Recent progress on neural PDFs

NNPDF Collaboration

HERA and the LHC Workshop

21st March 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

DIS data $\rightarrow$ Structure Functions

- Structure Function=Hard. Coeff. $\otimes$ Parton Distn.

$$F^{NC}(x,Q^2) = x\sum_f e_f^2 (q_f + \bar{q}_f) + \alpha_s \left[C_f\left(\alpha_s\right) \otimes (q_f + \bar{q}_f) C_g\left(\alpha_s\right) \otimes g \right]$$

- Trivial complication: disentangle quark flavors and gluon, evolve to common scale, deconvolute
- Serious complication: determine errors on PDFs

The Problem

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

Expectation values $\rightarrow$ Functional integrals

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

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 $\longrightarrow$ Done
- Determination of the Parton Distributions $\longrightarrow$ Working on it ...

General strategy: I

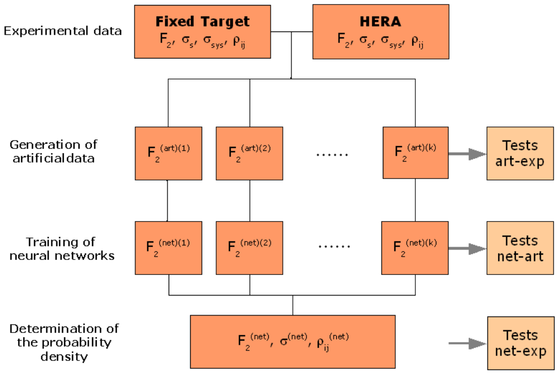
- Monte Carlo sampling of data (Generation of replicas of experimental data) $\longrightarrow$ Faithful representation of uncertainties
- Neural network training over Monte Carlo replicas $\longrightarrow$ Unbiased parametrization

Expectation values $\longrightarrow$ Sum over the Nets

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

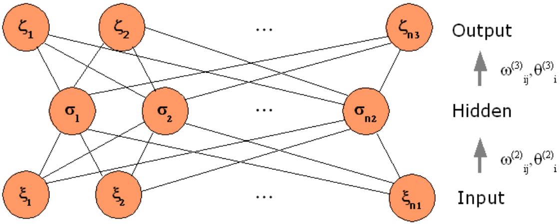
$\mathcal{P}\left[F(x)\right]$ validated through statistical estimators

General strategy: II



Structure I

Neural networks: a class of algorithms providing robust, universal, unbiased approximants to incomplete or noisy data



Structure II

- Building blocks: neurons, *i. e.* input/output units characterized by sigmoid activation

$$\xi_i^{(l)} = g\left(\sum_{j=1}^{n_l-1} \omega_{ij}^{(l-1)} \xi_j^{(l-1)} - \theta_i^{(l)}\right) \qquad g(x) = \frac{1}{1+e^{-x}}$$

- Parameters: weights $\omega_{ij}^{(l)}$ and thresholds $\theta_i^{(l)}$.
- Architecture: multilayer feed-forward NN. Each neuron receives input from neurons in preceding layer and feeds output to neurons in successive layer
- Assumption: smooth function

Training

- Architecture: redundant to avoid smoothing bias
- Learning: supervised training on covariance matrix error (highly nonlocal error function)
- Training method: Genetic Algorithm (extremely effective to find the global minimum, but slow convergence rate)

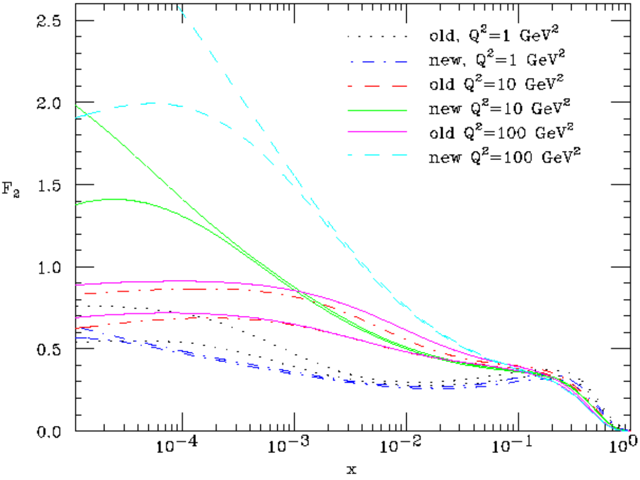
Credits

- S. Forte, L. Garrido, J. I. Latorre and A. P., *"Neural network parametrization of deep-inelastic structure functions,"* JHEP **0205** (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"*, [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 $F_2^p(x, Q^2)$



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

The probability measure $\mathcal{P}[q]$ contains all information from experimental data (central values, errors, correlations)

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:

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

Evolution kernel

- We want Mellin space evolution:

$$q(N,Q^2) = q(N,Q_0^2)\Gamma\left(N,\alpha_s\left(Q^2\right),\alpha_s\left(Q_0^2\right)\right)$$

