \sigma_{p,l} + r_p^{(k)} \sigma_{p,s} \right), \quad k = 1, \ldots, N_{\text{rep}}, \]

where

\[ S_{p,N}^{(k)} = \prod_{n=1}^{N_a} \left( 1 + r_{p,n}^{(k)} \sigma_{p,n} \right) \prod_{n=1}^{N_r} \sqrt{1 + r_{p,n}^{(k)} \sigma_{p,n}}. \]
Basis set

- Each independent PDF at the initial scale $Q_0^2 = 2\text{GeV}^2$ is parameterized by an individual NN.

- Little constraint on strange → Flavor Assumptions:
  - Symmetric strange sea $s(x) = \bar{s}(x)$
  - Strange sea proportional to non-strange sea $\bar{s}(x) = \frac{C}{2}(\bar{u}(x) + \bar{d}(x))$ (C = 0.5)
  - Intrinsic heavy quarks contributions neglected.

- Parametrization of $(4+1)$ combinations of PDFs at $Q_0^2 = 2\text{GeV}^2$:

  Singlet : $\Sigma(x)$ $\leftrightarrow$ NN$\Sigma(x)$ 2-5-3-1 37 pars

  Gluon : $g(x)$ $\leftrightarrow$ NN$g(x)$ 2-5-3-1 37 pars

  Total valence : $V(x) \equiv u_V(x) + d_V(x)$ $\leftrightarrow$ NN$V(x)$ 2-5-3-1 37 pars

  Non-singlet triplet : $T_3(x)$ $\leftrightarrow$ NN$T_3(x)$ 2-5-3-1 37 pars

  Sea asymmetry : $\Delta_S(x) \equiv \bar{d}(x) - \bar{u}(x)$ $\leftrightarrow$ NN$\Delta(x)$ 2-5-3-1 37 pars

185 parameters
Flavour assumptions

NNPDF1.0

\[ \Sigma(x, Q_0^2) = (1 - x)^{m_{\Sigma}} x^{-n_{\Sigma}} NN_{\Sigma}(x), \]
\[ V(x, Q_0^2) = A_{V} (1 - x)^{m_{V}} x^{-n_{V}} NN_{V}(x), \]
\[ T_3(x, Q_0^2) = (1 - x)^{m_{T_3}} x^{-n_{T_3}} NN_{T_3}(x), \]
\[ \Delta S(x, Q_0^2) = A_{\Delta S} (1 - x)^{m_{\Delta S}} x^{-n_{\Delta S}} NN_{\Delta S}(x), \]
\[ g(x, Q_0^2) = A_{g} (1 - x)^{m_{g}} x^{-n_{g}} NN_{g}(x), \]
\[ s(x, Q_0^2) = \bar{s}(x, Q_0^2) = C_s/2 (\bar{u}(x, Q_0^2) + \bar{d}(x, Q_0^2)) \]

Normalization factors \( \rightarrow \) fixed by valence and momentum sum rules

\[ \int_0^1 dx \ x (\Sigma(x) + g(x)) = 1 \]
\[ \int_0^1 dx \ (u(x) - \bar{u}(x)) = 2 \]
\[ \int_0^1 dx \ (d(x) - \bar{d}(x)) = 1 \]
Training strategy

- unbiased basis of functions, parametrized by a large number of parameters
- genetic algorithms for minimization
- might accommodate statistical fluctuations of the data
- optimal training, beyond which the fit is just adjusting to statistical fluctuations
- dynamical stopping by cross validation
- for each replica divide the data randomly into training and validation
- minimization performed on the training set only
- when the training $\chi^2$ still decreases while the validation $\chi^2$ stops decreasing $\rightarrow$ STOP
Dynamical stopping

![Graph showing iterations and energy (E) over iterations. The graph displays two datasets: Training Dataset and Validation Dataset. The stopping point is indicated by a green triangle.](image-url)
Distance between MC ensembles.

