A First Unbiased Determination of the Spin Content of the Proton

Juan Rojo

Instituto Nazionale di Fisica Nucleare, Sezione di Milano

Deep Inelastic Scattering 2010
Firenze, 21/04/2010

Work in collaboration

NNPDF collaboration

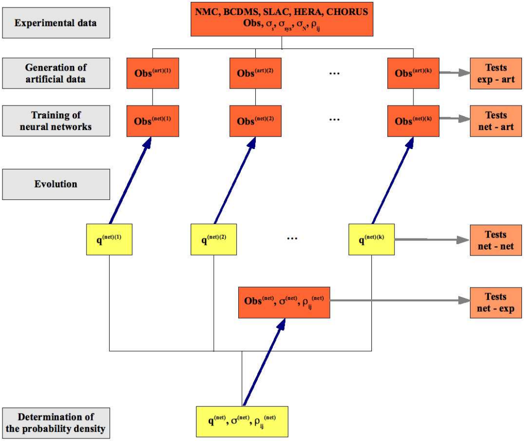
R.D.Ball[1], L.Del Debbio[1], S.Forte[2], A.Guffanti[3], J.I.Latorre[4], G.Ridolfi[5], J. R.[2], M. Ubiali[1]

[1] PPT Group, School of Physics, University of Edinburgh
[2] Dipartimento di Fisica, Università di Milano and INFN, Sezione Milano
[3] Physikalisches Institut, Albert-Ludwigs-Universität Freiburg
[4] Departament d'Estructura i Constituents de la Matèria, Universitat de Barcelona
[5] Dipartimento di Fisica, Università di Genoa and INFN, Sezione Genoa

Outline

Outline

The NNPDF approach to PDF determination



Monte Carlo errors

Non-gaussian errors and non trivial error propagation.

Neural Networks

Avoid bias from a restrictive fixed functional form.

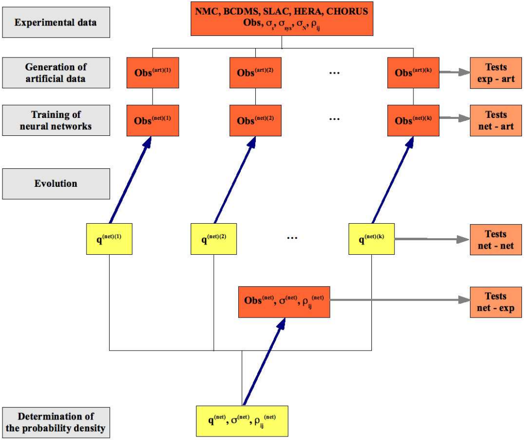
Dynamical Stopping

No looking for absolute minimum but learning from data.

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

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

Step 1: Monte Carlo Errors



Monte Carlo errors

Non-gaussian errors and non trivial error propagation.

Monte Carlo sample

Generate a $N_{\rm rep}$ Monte Carlo sets of artificial data, or "pseudo-data" of the original N_{data} data points

$$F_i^{(art)(k)}(x_p, Q_p^2) \equiv F_{i,p}^{(art)(k)} \qquad \begin{array}{rcl} i & = & 1, \ldots, N_{\rm data} \\ k & = & 1, \ldots, N_{\rm rep} \end{array}$$

Multi-gaussian distribution centered on each data point:

$$F_{i,p}^{(art)(k)} = S_{p,N}^{(k)} F_{i,p}^{\rm exp} \left(1 + r_p^{(k)} \sigma_p^{\rm stat} + \sum_{j=1}^{N_{\rm sys}} r_{p,j}^{(k)} \sigma_{p,j}^{\rm sys} \right)$$

If two points have correlated systematic uncertainties

$$r_{p,j}^{(k)} = r_{p',j}^{(k)}$$

Correlations are properly taken into account.

Monte Carlo Errors

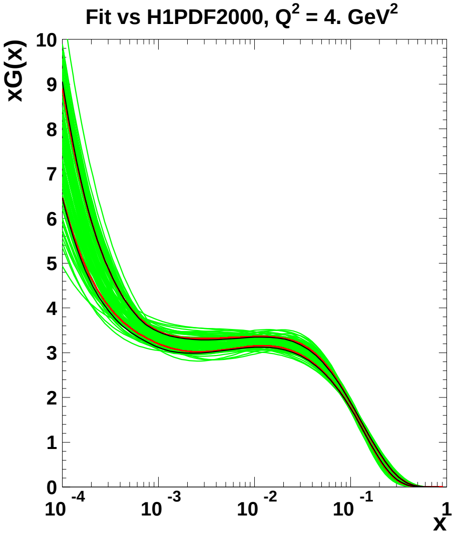
For each replica $^{(k)}$ of the experimental data we fit a set of independent PDFs

Ensemble of fitted replicas of PDFs: representation of the probability distribution in the space of PDFs

Uncertainties, central values and any other statistical property (e. g. correlations) of the PDFs (or any function of them) can be evaluated using standard statistical methods.

