NNPDF: results and comparisons

Luigi Del Debbio

luigi.del.debbio@ed.ac.uk

University of Edinburgh

NNPDF Collaboration, Nucl. Phys. B 809, 1 (2009), [arXiv:0808.1231], **NNPDF1.0**
arXiv:0811.2288, **NNPDF1.1**
in preparation, **NNPDF1.2**

Work in collaboration

The NNPDF Collaboration

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

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

Motivation

- robust input for analyses at LHC
- HERA-LHC, PDF4LHC workshops: problems with global fits
- determination of **unbiased** PDFs with **faithful** estimate of their error

$\hookrightarrow$ determine the probability density $\mathcal{P}[f_i(x)]$ in the space of PDFs $f_i(x)$

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

[kosower & giele 99]

NNPDF approach

- redundant parametrization
- MC determination of the measure $\mathcal{P}$
- take the experimental data at face value
- statistical estimators to assess size of systematics

Parametrization

NNPDF

$$f(x,Q_0^2) = Ax^n(1-x)^m \mathrm{NN}(x)$$

$\mathrm{NN}(x)$ depends on the weights w_{ij} (fit parameters)
architecture of the net: number of parameters, $O(100)$

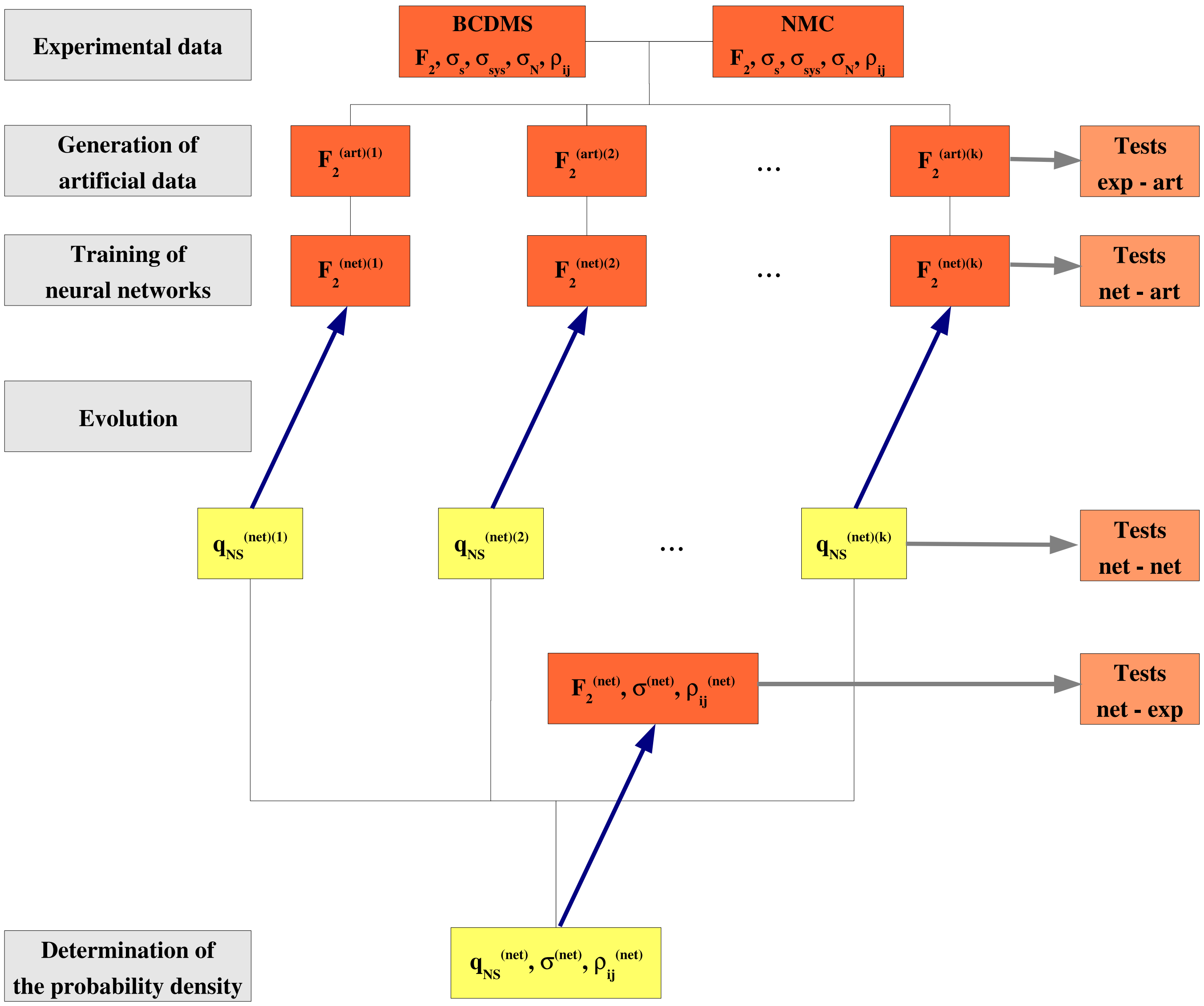
CTEQ/MSTW/Alekhin

$$f(x,Q_0^2) = Ax^\eta(1-x)^\xi(1+B\sqrt{x}+\ldots)$$

smaller number of parameters (30)

asymptotic behaviour fixed by a small number of them

The Neural Monte Carlo



The Monte Carlo ensemble

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

$\mathcal{F}$ can be any (LHC) observable that involves PDFs

in particular:

$$\bar{q}_\alpha = \frac{1}{N_{\mathrm{rep}}}\sum_{k=1}^{N_{\mathrm{rep}}}{q_\alpha}^{(\mathrm{net})(k)}$$

$$\mathrm{Var}\ q_\alpha = \frac{1}{N_{\mathrm{rep}}}\sum_{k=1}^{N_{\mathrm{rep}}}\left({q_\alpha}^{(\mathrm{net})(k)} - \bar{q}_\alpha\right)^2$$

Very different from the tolerance: $\Delta\chi^2 \simeq 50$ used by MSTW/CTEQ

PDFs uncertainties

- Monte Carlo prescription (NNPDF)

$$\sigma_{\mathcal{F}} = \left(\frac{N_{\rm set}}{N_{\rm set}-1} \left(\langle \mathcal{F}[\{f\}]^2 \rangle - \langle \mathcal{F}[\{f\}] \rangle^2 \right) \right)^{1/2}$$

- HEPDATA prescription (CTEQ and MRST/MSTW)

