Neural network determination of the non-singlet quark distribution

NNPDF Collaboration

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

[1] Particle Physics Theory Group, School of Physics, University of Edinburgh
[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

Motivation

- We want an accurate extraction of PDFs from data.
- We want to assess which is error due to PDFs when we evaluate an observable.

Some open problems:

- Errors combination and propagation from data to parameters and from parameters to observables is not trivial
- Theoretical bias due to the choice of a parametrization is difficult to assess (effects can be large if data are not precise or hardly compatible)

Motivation

- We want an accurate extraction of PDFs from data.
- We want to assess which is error due to PDFs when we evaluate an observable.

Some open problems:

- Errors combination and propagation from data to parameters and from parameters to observables is not trivial
- Theoretical bias due to the choice of a parametrization is difficult to assess (effects can be large if data are not precise or hardly compatible)

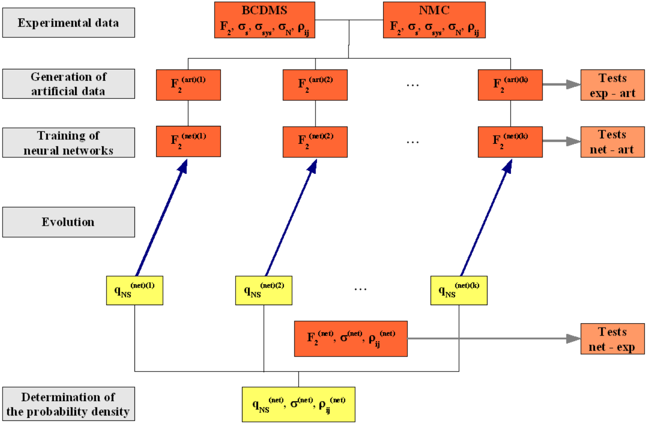
Motivation

- We want an accurate extraction of PDFs from data.
- We want to assess which is error due to PDFs when we evaluate an observable.

Some open problems:

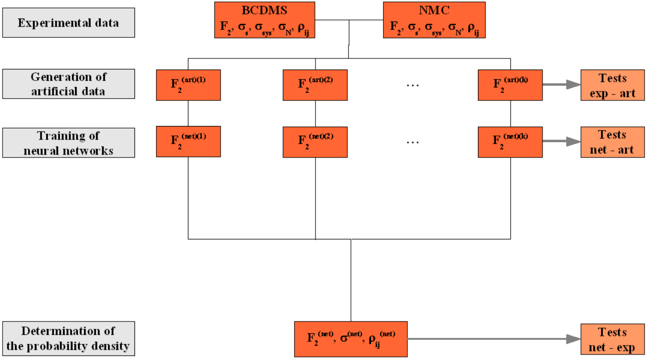
- Errors combination and propagation from data to parameters and from parameters to observables is not trivial
- Theoretical bias due to the choice of a parametrization is difficult to assess (effects can be large if data are not precise or hardly compatible)

The NNPDF approach



The NNPDF approach

[S. Forte et al., hep-ph/0204232 - L. Del Debbio et al., hep-ph/0501067]



Faithful error propagation: Data → Parametrization

- Monte Carlo sampling of data (generation of replicas of experimental data)

$$F_i^{(art)(k)} = \left(1 + r_N^{(k)} \sigma_N\right) \left[F_i^{(exp)} + r_i^s \sigma_i^{stat} + \sum_{l=1}^{N_{sys}} r^{l,(k)} \sigma_i^{sys,l} \right]$$

where σ_i are the experimantal errors, and r_i are random numbers choosen accordingly to the experimental correlation matrix.

Faithful error propagation: Parametrization → Observables

- Expectation values:

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

- Errors:

$$\sigma_{\mathcal{F}[g(x)]} = \sqrt{\left\langle \mathcal{F}\left[g(x)\right]^2 \right\rangle - \left\langle \mathcal{F}\left[g(x)\right]\right\rangle^2}$$

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

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

Unbiased parametrization

- A Neural network is trained over each MC replica.
- Neural networks are a class of algorithms very suitable to fit incomplete or noisy data [for HEP applications see ACAT 2005]
- Any continuous function can be uniformly approximated by a continuous neural network having only one internal layer, and with an arbitrary continuous sigmoid non-linearity [G. Cybenko (1989)].
- As an example, in a simple case (1-2-1) we have,