$$
\begin{aligned}
\langle \mathcal{F}[f(x)] \rangle &= \frac{1}{N_{\rm rep}} \sum_{k=1}^{N_{\rm rep}} \mathcal{F}[f^{(k)(\rm net)}(x)] \\
\sigma_{\mathcal{F}[f(x)]} &= \sqrt{\langle \mathcal{F}[f(x)]^2 \rangle - \langle \mathcal{F}[f(x)] \rangle^2} \\
\rho[f_a(x_1, Q_1^2), f_b(x_2, Q_2^2)] &= \frac{\langle f_a(x_1, Q_1^2) f_b(x_2, Q_2^2) \rangle - \langle f_a(x_1, Q_1^2) \rangle \langle f_b(x_2, Q_2^2) \rangle}{\sigma_a(x_1, Q_1^2) \sigma_b(x_2, Q_2^2)}
\end{aligned}
$$

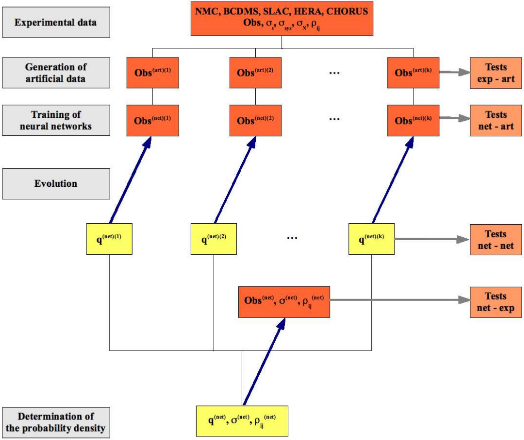
Monte Carlo vs. Hessian PDF uncertainties



HERA-LHC 2009 PDF benchmarks

- H1PDF2000 fit done with Hessian method and with Monte Carlo method
- The standard deviation of the 100 PDF replicas - MC method - in perfect agreement with **Hessian errors with** $\Delta\chi^2 = 1$
- The Monte Carlo method to estimate PDF uncertainties reproduces Hessian result when global χ^2 is quadratic

Step 2: Neural Network as unbiased and redundant parametrization

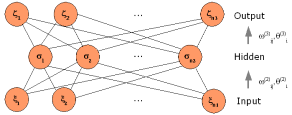


Neural Networks

Avoid bias from a restrictive fixed functional form.

What are neural networks?

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



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

$$\xi_i = g\Big(\sum_j \omega_{ij}\xi_j - \theta_i\Big), \qquad g(x) = \frac{1}{1+e^{-x}}$$

In a simple case (1-2-1) we have,

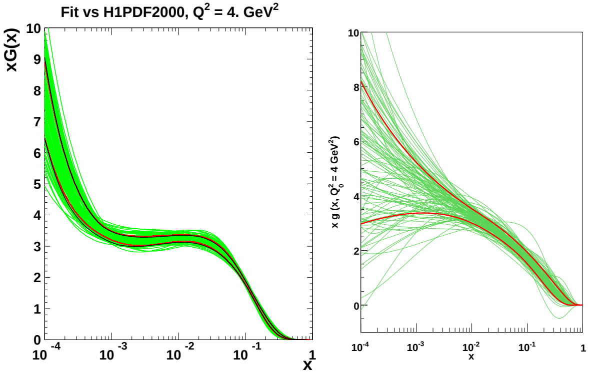
$$\xi_1^{(3)} = \frac{1}{1 + e^{\theta_1^{(3)} - \frac{\omega_{11}^{(2)}}{1+e^{\theta_1^{(2)} - \xi_1^{(1)}\omega_{11}^{(1)}}} - \frac{\omega_{12}^{(2)}}{1+e^{\theta_2^{(2)} - \xi_1^{(1)}\omega_{21}^{(1)}}}}}$$

7 parameters

...Just a convenient functional form which provides a redundant and flexible parametrization.

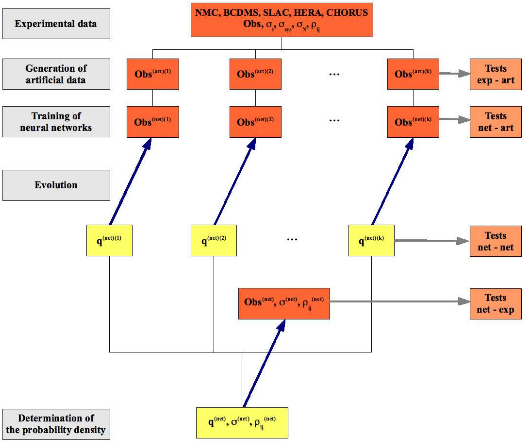
We want the best fit to be independent of any assumption made on the parametrization.