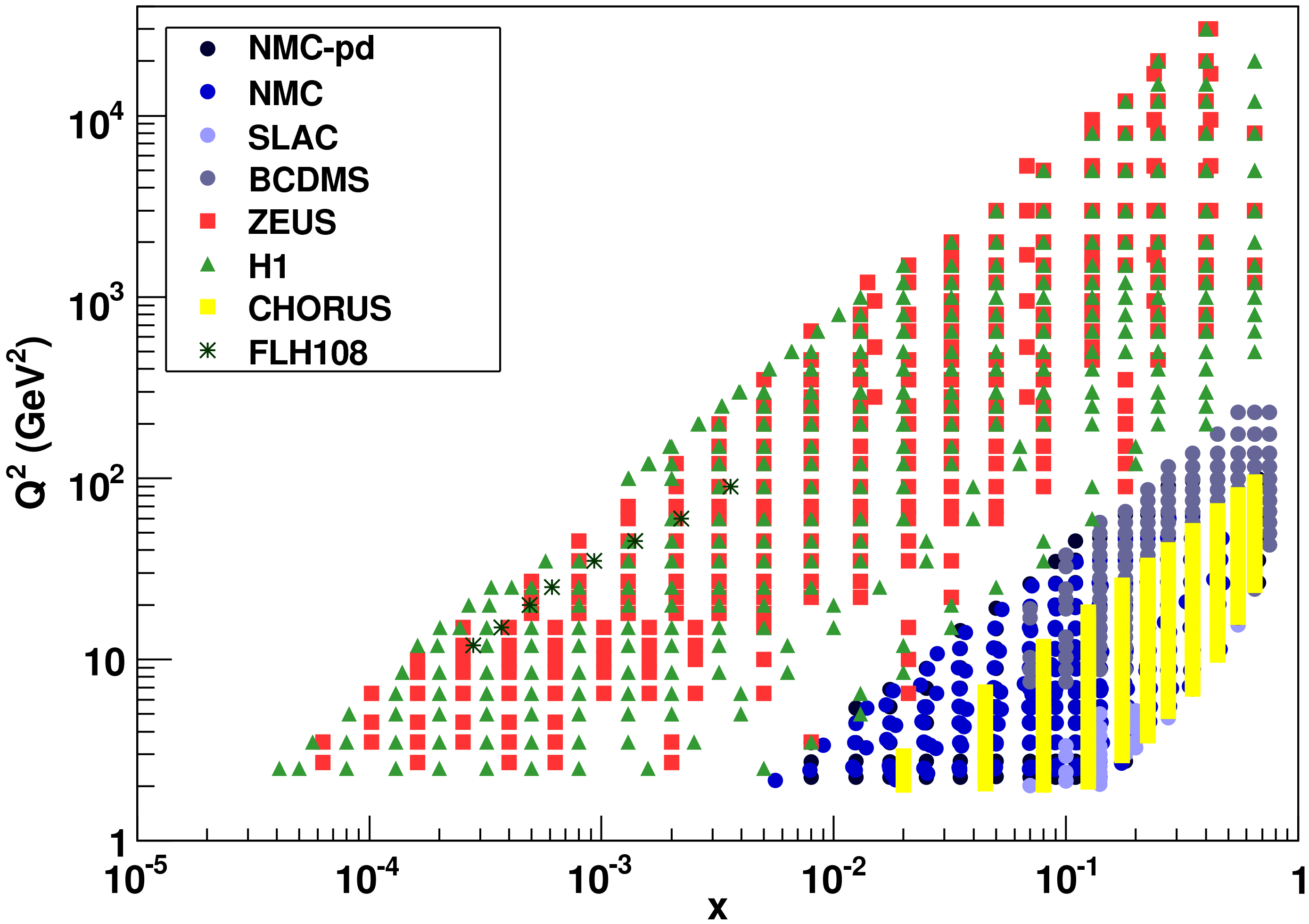
$$\sigma_{\mathcal{F}} = \frac{1}{2C_{90}} \left(\sum_{k=1}^{N_{\rm set}/2} \left(\mathcal{F}[\{f^{(2k-1)}\}] - \mathcal{F}[\{f^{(2k)}\}] \right)^2 \right)^{1/2} , \quad C_{90} = 1.64485$$

C_{90} accounts for the fact that the upper and lower parton sets correspond to 90% confidence levels rather than to one-σ uncertainties.

- HEPDATA* prescription (Alekhin)

$$\sigma_{\mathcal{F}} = \left(\sum_{k=1}^{N_{\rm set}} \left(\mathcal{F}[\{f^{(k)}\}] - \mathcal{F}[\{f^{(0)}\}] \right)^2 \right)^{1/2} .$$

Datasets



NNPDF1.0: full DIS fit + ZM-VFN + flavour assumptions, $Q^2 > 2\ \mathrm{GeV}^2, W^2 > 12\ \mathrm{GeV}^2$

