Neural networks approach to parton distributions fitting

NNPDF Collaboration

ACAT 2005

26th May 2005

The NNPDF Collaboration

Luigi Del Debbio[1], Stefano Forte[2],
José I. Latorre[3], Andrea Piccione[4] and Joan Rojo[3]

[1] Theory Division, CERN
[2] Dipartimento di Fisica, Università di Milano and INFN, Sezione di Milano
[3] Departament d'Estructura i Constituents de la Matèria, Universitat de Barcelona
[4] Dipartimento di Fisica Teorica, Università di Torino and INFN, Sezione di Torino

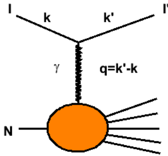
QCD and Hadrons

- QCD describes interactions between quarks and gluons.
  Experimentally we observe only hadrons $\longrightarrow$ Confinement
- Perturbative QCD is not trustable at low energies ($\sim$ GeV).
  We can not solve QCD in the non-perturbative region, but on a lattice ...
- We can extract information on the proton structure from a process with only one initial proton (DIS at HERA).
  Then we can use these as an input for a process where two initial protons are involved (DY at LHC) $\longrightarrow$ Factorization

QCD and Hadrons

- QCD describes interactions between quarks and gluons. Experimentally we observe only hadrons $\longrightarrow$ Confinement
- Perturbative QCD is not trustable at low energies ($\sim$ GeV). We can not solve QCD in the non-perturbative region, but on a lattice ...
- We can extract information on the proton structure from a process with only one initial proton (DIS at HERA). Then we can use these as an input for a process where two initial protons are involved (DY at LHC) $\longrightarrow$ Factorization

Deep Inelastic Scattering



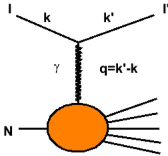
- The cross section

$$\frac{d^2\sigma}{dxdQ^2} = \frac{4\pi\alpha^2}{Q^4}\left[[1+(1-y)^2]F_1 + \frac{1-y}{x}(F_2 - 2xF_1)\right]$$

- The structure function (QCD + parton model)

$$F_2(x,\, Q^2) = x\left[\sum_{q=1}^{n_f} e_q^2\, C^q \otimes q_q(x,\, Q^2) + 2n_f\, C^g \otimes g(x,\, Q^2)\right]$$

Deep Inelastic Scattering



- The cross section

$$\frac{d^2\sigma}{dxdQ^2} = \frac{4\pi\alpha^2}{Q^4}\left[[1+(1-y)^2]F_1 + \frac{1-y}{x}\left(F_2 - 2xF_1\right)\right]$$

- The structure function (QCD + parton model)

$$F_2(x,\,Q^2) = x\left[\sum_{q=1}^{n_f} e_q^2\,\mathcal{C}^q \otimes q_q(x,\,Q^2) + 2n_f\,\mathcal{C}^g \otimes g(x,\,Q^2)\right]$$

The problem: I

- Structure function (or Xsec) is a convolution over x of PDFs and perturbative cross section → Deconvolution
- Each structure function (or Xsec) is a linear combination of many PDFs ($2n_f$ quarks + gluon) → Different processes
- Data are given at various scales, and we want PDFs as functions of x at a common scale Q^2 → Evolution
- TH uncertainties: resummation, nuclear corrections, higher twist, heavy quark thresholds, ...

Which is the uncertainty associated with a PDFs set?
[Djouadi and Ferrag 2003, Frixione and Mangano 2004, Tung 2004, HERA and the LHC Workshop 2004-2005]

The problem: I

- Structure function (or Xsec) is a convolution over x of PDFs and perturbative cross section $\rightarrow$ Deconvolution
- Each structure function (or Xsec) is a linear combination of many PDFs ($2n_f$ quarks + gluon) $\rightarrow$ Different processes
- Data are given at various scales, and we want PDFs as functions of x at a common scale Q^2 $\rightarrow$ Evolution
- TH uncertainties: resummation, nuclear corrections, higher twist, heavy quark thresholds, ...