Simple functional froms vs. NeuralNets



- PDFs parametrized with simple functional forms $\longrightarrow$ May result in systematic underestimation of PDF uncertainties
- The use of an universal interpolant like Artificial Neural Networks removes any bias from choice of input PDF functional form $q_i(x, Q^2)$

Step 3: Training and dynamical stopping

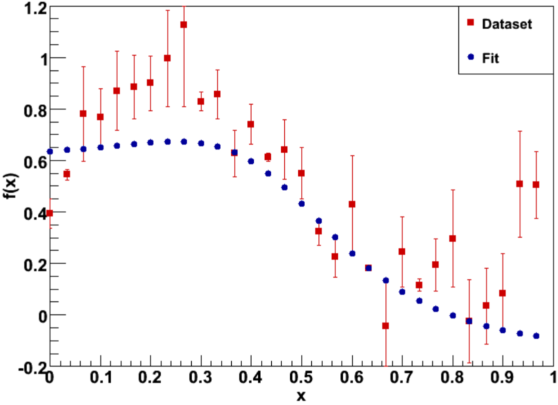


Dynamical Stopping

No looking for absolute minimum but learning from data.

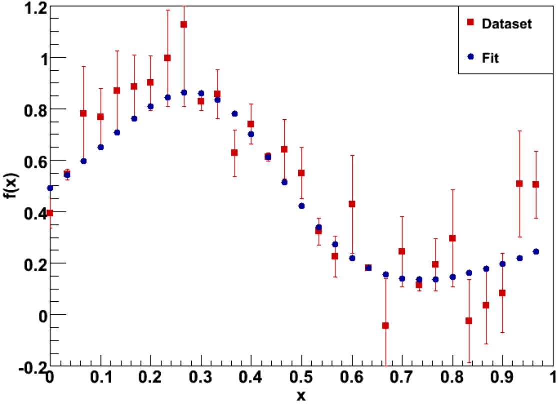
Neural network learning

We need to train to avoid under-learning ...



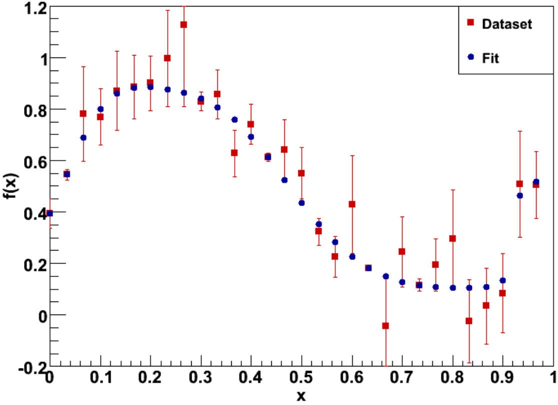
Neural network learning

... until we arrive to proper learning ....



Neural network learning

... but be careful to avoid overlearning!



Fitting Strategy

Our fitting strategy is very different from that used by other collaborations: instead of a set of basis functions with a small number of parameters, we have an unbiased basis of functions parameterized by a very large and redundant set of parameters.

DSSV,AAC,LSS

$\mathcal{O}$(10-20) parm

NNPDFpol

$\mathcal{O}$(200) parm

Not trivial because ...

A redundant parametrization might adapt not only to physical behavior but also to random statistical fluctuations of data.

Ingredients of fitting procedure

1. Flexible and redundant parametrization
2. Genetic Algorithm minimization
3. Dynamical stopping criterion

Genetic Algorithm

- Set neural network parameters randomly.
- Make clones of the parameter vector and mutate them.
- Evaluate the figure of merit for each clone:

Error function

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

${\rm cov}^{(t_0)}$ defined from an experimental covariance matrix which to include normalization errors with the t_0 method (arXiv:0912.2276)

$${\rm cov}_{ij}^{(t_0)} = \sigma_i^{\rm stat,2} F_i^{(\rm exp)2} + \sum_k^{N_{\rm sys}} \sigma_i^{\rm sys,k} \sigma_j^{\rm sys,k} F_i^{(\rm exp)} F_j^{(\rm exp)} + \sigma_i^{\rm N} \sigma_j^{\rm N} F_i^{(0)} F_j^{(0)},$$

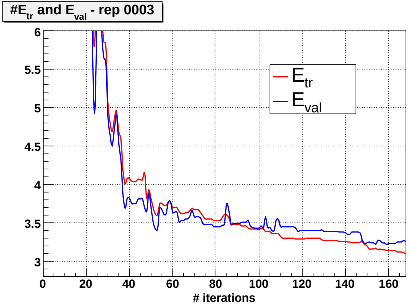
- Select the best ones and iterate the procedure until a stability is reached.

Dynamical Stopping Criterion

* Genetic Algorithms are monotonically decreasing by construction.
* The best fit is not given by the absolute minimum.
* The best fit is given by an optimal training beyond which the figure of merit improves only because we are fitting statistical noise of the data.

Cross-validation method

* Divide data in two sets: training and validation.
* Random division for each replica ($f_t = f_v = 0.5$).
* Minimization is performed only on the training set. The validation χ^2 for the set is computed.
* When the training χ^2 still decreases while the validation χ^2 stops decreasing → STOP.