NNPDF1.1: s and $\bar{s}$ distributions parametrized

NNPDF1.2: dimuon data added

Some more details

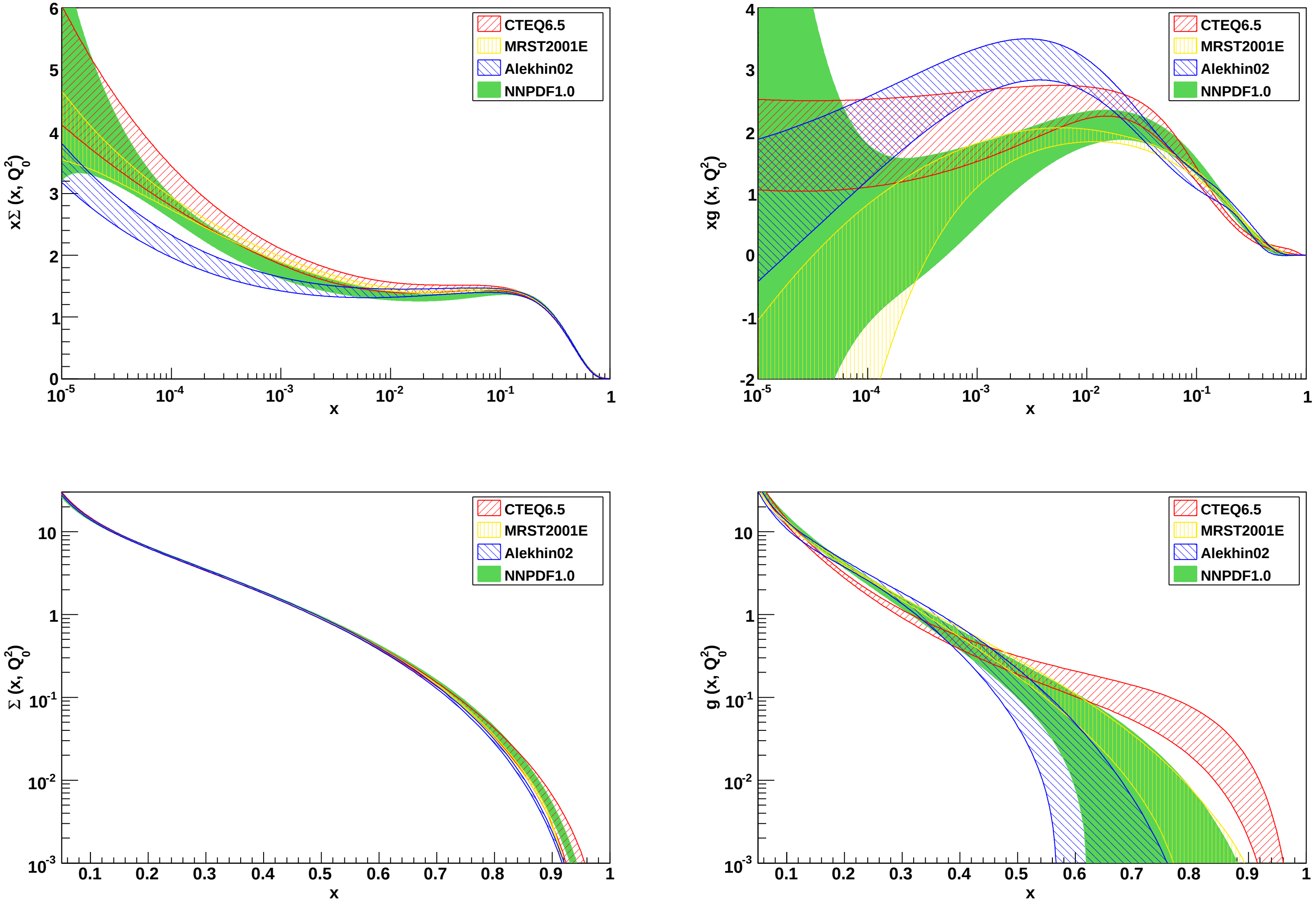
$$\chi^{2(k)}[\omega] = \sum_{i,j}^{N_{\rm dat}} (F_i^{({\rm art})(k)} - F_i^{({\rm net})(k)}) \left(\left(\overline{\rm cov}^{(k)} \right)^{-1} \right)_{ij} (F_j^{({\rm art})(k)} - F_j^{({\rm net})(k)})$$

- fully–correlated χ^2 is used
- $\overline{\rm cov}^{(k)}$ defined from an experimental covariance matrix which does not include normalization errors [D'Agostini 2003]

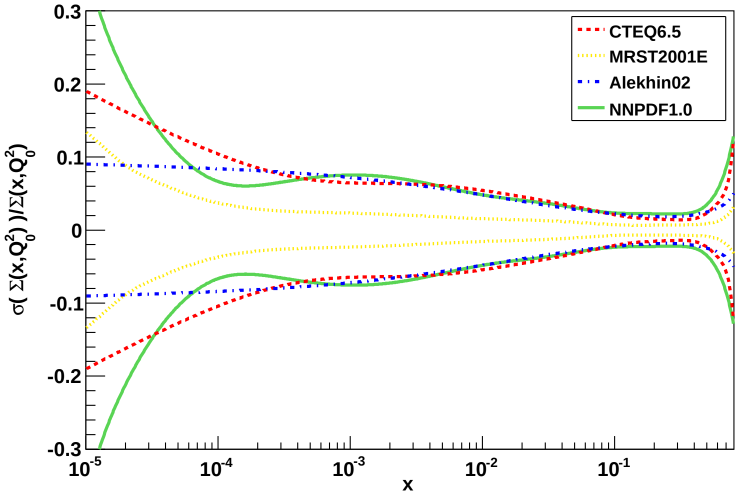
$$\overline{\rm cov}_{ij}^{(k)} = \overline{\rm cov}_{ij}^{(exp)} S_{i,N}^{(k)} S_{j,N}^{(k)}$$

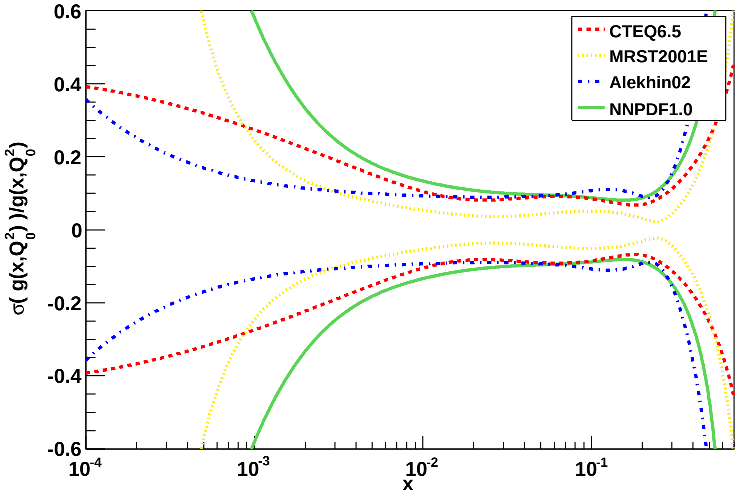
- $F_i^{({\rm net})}$ computed from PDFs using NLO, ZM-VFN scheme
- α_s is not fitted - kept fixed
- $N_{\rm rep} \simeq 100 \div 1000$ to obtain an accurate description of data