- We do not want complex neural networks:

$$\Gamma\left(x,\alpha_s\left(Q^2\right),\alpha_s\left(Q_0^2\right)\right) \equiv \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} dN\, x^{-N}\Gamma\left(N,\alpha_s\left(Q^2\right),\alpha_s\left(Q_0^2\right)\right)$$

- $\Gamma(x)$ is a distribution $\to$ must be regularized at $x=1$:

$$q(x,Q^2) = q(x,Q_0^2)\int_x^1 dy\,\Gamma(y) + \int_x^1 \frac{dy}{y}\Gamma(y)\left(q\left(\frac{x}{y},Q_0^2\right) - yq(x,Q_0^2)\right)$$

Some details

- At higher orders $\rightarrow$ Wilson coefficients $C\left(N,\alpha_s\left(Q^2\right)\right)$

$$\tilde{\Gamma}\left(x,\alpha_s\left(Q^2\right),\alpha_s\left(Q_0^2\right)\right)=\frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty}dN\; x^{-N}C\left(N,\alpha_s\left(Q^2\right)\right)\Gamma\left(N\right)$$

- Mellin transform inversion of evolution factor is crucial.
  We tested different paths and algorithms $\rightarrow$ Fixed Talbot

Interpolation

- During *pdf training* $\Gamma(x)$ is called many times $\rightarrow$ Interpolate $\Gamma(x)$ before the fit (hard numerical task)
- For each Q^2 bin x interpolation with ~ 100 Chebyshev polynomials

$$\Gamma(x) \approx \left[\sum_{k=1}^{N} c_k T_{k-1}(x) \right] - \frac{1}{2} c_1 \qquad T_n(x) = \cos(n \arccos x)$$

LHAPDF Benchmark

Next-to-Leading Order FFN, 2 $GeV^2 \rightarrow 10000\ GeV^2$			
x	$xu_v(x,Q^2)$ (LH)	$xu_v(x,Q^2)$ (CB)	Rel. error
0.001	$5.7926\ 10^{-2}$	$5.7932\ 10^{-2}$	$1.1\ 10^{-4}$
0.01	$2.3026\ 10^{-1}$	$2.3025\ 10^{-1}$	$4.7\ 10^{-5}$
0.1	$5.5452\ 10^{-1}$	$5.5452\ 10^{-1}$	$2.4\ 10^{-6}$
0.3	$3.5393\ 10^{-1}$	$3.5394\ 10^{-1}$	$2.1\ 10^{-5}$
0.5	$1.2271\ 10^{-1}$	$1.2273\ 10^{-1}$	$1.5\ 10^{-4}$
0.7	$2.0429\ 10^{-2}$	$2.0427\ 10^{-2}$	$1.1\ 10^{-4}$
0.9	$3.6096\ 10^{-4}$	$3.6086\ 10^{-4}$	$3.0\ 10^{-4}$

Non-Singlet PDF

- First application of the method:

$$F_2^{NS}(x,Q^2) \equiv 2\left(F_2^p - F_2^d\right)(x,Q^2)$$
$$= \frac{x}{6}\left(u+\bar{u}-d-\bar{d}\right)(x,Q^2) \equiv xq^{NS}(x,Q^2)$$

- In the NS sector $\int_0^1 dx\, \Gamma(x) = 1$ to all orders:

$$q^{NS}(x,Q^2) = q^{NS}(x,Q_0^2) + \int_x^1 \frac{dy}{y}\Gamma(y)\left(q^{NS}\left(\frac{x}{y},Q_0^2\right) - yq^{NS}(x,Q_0^2)\right)$$
$$- \; q^{NS}(x,Q_0^2)\int_0^x dy\, \Gamma(y)$$

$x = 1$

- Structure functions $\longrightarrow$ artificial points:

$$F_2(x = 1, Q^2) = 0$$

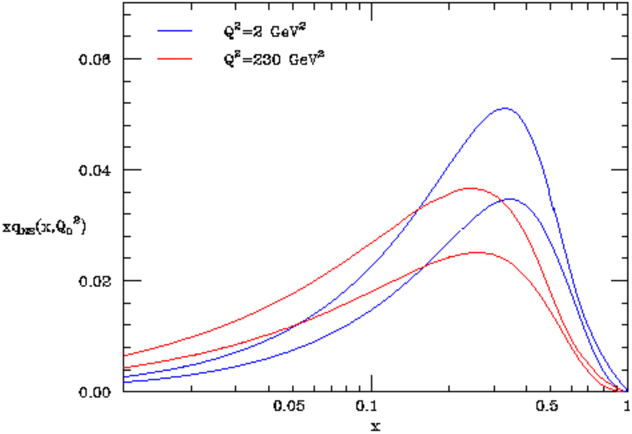
- Parton Distributions $\longrightarrow$ Lagrange multiplier:

$$\chi^2 = \chi^2 + \lambda \left(q^{NS}(x = 1, Q_0^2) \right)^2 , \quad \lambda = 10^6$$