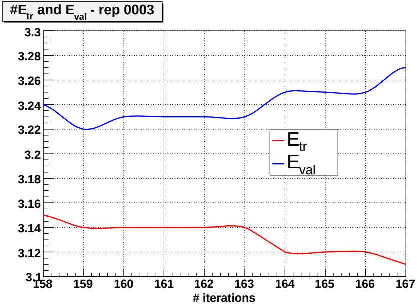


Dynamical Stopping Criterion

* Genetic Algorithms are monotonically decreasing by construction.
* The best fit is not given by the absolute minimum.
* The best fit is given by an optimal training beyond which the figure of merit improves only because we are fitting statistical noise of the data.

Cross-validation method

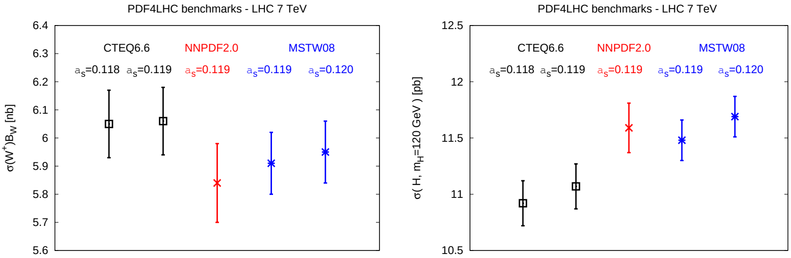
* Divide data in two sets: training and validation.
* Random division for each replica ($f_t = f_v = 0.5$).
* Minimization is performed only on the training set. The validation χ^2 for the set is computed.
* When the training χ^2 still decreases while the validation χ^2 stops decreasing → STOP.



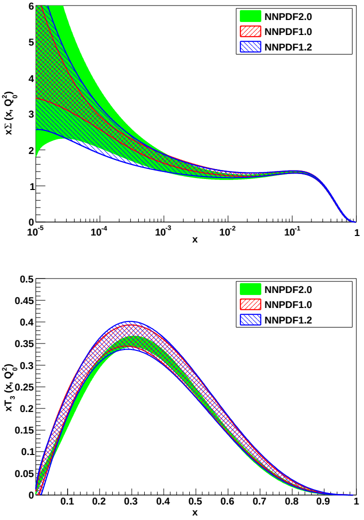
Outline

The NNPDF (Unpolarized) Roadmap

1. NNPDF1.0 (arxiv:arXiv:0808.1231) → Determination of PDFs from inclusive DIS data, $N_{\mathrm{PDF}} = 5$
2. NNPDF1.2 (arxiv:0906.1958) → Determination of PDFs from inclusive DIS data and neutrino charm production for $s^{\pm}(x, Q^2)$, $N_{\mathrm{PDF}} = 7$
3. NNPDF2.0 (arxiv:1002.4407) → Determination of PDFs from DIS and hadronic data (Drell-Yan, Weak boson production, inclusive jets)
4. In preparation: NNLO PDFs, PDFs for MCs, implementation of HQ effects



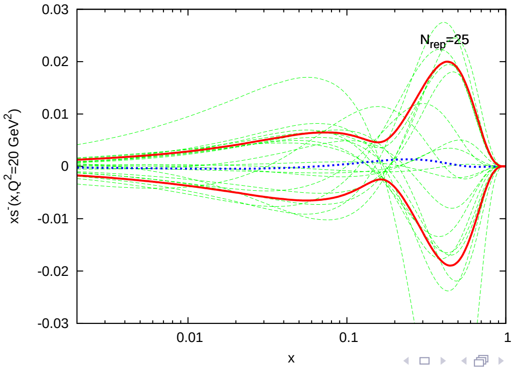
The NNPDF (Unpolarized) Roadmap



Strange asymmetry PDF: $s^{-}(x, Q^2)$ in NNPDF1.2

- No theoretical constraints on $s^{-}(x, Q_0^2)$ apart from valence sum rule
- At least one crossing required by sum rule, but some replicas have two crossings
- Compare with more restrictive parametrizations

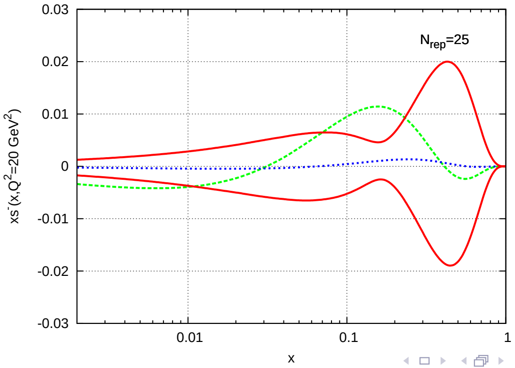
$$xs^{-}_{\rm mstw} = A_{-} x^{0.2} (1-x)^{\eta_{-}} (1 - x/x_0)$$



Strange asymmetry PDF: $s^{-}(x, Q^2)$ in NNPDF1.2

- No theoretical constraints on $s^{-}(x, Q_0^2)$ apart from valence sum rule
- At least one crossing required by sum rule, but some replicas have two crossings
- Compare with more restrictive parametrizations

$$xs^{-}_{\mathrm{mstw}} = A_{-}x^{0.2}\left(1-x\right)^{\eta_{-}}\left(1-x/x_0\right)$$