Which is the uncertainty associated with a PDFs set?
[Djouadi and Ferrag 2003, Frixione and Mangano 2004, Tung 2004, HERA and the LHC Workshop 2004-2005]

The problem: I

- Structure function (or Xsec) is a convolution over x of PDFs and perturbative cross section $\rightarrow$ Deconvolution
- Each structure function (or Xsec) is a linear combination of many PDFs ($2n_f$ quarks + gluon) $\rightarrow$ Different processes
- Data are given at various scales, and we want PDFs as functions of x at a common scale Q^2 $\rightarrow$ Evolution
- TH uncertainties: resummation, nuclear corrections, higher twist, heavy quark thresholds, ...

Which is the uncertainty associated with a PDFs set?
[Djouadi and Ferrag 2003, Frixione and Mangano 2004, Tung 2004, HERA and the LHC Workshop 2004-2005]

The Problem: II

- For a single quantity $\rightarrow$ 1 sigma error
- For a pair of numbers $\rightarrow$ 1 sigma ellipse
- For a function $\rightarrow$ We need the probability measure $\mathcal{P}[f]$ in the space of functions $f(x)$

Expectation values $\rightarrow$ Functional integrals

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

Determine an infinite-dimensional object (a function) from finite set of data points $\rightarrow$ Mathematically ill-posed problem

The standard approach

1. Choose a simple functional form with enough free parameters
2. Fit parameters by minimizing χ^2

Some difficulties arise:

- Errors and correlations of parameters require at least fully correlated analysis of data errors
- Error propagation to observables is difficult: many observables are nonlinear/nonlocal functional of parameters
- Theoretical bias due to choice of parametrization is difficult to assess (effects can be large if data are not precise or hardly compatible)

The NNPDF approach

- Determination of the Structure Functions:
  this is the easiest case, since no evolution is required, but only data fitting. A good application to test the technique → Done
- Determination of the Parton Distributions:
  the real stuff → Working on it ...

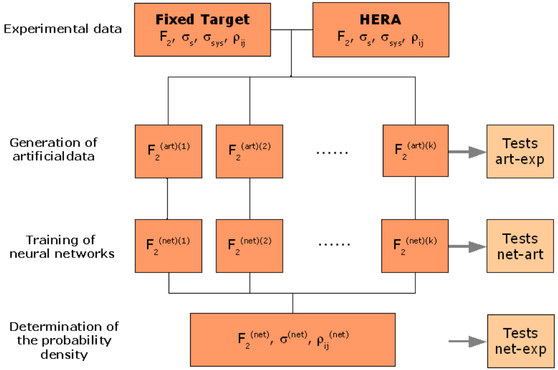
General strategy: I

- Monte Carlo sampling of data (generation of replicas of experimental data) $\longrightarrow$ Faithful representation of uncertainties
- NN training over MC replicas $\longrightarrow$ Unbiased parametrization

Expectation values $\longrightarrow$ Sum over the Nets

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

General strategy: II



Training: I

- Architecture: 4-5-3-1
- Back Propagation ($\sim 10^8$ training cycles):

$$\chi^{2\,(k)}_{\mathrm{diag}} = \frac{1}{N_{\mathrm{dat}}} \sum_{i=1}^{N_{\mathrm{dat}}} \frac{\left(F_i^{(\mathrm{art})(k)} - F_i^{(\mathrm{net})(k)}\right)^2}{\sigma_{i,t}^{(\mathrm{exp})2}}$$

- Genetic Algorithm ($\sim 10^4$ generations):

$$\chi^{2\,(k)} = \frac{1}{N_{\mathrm{dat}}} \sum_{i,j=1}^{N_{\mathrm{dat}}} \left(F_i^{(\mathrm{art})(k)} - F_i^{(\mathrm{net})(k)}\right) \mathrm{cov}_{ij}^{-1} \left(F_j^{(\mathrm{art})(k)} - F_j^{(\mathrm{net})(k)}\right)$$