- Stability of the NNPDF parton set can be assessed by using standard statistical tools.

- Distances between two probability distributions: \( \left\{ f_{ik}^{(1)} = f_{k}^{(1)}(x_i, Q_0^2) \right\} \)

\[
\langle d[f] \rangle = \sqrt{\left\langle \left( \frac{\langle f_i \rangle^{(1)} - \langle f_i \rangle^{(2)} \rangle^2}{\sigma^2[f_i^{(1)}] + \sigma^2[f_i^{(2)}]} \right)_{\text{pts}} \right\rangle}
\]

- where:

\[
\langle f_i \rangle^{(1)} \equiv \frac{1}{N_{\text{rep}}} \sum_{k=1}^{N_{\text{rep}}} f_{ik}^{(1)},
\]

\[
\sigma^2[f_i^{(1)}] \equiv \frac{1}{N_{\text{rep}}(N_{\text{rep}} - 1)} \sum_{k=1}^{N_{\text{rep}}} \left( f_{ik}^{(1)} - \langle f_i \rangle^{(1)} \right)^2
\]

- For statistically equivalent PDF sets: \( \langle d[f] \rangle \sim \langle d[\sigma_f] \rangle \sim 1 \)
Stability under variation of the parametrization

- Stability under change of architecture of the nets:
  - 37 pars $\rightarrow$ 31 pars
- Independence on the parametrization!

<table>
<thead>
<tr>
<th></th>
<th>Data</th>
<th>Extrapolation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Sigma(x, Q_0^2)$</td>
<td>$5 \times 10^{-4} \leq x \leq 0.1$</td>
<td>$10^{-5} \leq x \leq 10^{-4}$</td>
</tr>
<tr>
<td>$\langle d[f] \rangle$</td>
<td>0.98</td>
<td>1.25</td>
</tr>
<tr>
<td>$\langle d[\sigma] \rangle$</td>
<td>1.14</td>
<td>1.34</td>
</tr>
<tr>
<td>$g(x, Q_0^2)$</td>
<td>$5 \times 10^{-4} \leq x \leq 0.1$</td>
<td>$10^{-5} \leq x \leq 10^{-4}$</td>
</tr>
<tr>
<td>$\langle d[f] \rangle$</td>
<td>1.52</td>
<td>1.15</td>
</tr>
<tr>
<td>$\langle d[\sigma] \rangle$</td>
<td>1.16</td>
<td>1.07</td>
</tr>
<tr>
<td>$T_3(x, Q_0^2)$</td>
<td>$0.05 \leq x \leq 0.75$</td>
<td>$10^{-3} \leq x \leq 10^{-2}$</td>
</tr>
<tr>
<td>$\langle d[f] \rangle$</td>
<td>1.00</td>
<td>1.11</td>
</tr>
<tr>
<td>$\langle d[\sigma] \rangle$</td>
<td>1.76</td>
<td>2.27</td>
</tr>
<tr>
<td>$V(x, Q_0^2)$</td>
<td>$0.1 \leq x \leq 0.6$</td>
<td>$3 \times 10^{-3} \leq x \leq 3 \times 10^{-2}$</td>
</tr>
<tr>
<td>$\langle d[f] \rangle$</td>
<td>1.30</td>
<td>0.90</td>
</tr>
<tr>
<td>$\langle d[\sigma] \rangle$</td>
<td>1.10</td>
<td>0.98</td>
</tr>
<tr>
<td>$\Delta_S(x, Q_0^2)$</td>
<td>$0.1 \leq x \leq 0.6$</td>
<td>$3 \times 10^{-3} \leq x \leq 3 \times 10^{-2}$</td>
</tr>
<tr>
<td>$\langle d[f] \rangle$</td>
<td>1.04</td>
<td>1.91</td>
</tr>
<tr>
<td>$\langle d[\sigma] \rangle$</td>
<td>1.44</td>
<td>1.80</td>
</tr>
</tbody>
</table>