Impact on NuTeV anomaly

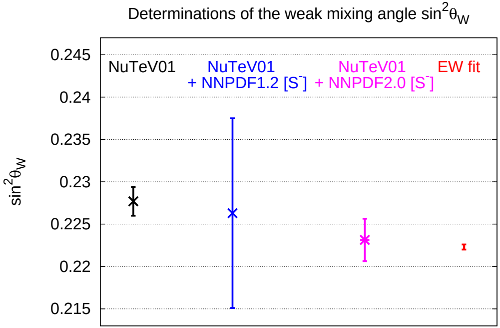
NuTeV anomaly: Discrepancy ($\geq 3\sigma$) between indirect (global fit)and direct (NuTeV neutrino scattering) determinations of $\sin^2\theta_W$

EW fit

$\sin^2\theta_W = 0.2223 \pm 0.0003$

NuTeV

$\sin^2\theta_W = 0.2277 \pm 0.0017$



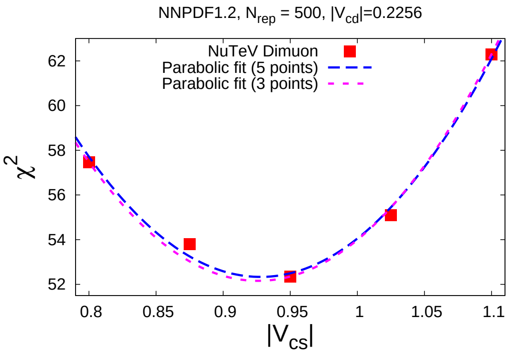
NuTeV + NNPDF1.2 $[S^-]$

$\sin^2\theta_W = 0.2263 \pm 0.0017^{\mathrm{exp}} \pm 0.0107^{\mathrm{PDFs}}$

NuTeV + NNPDF2.0 $[S^-]$

$\sin^2\theta_W = 0.22314 \pm 0.00251$

Direct $|V_{cs}|$ determination



CKM global fit

$V_{cs} = 0.97334 \pm 0.00023, \; \Delta V_{cs} \sim 0.02\%$

Direct determination-D and B decays

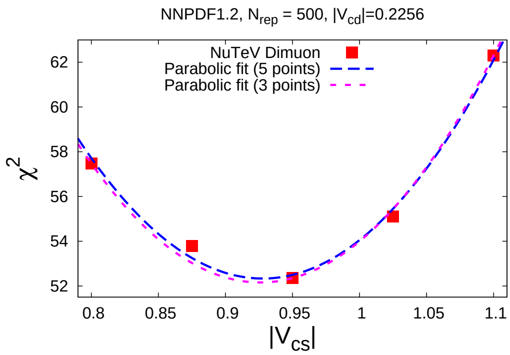
$V_{cs} = 1.04 \pm 0.06, \qquad \Delta V_{cs} \sim 6\%$

Direct det from $\nu-$DIS (CCFR)

$V_{cs} \geq 0.74 \quad (90\%\mathrm{CL})$

[PDG, Amsler et al, Phys. Lett. B67(2008) 1.]

Direct $|V_{cs}|$ determination



Direct det from ν−DIS (CCFR)

$$V_{cs} \geq 0.74 \quad (90\%\mathrm{CL})$$

Direct determination-D and B decays

$$V_{cs} = 1.04 \pm 0.06, \qquad \Delta V_{cs} \sim 6\%$$

Direct det NNPDF1.2

$$V_{cs} = 0.96 \pm 0.07, \; \Delta V_{cd} \sim 7\%$$

[PDG, Amsler et al, Phys. Lett. B67(2008) 1.]

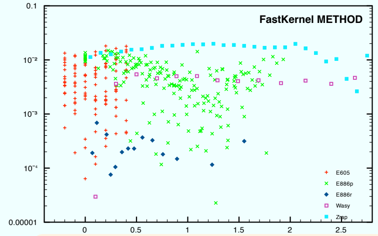
- $|V_{cs}|$ determination from neutrino DIS affected by $s^+(x)$ uncertainties
- Unbiased parametrizations for PDFs allow to discriminate variations in $s^+(x)$ from variations in CKM matrix elements

NNPDF2.0

FastKernel

- NLO computation of hadronic observables too slow for parton global fits.
- MSTW08 and CTEQ include Drell-Yan NLO as (local) K-factors rescaling the LO cross section

* NNPDF2.0 includes full NLO calculation of hadronic observables.
* Use available fastNLO interface for jet inclusive cross-sections. [hep-ph/0609285]
* Built up our own **FastKernel** computation of DY observables.