Results – Singlet



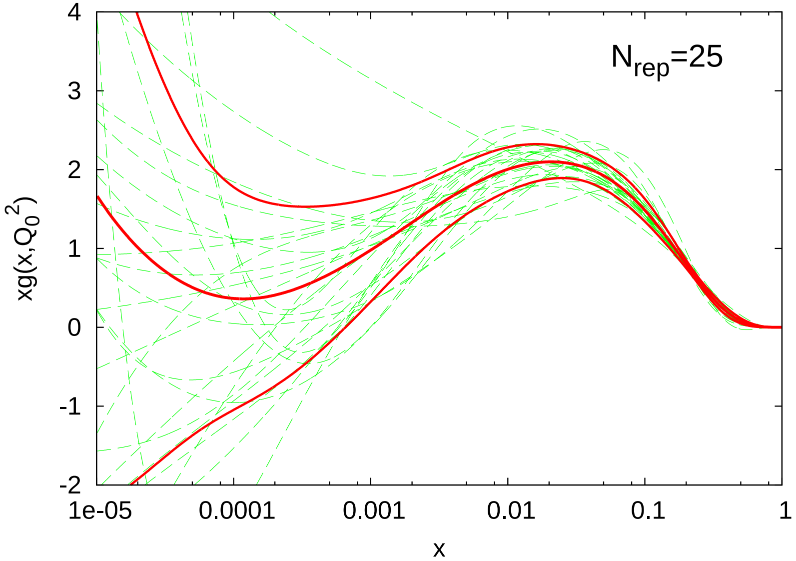
Relative Error

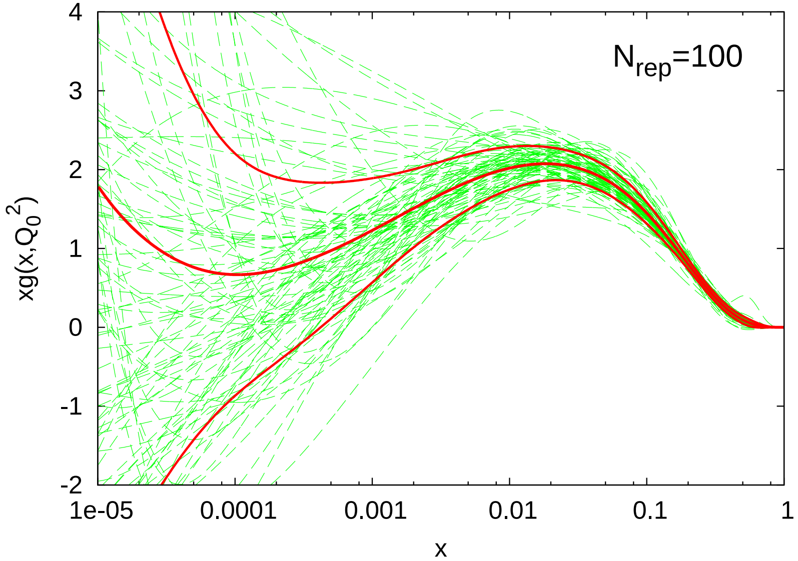


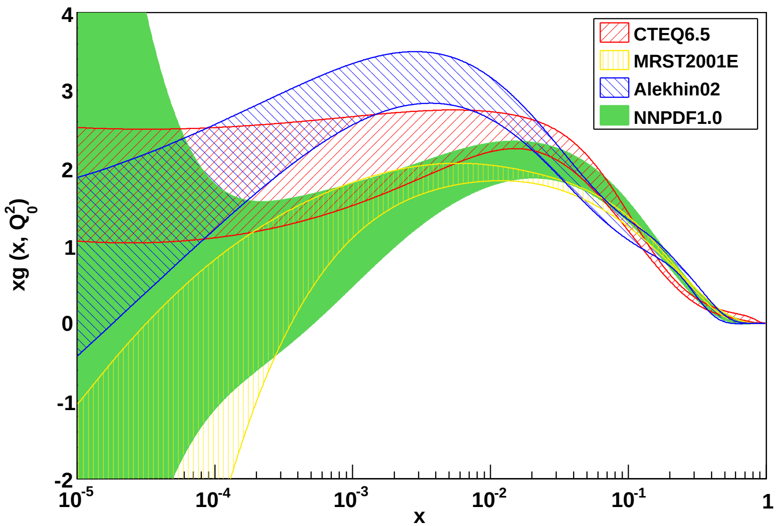


- NNPDF1.0 uncertainties correspond to a genuine 68% CL
- PDF error larger than other PDF sets in some regions (extrapolation), smaller in others (not artificially inflated by large $\Delta\chi^2 \sim 50/100$)
- in general close to CTEQ6.5 in data region