Training: I

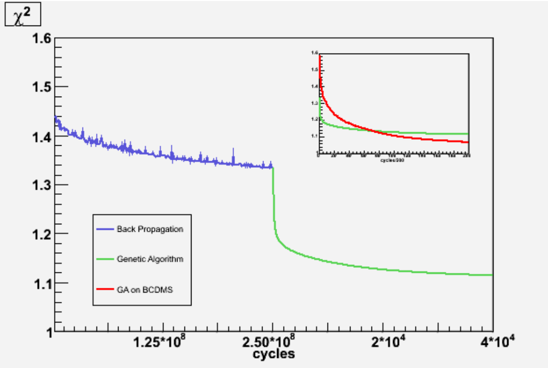
- Architecture: 4-5-3-1
- Back Propagation ($\sim 10^8$ training cycles):

$$\chi^{2\,(k)}_{\mathrm{diag}} = \frac{1}{N_{\mathrm{dat}}} \sum_{i=1}^{N_{\mathrm{dat}}} \frac{\left(F_i^{(\mathrm{art})(k)} - F_i^{(\mathrm{net})(k)}\right)^2}{\sigma_{i,t}^{(\mathrm{exp})2}}$$

- Genetic Algorithm ($\sim 10^4$ generations):

$$\chi^{2\,(k)} = \frac{1}{N_{\mathrm{dat}}} \sum_{i,j=1}^{N_{\mathrm{dat}}} \left(F_i^{(\mathrm{art})(k)} - F_i^{(\mathrm{net})(k)}\right) \mathrm{cov}_{ij}^{-1} \left(F_j^{(\mathrm{art})(k)} - F_j^{(\mathrm{net})(k)}\right)$$

Training: II


χ^2
1.6
1.5
1.4
1.3
1.2
1.1
1
1.25*10^8
2.50*10^8
2*10^4
4*10^4
cycles
Back Propagation
Genetic Algorithm
GA on BCDMS

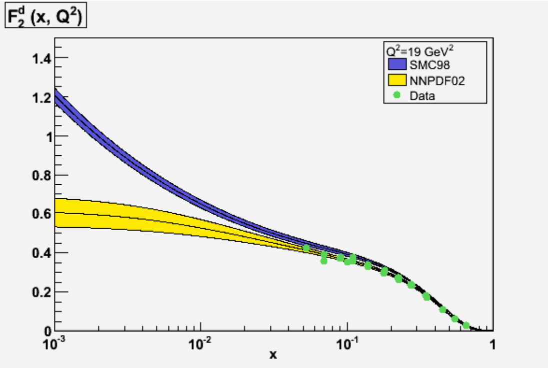
Credits

- S. Forte, L. Garrido, J. I. Latorre and A. P., *"Neural network parametrization of deep-inelastic structure functions,"* JHEP05 (2002) 062 [arXiv:hep-ph/0204232]
- L. Del Debbio, S. Forte, J. I. Latorre, A. P. and J. Rojo [NNPDF Collaboration], *"Unbiased determination of the proton structure function F_2^p with faithful uncertainty estimation"*, JHEP03 (2005) 080 [arXiv:hep-ph/0501067]

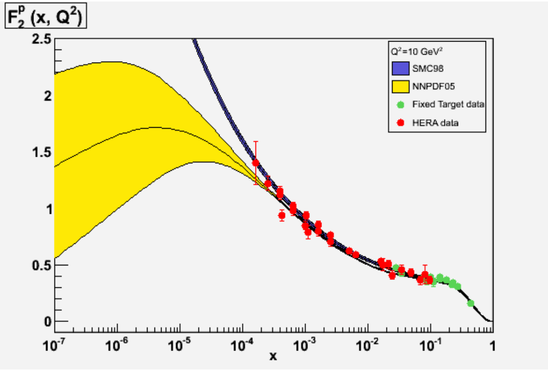
Source code, driver program and graphical web interface for F_2 plots and numerical computations available