- Both PDFs evolution and double convolution sped up
- Use high-orders polynomial interpolation
- Precompute all Green Functions

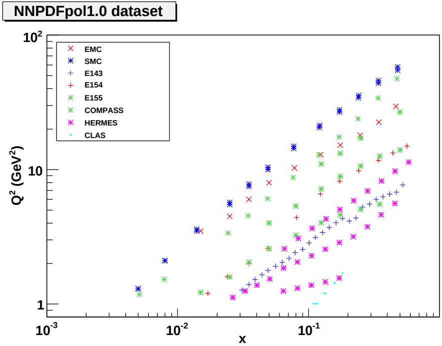
$$\int_{x_{0,1}}^{1} dx_1 \int_{x_{0,2}}^{1} dx_2\ f_a(x_1) f_b(x_2) C^{ab}(x_1, x_2) \rightarrow \sum_{\alpha,\beta=1}^{N_x} f_a(x_{1,\alpha}) f_b(x_{2,\beta}) \int_{x_{0,1}}^{1} dx_1 \int_{x_{0,2}}^{1} dx_2\ \mathcal{I}^{(\alpha,\beta)}(x_1, x_2) C^{ab}(x_1, x_2)$$

Outline

Experimental data

- Inclusive polarized structure function data $g_1(x, Q^2)$ on proton, deuteron and neutron targets from spin asymmetries

$$g_1(x, Q^2) = A_1(x, Q^2)\frac{F_2(x, Q^2)}{2x(1 + R(x, Q^2))}\left(1 + \gamma^2\right) , \qquad \gamma^2 \equiv \frac{4M_N^2 x^2}{Q^2}$$



$$g_1^p = \left[\Delta C_{2,q} \otimes \sum e_i^2 \Delta q_i^+ + \Delta C_{2,g} \otimes \Delta g\right]$$

Kinematical cuts:

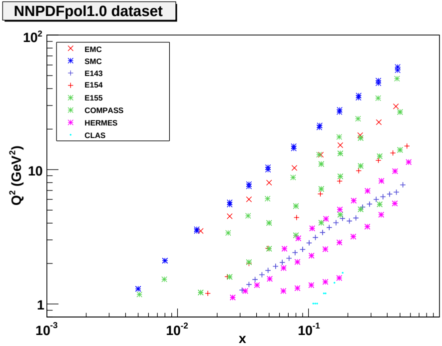
- $Q^2 \geq 1$ GeV^2
- $W^2 \geq 6.25$ GeV^2

Minimize impact of Higher Twist terms while maximizing amount of data included (C. Simolo, Ph.D. Thesis, arXiv:0807.1501)

Experimental data

- Inclusive polarized structure function data $g_1(x, Q^2)$ on proton, deuteron and neutron targets from spin asymmetries

$$g_1(x, Q^2) = A_1(x, Q^2)\frac{F_2(x, Q^2)}{2x(1+R(x, Q^2))}\left(1+\gamma^2\right) \ , \qquad \gamma^2 \equiv \frac{4M_N^2 x^2}{Q^2}$$



Theoretical constraints:

* Sum rules

$$\left[\Delta T_3(Q_0^2)\right] \equiv \int_0^1 dx\ \Delta T_3(x, Q_0^2) = a_3 \ ,$$

$$\left[\Delta T_8(Q_0^2)\right] \equiv \int_0^1 dx\ \Delta T_8(x, Q_0^2) = a_8 \ ,$$

* Positivity of polarized PDFs $\longrightarrow$ Constraints on polarized SFs: $|g_1(x, Q^2)| \leq F_1(x, Q^2)$ for all targets

- F_1^p, F_1^d, F_1^n computed consistently from NNPDF1.0

Input PDF basis

Polarized PDFs are parametrized at $Q_0^2 = 1$ GeV^2 in the basis:

- Singlet $\Delta\Sigma(x) \equiv \sum_{i=1}^{n_f} (\Delta q_i(x) + \Delta\bar{q}_i(x))$,
- Triplet $\Delta T_3(x) \equiv (\Delta u(x) + \Delta\bar{u}(x)) - (\Delta d(x) + \Delta\bar{d}(x))$,
- Octet $\Delta T_8(x) \equiv (\Delta u(x) + \Delta\bar{u}(x)) + (\Delta d(x) + \Delta\bar{d}(x)) - 2\,(\Delta s(x) + \Delta\bar{s}(x))$,
- Gluon $\Delta g(x)$.

PDFs are parametrized with Artificial Neural Networks

$$
\begin{aligned}
\Delta\Sigma(x, Q_0^2) &= (1-x)^{m_{\Delta\Sigma}} x^{-n_{\Delta\Sigma}} \mathrm{NN}_{\Delta\Sigma}(x)\ , \\
\Delta T_3(x, Q_0^2) &= A_{\Delta T_3}(1-x)^{m_{\Delta T_3}} x^{-n_{\Delta T_3}} \mathrm{NN}_{\Delta T_3}(x)\ , \\
\Delta T_8(x, Q_0^2) &= A_{\Delta T_8}(1-x)^{m_{\Delta T_8}} x^{-n_{\Delta T_8}} \mathrm{NN}_{\Delta T_8}(x)\ , \\
\Delta g(x, Q_0^2) &= (1-x)^{m_{\Delta g}} x^{-n_{\Delta g}} \mathrm{NN}_{\Delta g}(x)\ .
\end{aligned}
$$