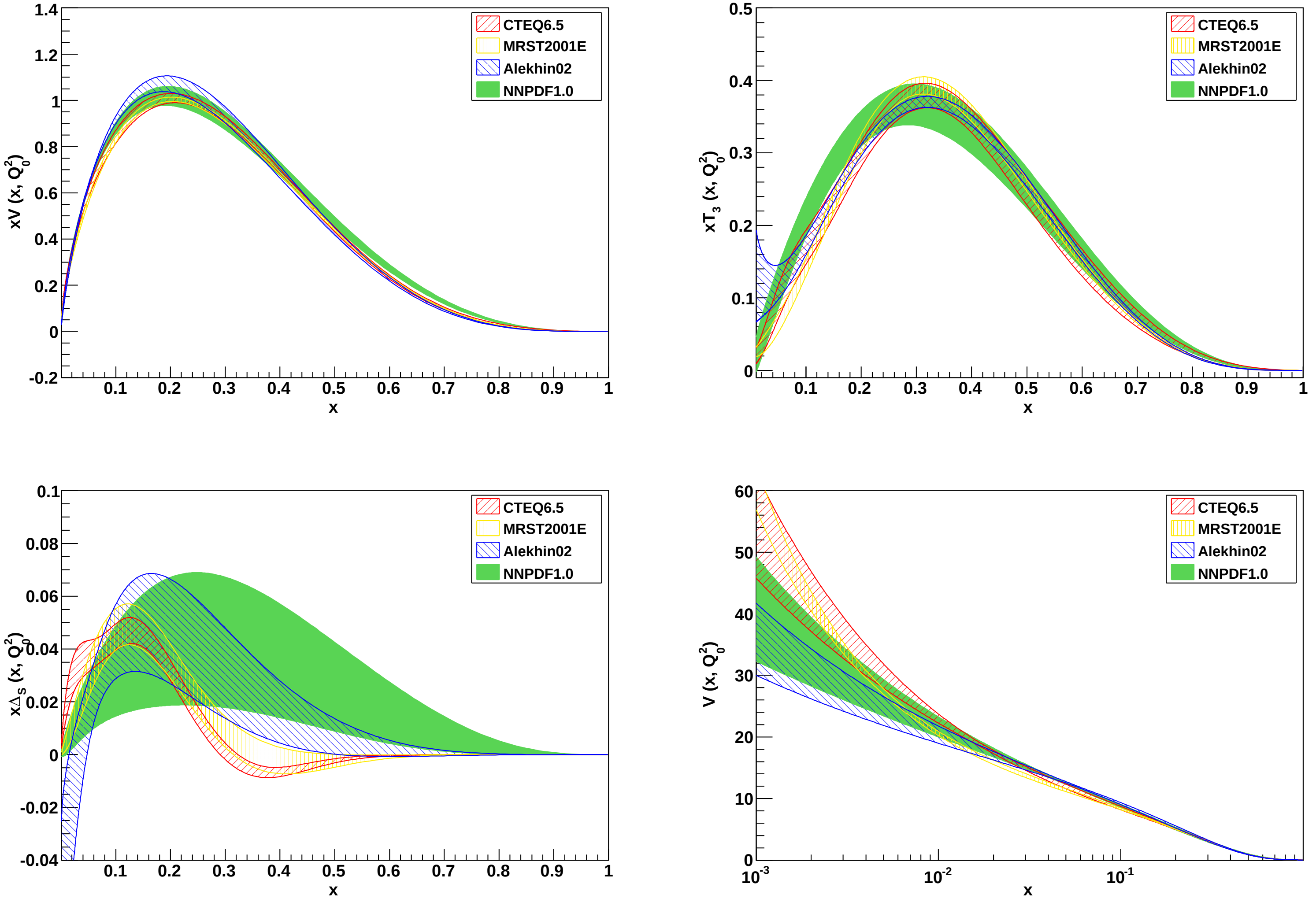
Monte Carlo sampling of PDFs



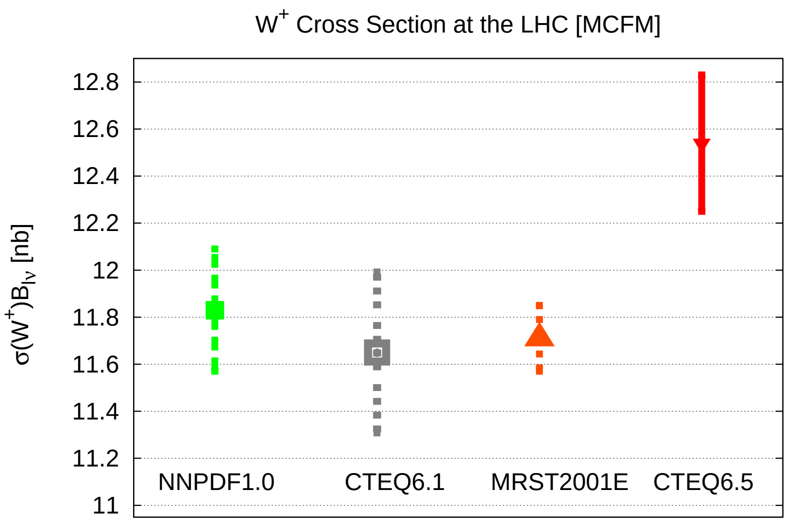


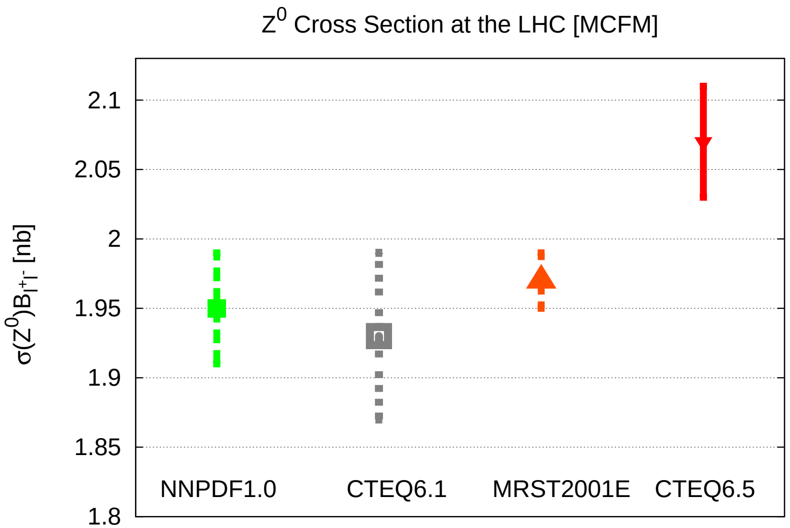


Results – Valence



W & Z cross-sections



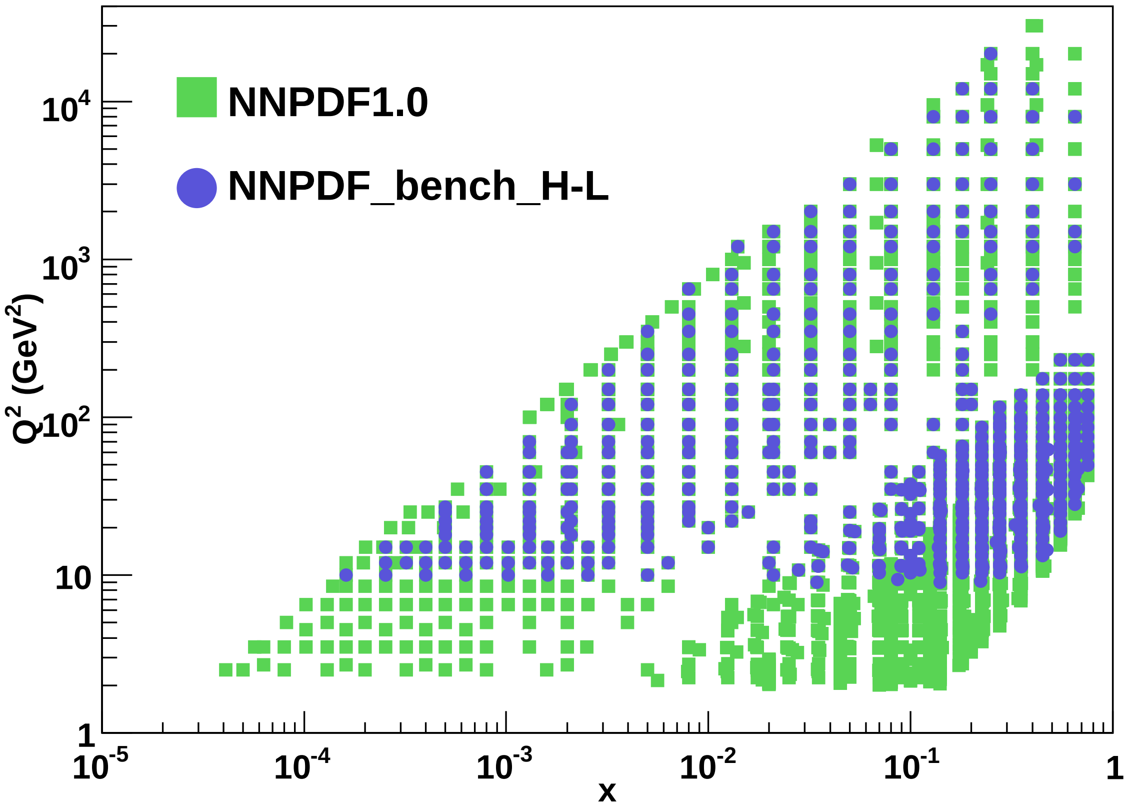


	$\sigma_{W^+}\mathcal{B}_{l^+\nu_l}$	$\Delta\sigma_{W^+}/\sigma_{W^+}$	$\sigma_{W^-}\mathcal{B}_{l^-\nu_l}$	$\Delta\sigma_{W^-}/\sigma_{W^-}$	$\sigma_Z\mathcal{B}_{l^+l^-}$	$\Delta\sigma_Z/\sigma_Z$
NNPDF1.0	11.83 ± 0.26	2.2%	8.41 ± 0.20	2.4%	1.95 ± 0.04	2.1%
CTEQ6.1	11.65 ± 0.34	2.9%	8.56 ± 0.26	3.0%	1.93 ± 0.06	3.1%
MRST01	11.71 ± 0.14	1.2%	8.70 ± 0.10	1.1%	1.97 ± 0.02	1.0%
CTEQ6.5	12.54 ± 0.29	2.3%	9.19 ± 0.22	2.4%	2.07 ± 0.04	1.9%