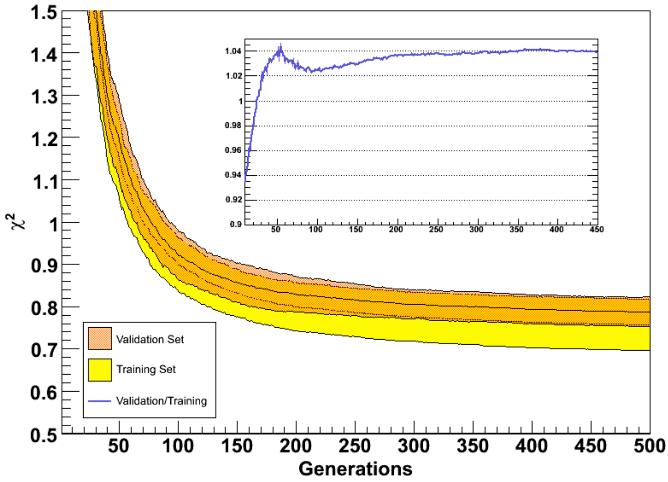
$$\xi_1^{(3)} = \frac{1}{1 + e^{\theta_1^{(3)} - \frac{\omega_{11}^{(2)}}{1+e^{\theta_1^{(2)} - \xi_1^{(1)}\omega_{11}^{(1)}}} - \frac{\omega_{12}^{(2)}}{1+e^{\theta_2^{(2)} - \xi_1^{(1)}\omega_{21}^{(1)}}}}}$$

Unbiased parametrization

- A Neural network is trained over each MC replica.
- Neural networks are a class of algorithms very suitable to fit incomplete or noisy data [for HEP applications see ACAT 2005]
- Any continuous function can be uniformly approximated by a continuous neural network having only one internal layer, and with an arbitrary continuous sigmoid non-linearity [G. Cybenko (1989)].
- As an example, in a simple case (1-2-1) we have,

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

Minimization strategy



A new evolution framework

- We want Mellin space evolution (numerically efficient):

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

- We do not want complex neural networks:

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

- The evolved PDF is given by

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

Some details

- Fitted quantity: $F_2^p(x, Q^2) - F_2^d(x, Q^2)$
- Kinematical cuts: $Q^2 \geq 3\ GeV^2$, $W^2 \geq 6.25\ GeV^2$
- Experimental data: NMC (229 pts) and BCDMS (254 pts)
- TMC: Georgi-Politzer, F_2 integral evaluated with NN F_2
- Neural network architecture: 2-5-3-1 (37 parameters).
- $x\,q(x, Q_0^2 = 2\ GeV^2) = NN(x, \log x)(1-x)^a$
- ZM-VFN: $m_c = 1.4 GeV$, $m_b = 4.5 GeV$, $m_t = 175 GeV$

Delivery

- # MC reps: 1000
- Strong coupling: $\alpha_s\left(M_Z^2\right) = 0.118 \pm 0.002$
- Perturbative order: LO, NLO, NNLO
- With $\alpha_s = 0.118$ we have:

χ^2/d.o.f.	Total	NMC	BCDMS
LO	0.94	0.93	0.95
NLO	0.95	0.92	0.97
NNLO	0.96	0.92	1.00

Non-singlet QCD analysis of $F_2(x, Q^2)$ up to NNLO

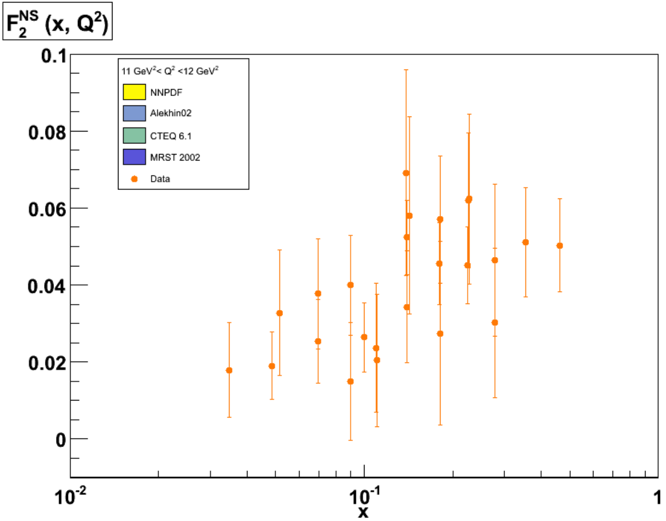
[M. Glück, E. Reya and C. Shuck, hep-ph/0604116]