Preprocessing makes learning more efficient

$A_{\Delta T_3}, A_{\Delta T_8}$ determined from sum rules

Polarized PDF evolution

In Mellin space the DGLAP equations

$$\begin{aligned}\mu^2 \frac{\partial}{\partial \mu^2} \Delta q_{NS}^{\pm,v}(N,\mu^2) &= \Delta\gamma_{NS}^{\pm,v} q_{NS}^{\pm,v}(N,\mu^2) \\ \mu^2 \frac{\partial}{\partial \mu^2} \begin{pmatrix} \Delta\Sigma \\ \Delta g \end{pmatrix} (N,\mu^2) &= \begin{pmatrix} \Delta\gamma_{qq}\left(N,\alpha_s(Q^2)\right) & \Delta\gamma_{qg}\left(N,\alpha_s(Q^2)\right) \\ \Delta\gamma_{gq}\left(N,\alpha_s(Q^2)\right) & \Delta\gamma_{gg}\left(N,\alpha_s(Q^2)\right) \end{pmatrix} \begin{pmatrix} \Delta\Sigma \\ \Delta g \end{pmatrix}\end{aligned}$$

Can easily be solved analytically

$$\Delta q_{NS}^{\pm,v}(N,Q^2) = \Gamma_{NS}^{\pm,v}(N,a_s,a_0)\Delta q_{NS}^{\pm,v}(N,Q_0^2)\ , \quad a_s \equiv \alpha_s/2\pi$$

where at Next-to-Leading order

$$\Gamma_{\rm NS,NLO}^{\pm,v}(N,a_s,a_0) = \exp\left\{ \frac{U_1^{\pm,v}}{b_1} \ln\left(\frac{1+b_1 a_s}{1+b_1 a_0} \right) \right\} \left(\frac{a_s}{a_0} \right)^{-R_0^{\rm ns}} .$$

Polarized PDF evolution

$x-$space Polarized PDFs are typically obtained by Mellin inversion

$$\Delta q_{\mathrm{NS}}(x,Q^2) = \int_{c-i\infty}^{c+i\infty} \frac{dN}{2\pi i} x^{-N} \Delta q_{\mathrm{NS}}(N,Q^2)$$

Problem: Mellin space expressions $q_{\mathrm{NS}}(N,Q^2)$ only exist for simple parametrizations

FastKernel method $\rightarrow$ Mellin invert the evolution kernels only

$$\Gamma_{\mathrm{NS}}(x,a_s,a_0) = \int_{c-i\infty}^{c+i\infty} \frac{dN}{2\pi i} x^{-N} \Gamma_{\mathrm{NS}}(N,a_s,a_0)$$

$$\Delta q_{NS}(x,Q^2) = \int_x^1 \frac{dy}{y} \Gamma_{qq}(y,a_s,a_0)\, \Delta q_{NS}\left(\frac{x}{y},Q_0^2\right)$$

Now any $x-$space parametrization of $q(x,Q_0^2)$ is allowed

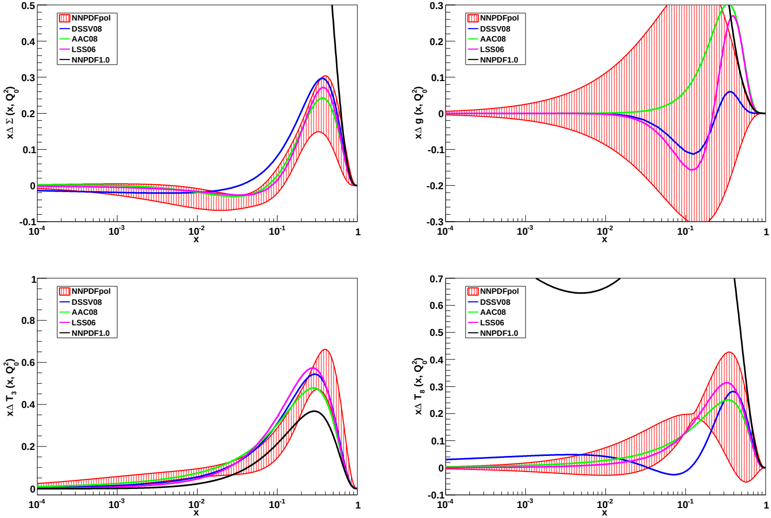
Polarized PDF evolution

NNPDF NLO polarized PDF evolution (**FastKernel method**) benchmarked with the Les Houches PDF benchmarks, G. Salam and A. Vogt, hep-ph/0511119

