Neural Network Determination of Parton Distribution Functions

Maria Ubiali

School of Physics, University of Edinburgh

Moriond QCD and High Energy Interactions
La Thuile, 13 March 2008

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{5}, M.U.\textsuperscript{1}

\textsuperscript{1} PPT Group, School of Physics, University of Edinburgh
\textsuperscript{2} Dipartimento di Fisica, Universit\`a di Milano
\textsuperscript{3} Physikalisches Institut, Albert-Ludwigs-Universit\`at Freiburg
\textsuperscript{4} Departament d’Estructura i Constituents de la Matèria, Universitat de Barcelona
\textsuperscript{5} LPTHE, Université Paris VI et Paris VII
Given a set of data points we must determine a set of functions with error.

We need an error band in the space of functions, i.e. a probability density $P[q(x)]$ in the space of PDFs, $q(x)$. For an observable $F$ depending on PDFs:

$$\langle F[q(x)] \rangle = \int [Dq] F[q(x)] P[q(x)]$$

Standard approach, choose a basis of functions and project PDFs on it: the $\infty$-dimensional space of function reduces to a finite-dimensional space of parameters.

Issues:

- Non trivial propagation of errors: non-gaussian errors and incompatible data.
- The error associated to the choice of parametrisation is difficult to assess.
NNPDF approach

\[ \langle F[q(x)] \rangle = \int [Dq] F[q(x)] P[q(x)] \quad \rightarrow \quad \langle F[q(x)] \rangle = \frac{1}{N_{\text{rep}}} \sum_{k=1}^{N_{\text{rep}}} F[q^{(k)(\text{net})}(x)] \]
Monte Carlo determination of errors:

After fitting, the error of an observable depending on PDFs →

\[
\sigma_{\mathcal{F}[q(x)]} = \sqrt{\langle \mathcal{F}[q(x)]^2 \rangle - \langle \mathcal{F}[q(x)] \rangle^2}
\]

Neural Networks as redundant and unbiased parametrisation of PDFs:

* Each neuron receives input from neurons in preceding layer.
* Activation determined by weights and thresholds according to a non linear function:

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

Dynamical stopping criterion in order to fit data and not statistical noise.

* Divide data in two sets: training and validation.
* Minimisation is performed only on the training set. The validation $\chi^2$ for the set is computed.
* When the training $\chi^2$ still decreases while the validation $\chi^2$ stops decreasing → STOP.
- **NLO fit.**
- **ZM-VFN** treatment of heavy quarks.
- All DIS data included.
- **Flavor Assumptions:**
  - Symmetric strange sea $s(x) = \bar{s}(x)$
  - Strange sea proportional to non-strange sea $\bar{s}(x) = \frac{C}{2}(\bar{u}(x) + \bar{d}(x))$ ($C = 0.5$)

- **Parametrization of 4+1 combinations of PDFs at $Q_0^2 = 2$ GeV$^2$:**
  - Singlet : $\Sigma(x)$ $\mapsto NN_{\Sigma}(x)$ 2-3-2-1 20 pars
  - Gluon : $g(x)$ $\mapsto NN_{g}(x)$ 2-3-2-1 20 pars
  - Total valence : $V(x) = u_V(x) + d_V(x)$ $\mapsto NN_V(x)$ 2-3-2-1 20 pars
  - Non-singlet triplet : $T_3(x)$ $\mapsto NN_{T3}(x)$ 2-3-2-1 20 pars
  - Sea asymmetry : $\Delta_S(x) = \bar{d}(x) - \bar{u}(x)$ $\mapsto NN_{\Delta}(x)$ 2-3-1 13 pars

**93 parameters**
Some Very Preliminary Results

**Singlet PDF - Log scale**

**Gluon PDF - Log scale**

**ValTot PDF - Lin scale**

**SeaAsymm PDF - Lin scale**

**Triplet PDF - Lin scale**

**F_2 Proton**

**Neutrino cross section**

**CC reduced xsec**

pdfs

observables

Maria Ubiali
Neural PDFs for LHC
Conclusions

- Standard approaches to PDFs fitting might lead to underestimation of errors associated with parton densities.
- Combination of Monte Carlo techniques and Neural Networks as unbiased interpolating functions has proved to be a fast and robust alternative method.
- A non singlet fit has been published [hep-ph/0701127] and a full DIS fit will be published very soon.