By following the same analysis of BBG (see Guffanti's talk), they find

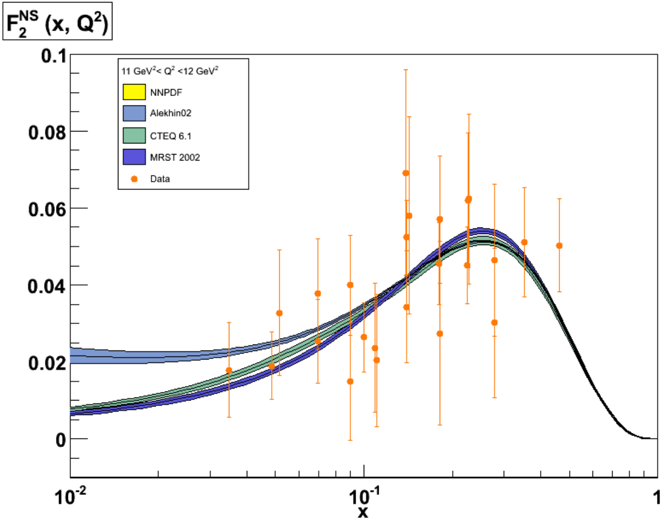
χ^2	Total	+HT
LO	0.98	0.82
NLO	0.93	0.83
NNLO	0.89	0.83

- If the data are not abundant enough the effect of the evolution can be reabsorbed in the initial condition.
- Once $\chi^2 \leq 1$ there isn't much room left for HT.

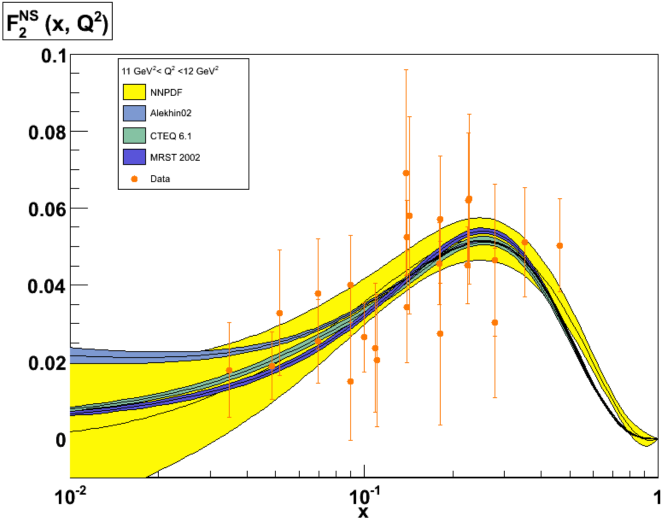
Reconstructing F_2^{NS} @ NLO with errors



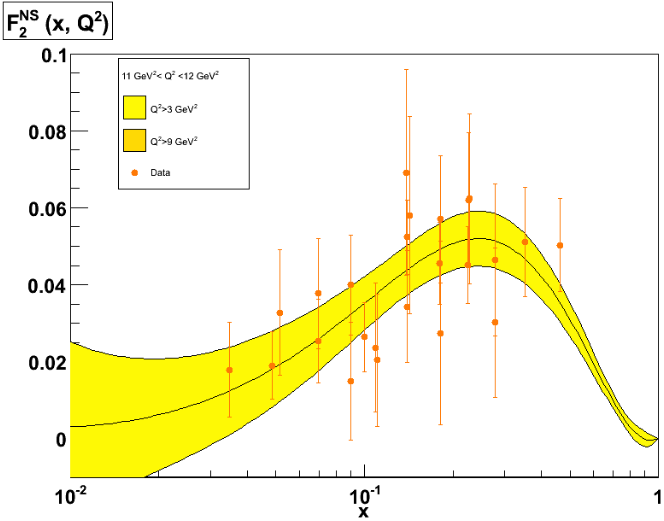
Reconstructing F_2^{NS} @ NLO with errors