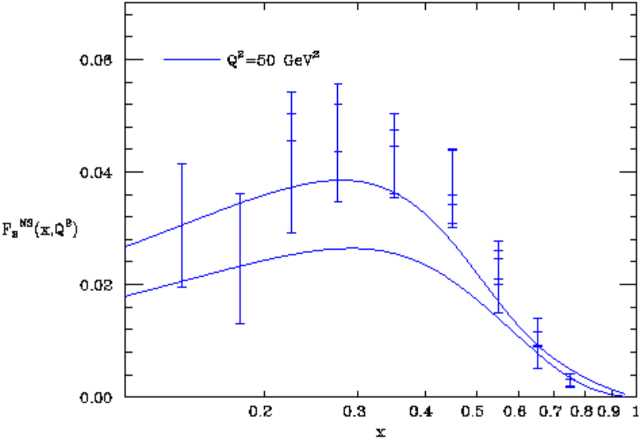
Details

- Experimental data: NMC (202 pts) and BCDMS (254 pts)
- Kinematical cuts: $Q^2 \geq 4\ GeV^2$, $W^2 \geq 6.25\ GeV^2$
- Neural network architecture: 2-2-2-1 (15 params.)
- Strong coupling: $\alpha_s\left(M_Z^2\right) = 0.1182$
- 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: 50
- Time: $\sim$ 2 hours per replica on CPU 2.6 GHz
  ($\sim$ 1000 GA generations)

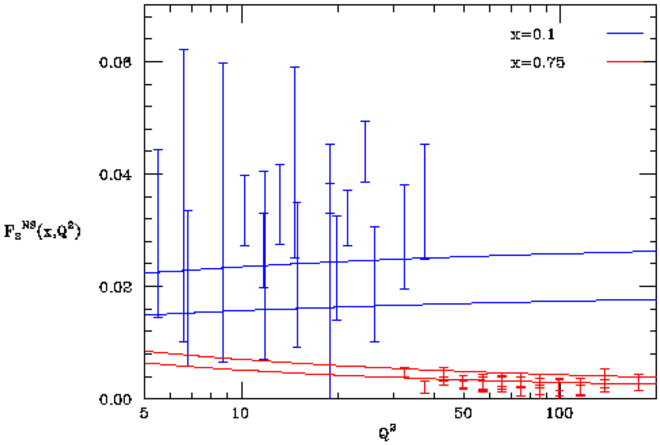
NLO $q^{NS}(x, Q^2)$



F_2^{NS} vs. x



F_2^{NS} vs. Q^2



Summary

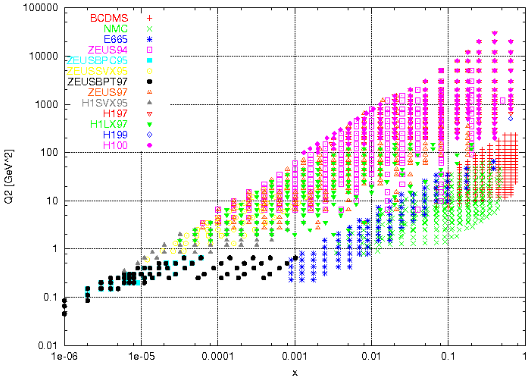
- Unbiased determination of structure functions with faithful estimation of uncertainties
- Successful implementation of neural parton fitting at NLO (errors and correlations are forthcoming)

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

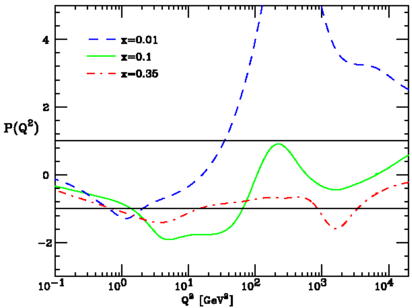
SF: Extras I

Kinematic of data used in the NN fit:



SF: Extras II

Comparison between the old and the new NN fit of F_2^p:



$$P(x,Q^2) \equiv \frac{F_2^{new}(x,Q^2) - F_2^{old}(x,Q^2)}{\sqrt{\sigma^2_{old}(x,Q^2) + \sigma^2_{new}(x,Q^2)}}$$

PDF: Extras

Mellin Inversion with the Fixed Talbot algorithm:

$$f(t) = \frac{1}{2\pi i}\int_C ds\ e^{ts}\tilde{f}(s), \quad t = -\ln x$$

$$s(\theta) = r\theta\left(\cot\theta + i\right), \quad -\pi \leq \theta \leq \pi$$

$$f(t) = \frac{r}{\pi}\int_0^{\pi} d\theta\ Re\left[\exp(ts(\theta))\tilde{f}(s(\theta))(1 + i\sigma(\theta))\right]$$

$$\sigma(\theta) = \theta + (\theta\cot\theta - 1)cot\theta$$

$$f(t, M) = \frac{r}{M}\left[\frac{1}{2}\tilde{f}(r)e^{rt} + \sum_{k=1}^{M-1} Re\left[\exp(ts(\theta_k))\tilde{f}(s(\theta_k))(1 + i\sigma(\theta_k))\right]\right]$$

$$r = \frac{2M}{5t}, \qquad \theta_k = \frac{k\pi}{M}$$