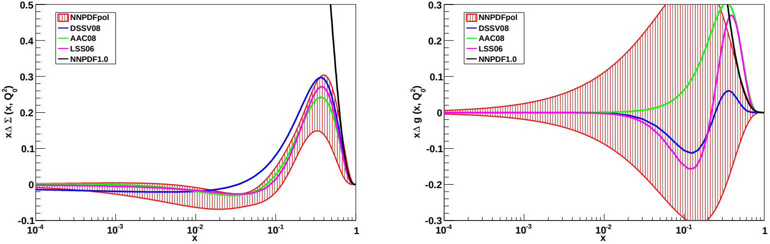
x	$\epsilon_{\rm rel}\left(\Delta u_V\right)$	$\epsilon_{\rm rel}\left(\Delta d_V\right)$	$\epsilon_{\rm rel}\left(\Delta\Sigma\right)$	$\epsilon_{\rm rel}\left(\Delta g\right)$
10^{-3}	$1.1\,10^{-4}$	$9.2\,10^{-5}$	$9.9\,10^{-5}$	$1.1\,10^{-4}$
10^{-2}	$1.4\,10^{-4}$	$1.9\,10^{-4}$	$3.5\,10^{-4}$	$9.3\,10^{-5}$
0.1	$1.2\,10^{-4}$	$1.6\,10^{-4}$	$5.4\,10^{-6}$	$1.7\,10^{-4}$
0.3	$2.3\,10^{-6}$	$1.1\,10^{-5}$	$7.5\,10^{-6}$	$1.7\,10^{-5}$
0.5	$5.6\,10^{-6}$	$9.6\,10^{-6}$	$1.6\,10^{-5}$	$2.5\,10^{-5}$
0.7	$1.2\,10^{-4}$	$9.2\,10^{-7}$	$1.6\,10^{-4}$	$7.8\,10^{-5}$
0.9	$3.5\,10^{-3}$	$1.1\,10^{-2}$	$4.1\,10^{-3}$	$7.8\,10^{-3}$

The `FastKernel` method leads to very fast and accurate PDF evolution also in the polarized case

Polarized NNPDFs - Very preliminary results!



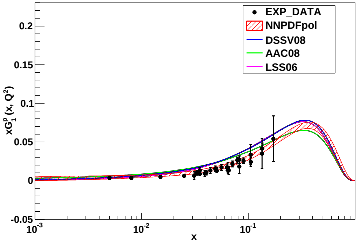
Polarized NNPDFs - Very preliminary results!

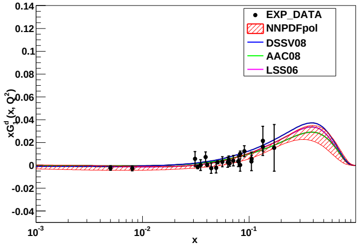


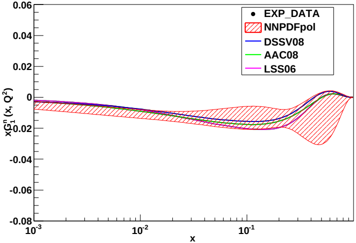
- The polarized gluon $\Delta g(x)$ is essentially unconstrained from inclusive polarized DIS only
- Reasonable agreement with other PPDF sets
- Much more work required for quantitative phenomenology (polarized moments, predictions for RHIC, ...)

Polarized NNPDFs - Very preliminary results!

Polarized structure functions at $Q^2 = 2\ \mathrm{GeV}^2$







Outline

Conclusions

- The NNPDF approach provides an unbiased determination of parton distributions with a faithful uncertainty estimation
- In the unpolarized case, the status of the art is NNPDF2.0, a global PDF analysis of DIS, DY and jet data with uses NLO QCD through without K-factors
- Once PDF uncertainties are statistically meaningful, important phenomenology is possible $\longrightarrow$ Solution to NuTeV anomaly, $|V_{cs}|$ determination, ...
- The NNPDF approach is being extended to the polarized case $\longrightarrow$ Very preliminary results indicate the uncertainties on $\Delta g(x, Q^2)$ from DIS-only are larger than hitherto thought

Thanks for your attention!

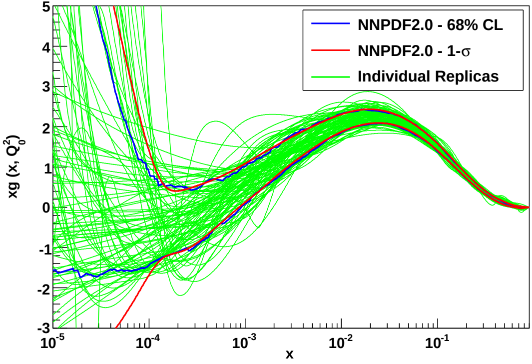
Conclusions

- The NNPDF approach provides an unbiased determination of parton distributions with a faithful uncertainty estimation
- In the unpolarized case, the status of the art is NNPDF2.0, a global PDF analysis of DIS, DY and jet data with uses NLO QCD through without K-factors
- Once PDF uncertainties are statistically meaningful, important phenomenology is possible $\longrightarrow$ Solution to NuTeV anomaly, $|V_{cs}|$ determination, ...
- The NNPDF approach is being extended to the polarized case $\longrightarrow$ Very preliminary results indicate the uncertainties on $\Delta g(x, Q^2)$ from DIS-only are larger than hitherto thought

Thanks for your attention!

Outline

Individual replicas vs Average quantities



Even though individual replicas may fluctuate significantly, average quantities such as central values and error bands are smooth inasmuch as stability is reached due to the number of replicas increasing.