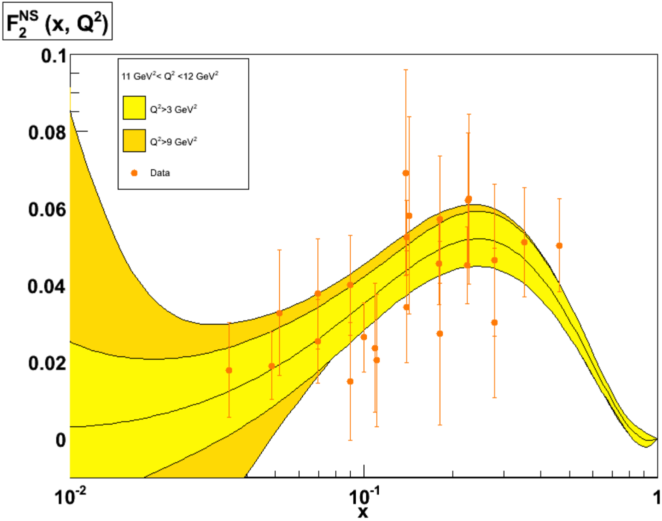
Reconstructing F_2^{NS} @ NLO with errors



The effect of kinematical cuts (only 100 nets)



The effect of kinematical cuts (only 100 nets)



Re-evaluation of the Gottfried sum rule

[R. Abbate and S. Forte, hep-ph/0511231]

- NMC:
$$S_G(0.004 < x < 0.8,\ 4\ \mathrm{GeV}^2) = 0.2281 \pm 0.0201$$
- NNPDF:
$$S_G(0.004 < x < 0.8,\ 4\ \mathrm{GeV}^2) = 0.2281 \pm 0.0437$$
- The two estimations perfectly agree for all $x_{min} < x < 0.8$ ranges, but the for the smallest $x_{min} = 0.004$.
- NMC uncertainty at the boundary of the measured region is evaluated assuming that the error is linear across the bins, and this results in an underestimation of the error on the last bin.
- The inclusion of the (assumed/unknown) small-x contribution yields
$$S_G(1.5\ \mathrm{GeV}^2 < Q^2 < 4.5\ \mathrm{GeV}^2) = 0.244 \pm 0.045$$

Re-evaluation of the Gottfried sum rule

[R. Abbate and S. Forte, hep-ph/0511231]

- NMC:

$$S_G(0.004 < x < 0.8,\ 4\ \mathrm{GeV}^2) = 0.2281 \pm 0.0201$$

- NNPDF:

$$S_G(0.004 < x < 0.8,\ 4\ \mathrm{GeV}^2) = 0.2281 \pm 0.0437$$

- The two estimations perfectly agree for all $x_{min} < x < 0.8$ ranges, but the for the smallest $x_{min} = 0.004$.
- NMC uncertainty at the boundary of the measured region is evaluated assuming that the error is linear across the bins, and this results in an underestimation of the error on the last bin.
- The inclusion of the (assumed/unknown) small-x contribution yields

$$S_G(1.5\ \mathrm{GeV}^2 < \mathrm{Q}^2 < 4.5\ \mathrm{GeV}^2) = 0.244 \pm 0.045$$

Perspectives

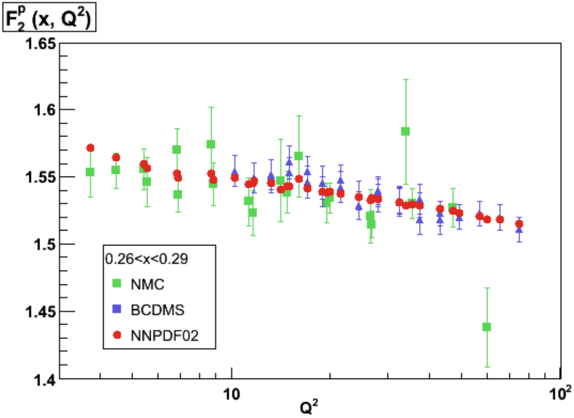
- Construct full set of NNPDF parton distributions from all available data (starting from DIS)
- Assess impact of uncertainties for relevant observables at LHC
- Make formalism compatible with standard interfaces (LHAPDF, PDFLIB) $\rightarrow$ NNPDF partons available for use in Monte Carlo generators

Minimization with a Genetic Algorithm

1. Set the parameters randomly.
2. Make clones of the set of parameters.
3. Mutate randomly each clone.
4. Evaluate χ^2 for all the clones.
5. Select clones with the lowest χ^2.
6. Back to 2, till $\chi^2 \sim 1$.