http://sophia.ecm.ub.es/f2neural

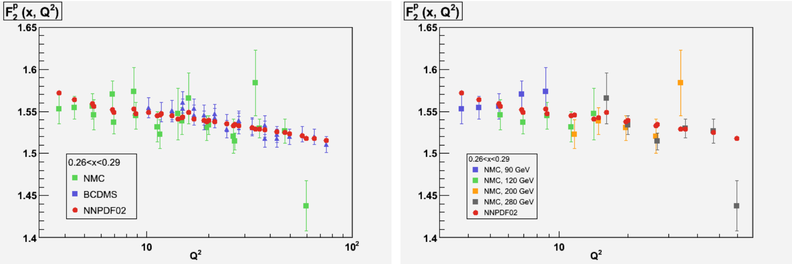
Fit of $F_2^d(x, Q^2)$ *[NNPDF 2002]*



Fit of $F_2^p(x, Q^2)$ *[NNPDF 2005]*



Incompatible data *[NNPDF 2002]*



Strategy

Same strategy as with SF + Altarelli-Parisi evolution

- Monte Carlo sampling of data
- Parametrize parton distributions with neural networks
- Evolution of parton distributions to experimental data scale and training over Monte Carlo replica sample

Strategy

Same strategy as with SF + Altarelli-Parisi evolution

- Monte Carlo sampling of data
- Parametrize parton distributions with neural networks
- Evolution of parton distributions to experimental data scale and training over Monte Carlo replica sample

Examples

- Expectation values:

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

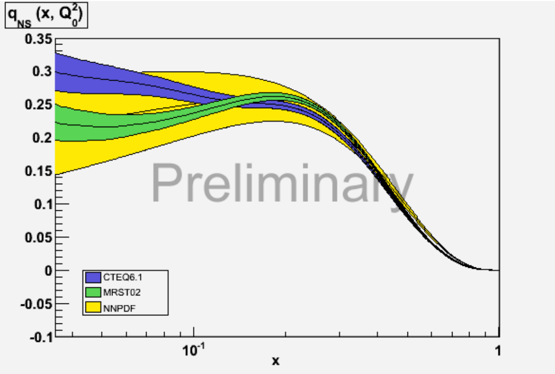
- Correlations between pairs of different parton distributions at different points:

$$\langle u(x_1)d(x_2)\rangle = \frac{1}{N_{rep}} \sum_{k=1}^{N_{rep}} u^{(net)(k)}(x_1, Q_0^2)d^{(net)(k)}(x_2, Q_0^2)$$

Details

- $q_{NS}(x, Q^2) \equiv \frac{1}{6}\left(u + \bar{u} - d - \bar{d}\right)(x, Q^2)$
- Experimental data: NMC (94 pts) and BCDMS (253 pts)
- Kinematical cuts: $Q^2 \geq 9\ GeV^2$, $W^2 \geq 6.25\ GeV^2$
- Neural network architecture: 2-2-2-1 (15 params.)
- Strong coupling: $\alpha_s\left(M_Z^2\right) = 0.1182$
- Perturbative order: NNLO
- VFN: $m_c = 1.5 GeV$, $m_b = 4.5 GeV$, $m_t = 175 GeV$
- TMC: F_2 integral evaluated with NN F_2
- # replica: 25 (should be 1000)

Non-Singlet


$q_{NS}(x, Q_0^2)$
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
-0.05
-0.1
10^{-1}
1
x
Preliminary
CTEQ6.1
MRST02
NNPDF

Summary

- Unbiased determination of structure functions with faithful estimation of uncertainties
- Successful implementation of neural parton fitting at NNLO

Outlook

- Construct full set of NNPDF parton distributions from all available data
- Estimate impact of theoretical uncertainties
- Assess impact of uncertainties of PDFs for relevant observables at LHC
- Perform a benchmark set of pdfs, to compare the different fitting programs (CTEQ,MRST, Alekhin)
- Make formalism compatible with standard interfaces (LHAPDF, PDFLIB) $\rightarrow$ NNPDF partons available for use in Monte Carlo generators