Comparing PDFs

HERA-LHC benchmark

Benchmark PDF fit to a reduced consistent set of DIS data [hep-ph/0511119]

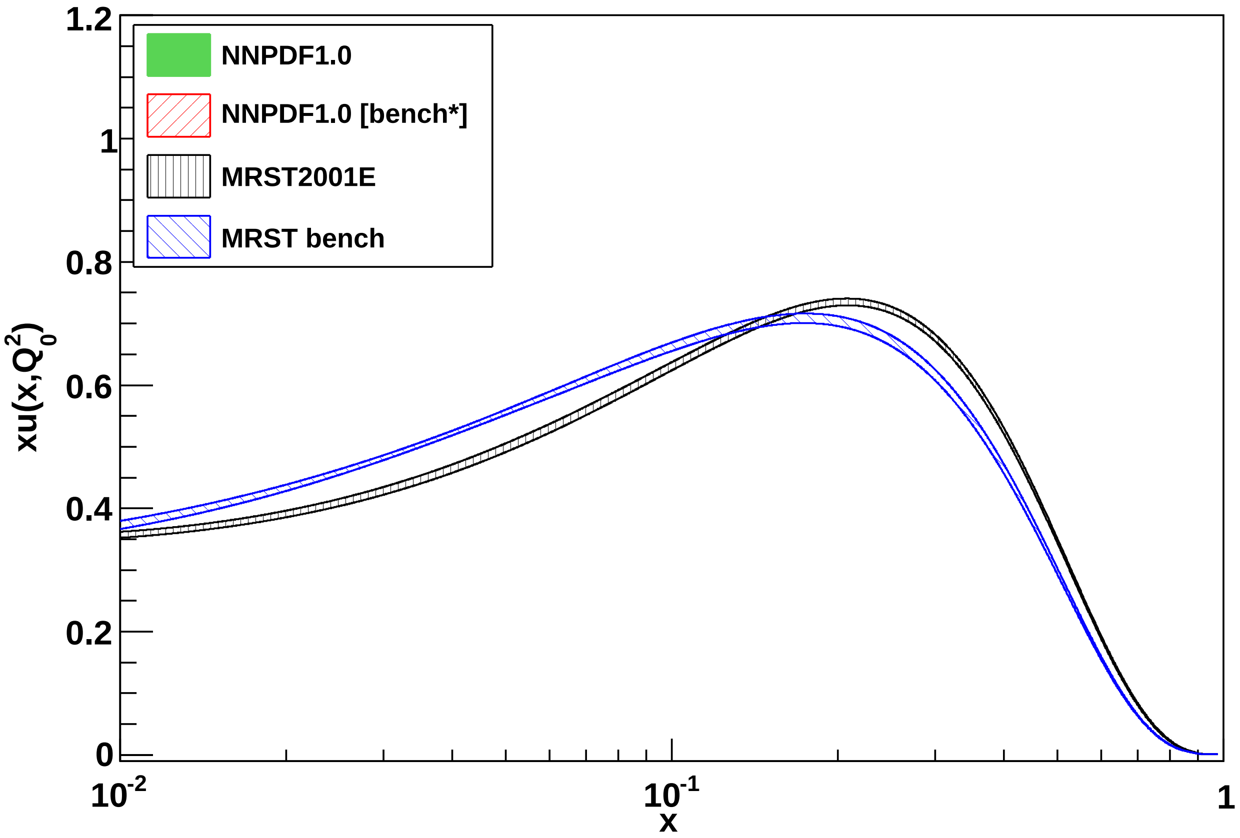


3163 data ⟶ **773** data

Comparing PDFs

HERA-LHC benchmark

Comparison between collaborations and between benchmark/global partons.
$u(x, Q^2 = 2\mathrm{GeV}^2)$: MRST data region

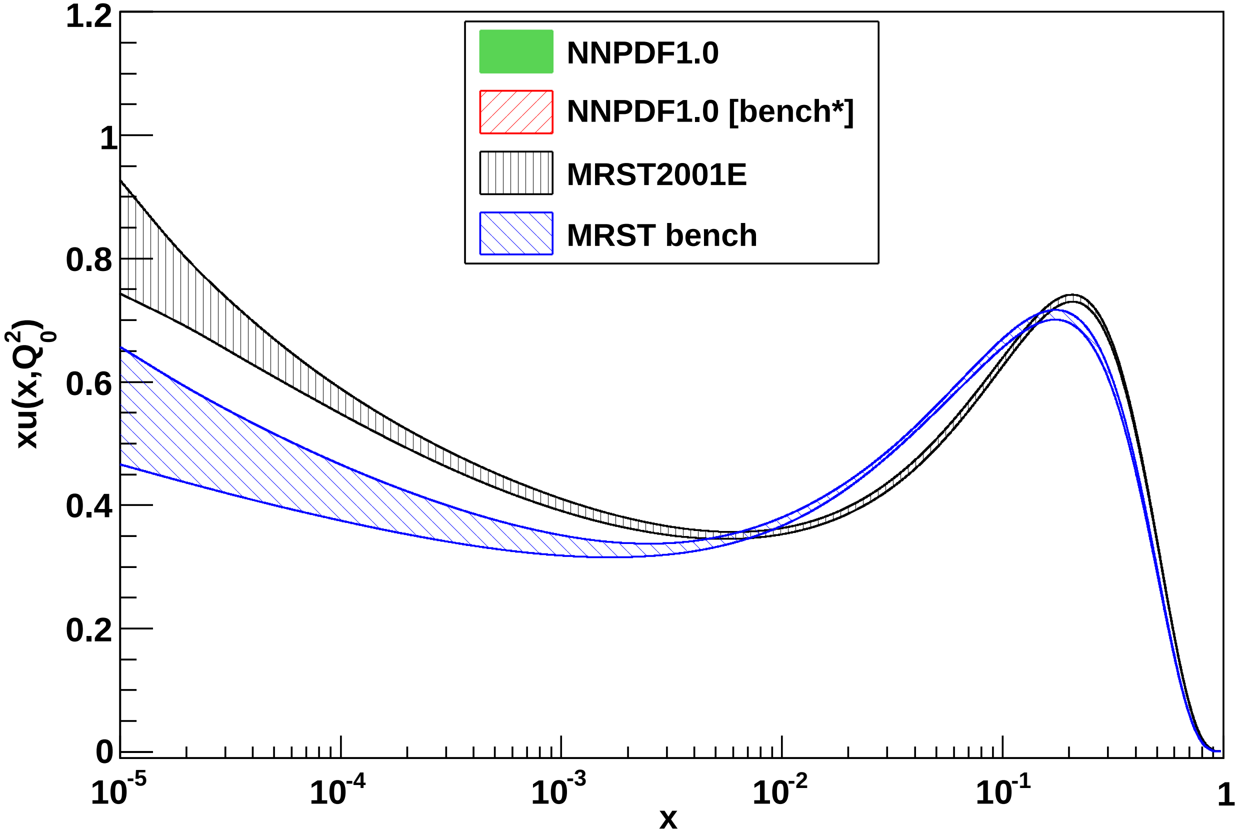


Comparing PDFs

HERA-LHC benchmark

Comparison between collaborations and between benchmark/global partons.