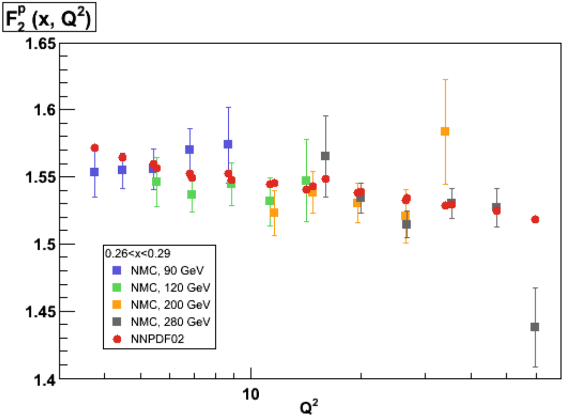
Incompatible data

[S. Forte et al., hep-ph/0204232]



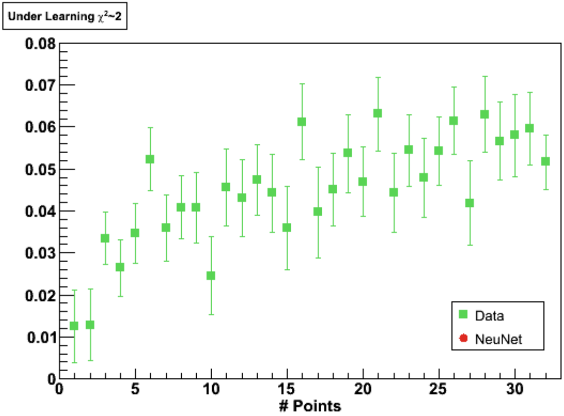
Incompatible data

[S. Forte et al., hep-ph/0204232]



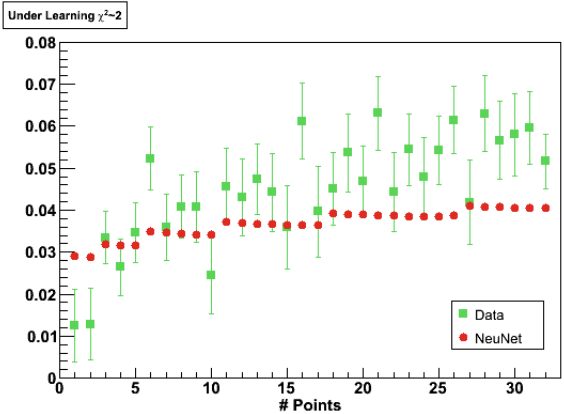
Learning of data

[S. Forte et al., hep-ph/0204232]



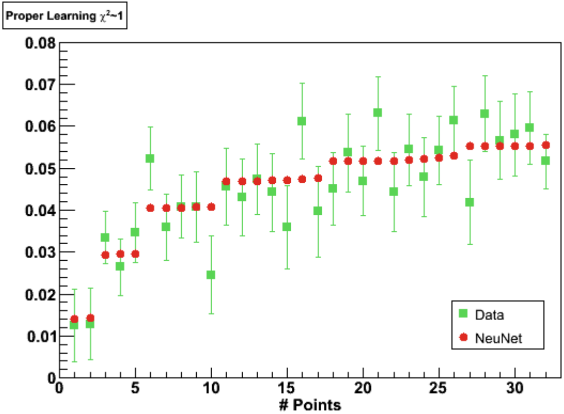
Learning of data

[S. Forte et al., hep-ph/0204232]



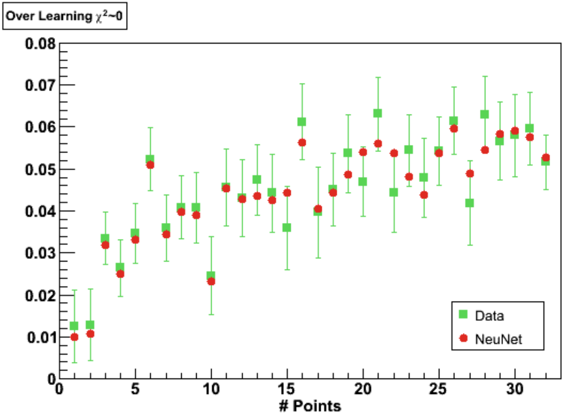
Learning of data

[S. Forte et al., hep-ph/0204232]



Learning of data

[S. Forte et al., hep-ph/0204232]



Reconstructing F_2^{NS}