$u(x, Q^2 = 2\text{GeV}^2)$: MRST extrapolation region



Comparing PDFs

HERA-LHC benchmark

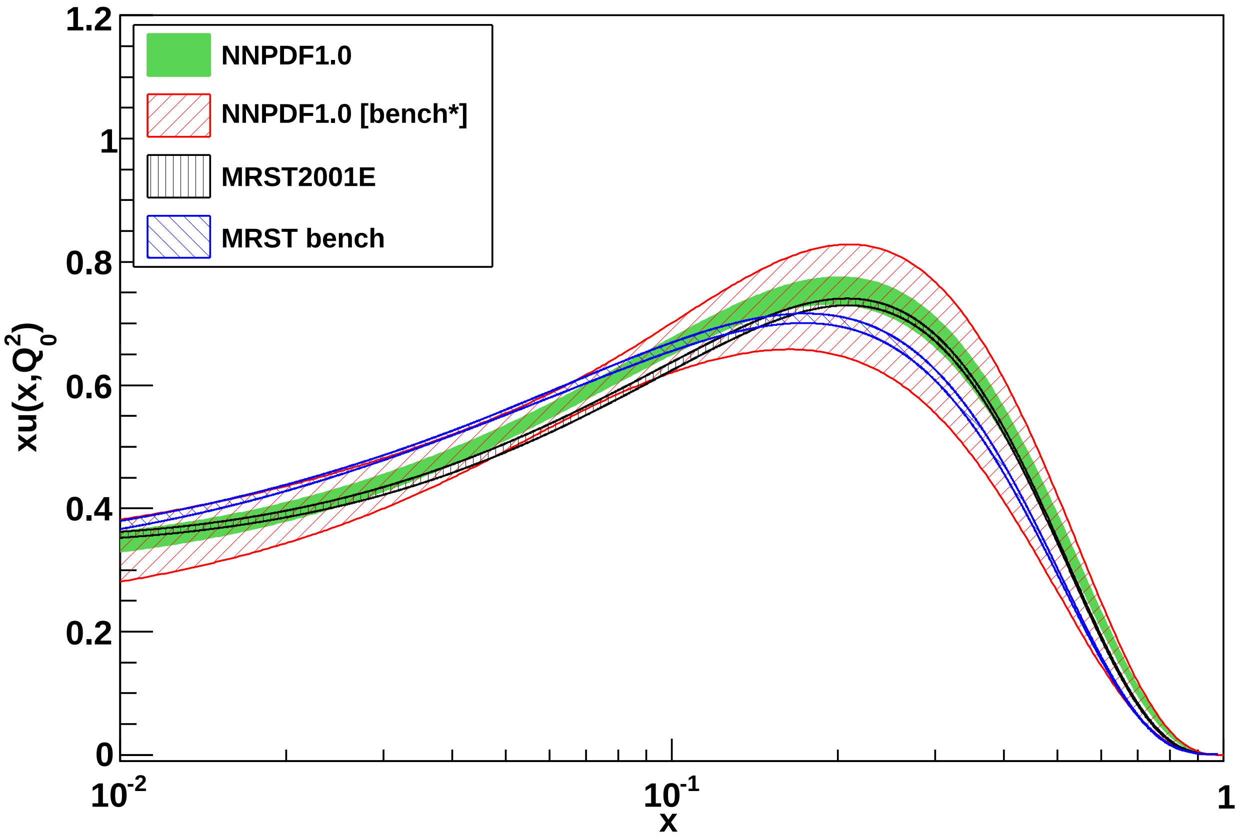
- benchmark partons and global partons do not agree within error!
- note that PDFs input parametrization, flavor assumptions and statistical treatment ($\Delta\chi^2_{\rm global} = 50$, $\Delta\chi^2_{\rm bench} = 1$) are tuned to data.
- not satisfactory especially to predict the behaviour of PDFs in the extrapolation region (LHC)

Comparing PDFs

HERA-LHC benchmark

Comparison between collaborations and between benchmark/global partons.

$u(x, Q^2 = 2\mathrm{GeV}^2)$: data region

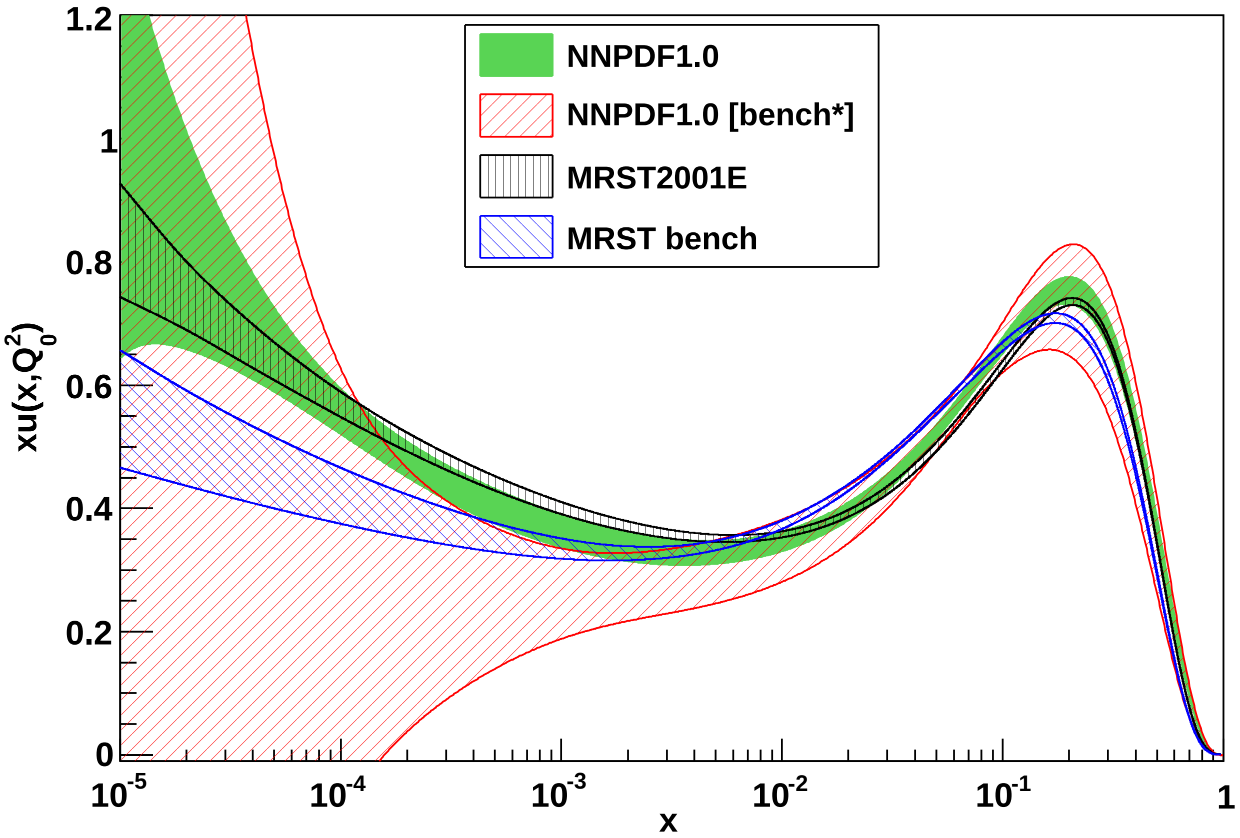


Comparing PDFs

HERA-LHC benchmark

Comparison between collaborations and between benchmark/global partons.

$u(x, Q^2 = 2\mathrm{GeV}^2)$: extrapolation region



Comparing PDFs

HERA-LHC benchmark

- NNPDF1.0 is consistent with MRST global fit.
- NNPDFbench is consistent with NNPDF1.0 and MRST.
- Same parametrization and flavour assumption.
- Same statistical treatment.
- Underestimation of the error in the standard approach.
- New MSTW fit has improved, but still tension btw global and benchmark fits

Conclusions

- Standard approaches with fixed parametrization tend to underestimate uncertainties unless experimental errors are inflated.
- Monte Carlo ensemble
  - Any statistical property of PDFs can be calculated using standard statistical methods.
  - No need of any tolerance criterion.
- The Neural Network parametrization
  - Small uncertainties come from an underlying physical law, not from parametrization bias.
  - Inconsistent data or underestimated uncertainties do not require a separate treatment and are automatically signalled by a larger value of the χ^2.
- The first NNPDF parton set [arXiv:0808.1231] is available on the common LHAPDF interface [http://projects.hepforge.org/lhapdf].

Conclusions

Unbiased PDFs with statistically meaningful error bars are important for LHC analyses

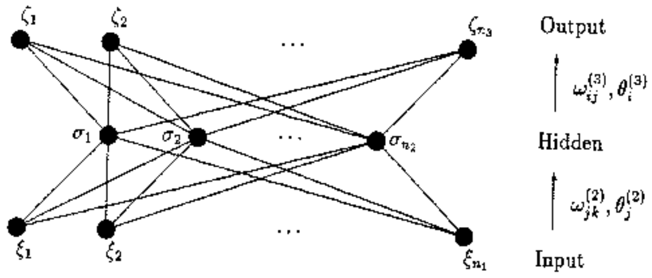
- **NNPDF1.0:** NLO DIS fit, available from LHAPDF
- **NNPDF1.1:** available from `http://sophia.ecm.ub.es/nnpdf`
- **NNPDF1.2:** dimuon data added, strange content fitted - to appear soon

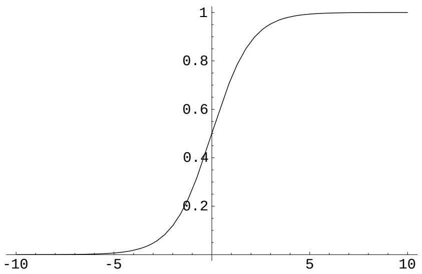
Forthcoming developments:

- inclusion of hadronic data
- heavy quarks effects

towards **NNPDF2.0:** global fit with faithful errors

Some details about Neural Networks





Multilayer feed-forward networks

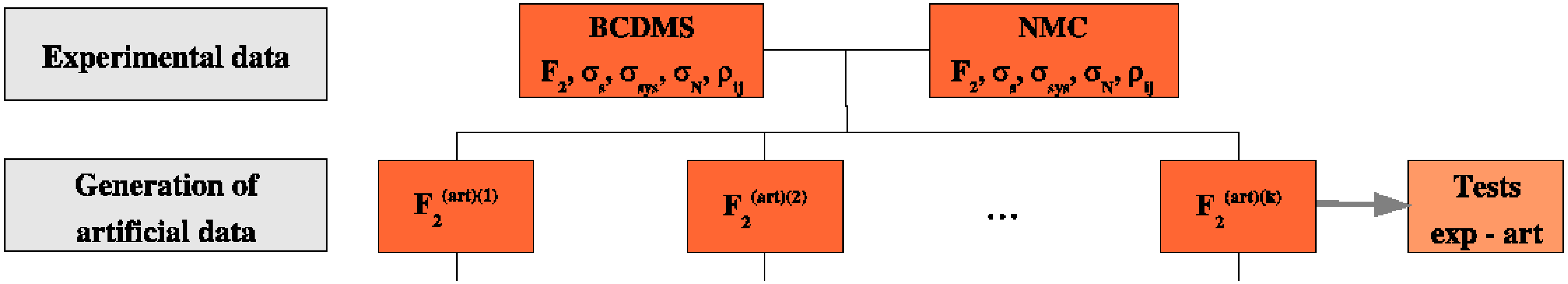
- Each neuron receives input from neurons in preceding layer and feeds output to neurons in subsequent layer
- Activation determined by weights and thresholds
  $\xi_i = g\left(\sum_j \omega_{ij}\xi_j - \theta_i\right)$
- Sigmoid activation function
  $g(x) = \frac{1}{1+e^{-\beta x}}$

... just another set of basis functions!

eg, a 1-2-1 NN: $\xi_1^{(3)}(\xi_1^{(1)}) = \frac{1}{1+\exp[\theta_1^{(3)} - \frac{\omega_{11}^{(2)}}{1+e^{\theta_1^{(2)} - \xi_1^{(1)}\omega_{11}^{(1)}}} - \frac{\omega_{12}^{(2)}}{1+e^{\theta_2^{(2)} - \xi_1^{(1)}\omega_{21}^{(1)}}}]}$

Thm: any function can be represented by a sufficiently big neural network

MC sampling of exp data



$$F_{I,p}^{(\mathrm{art})(k)} = S_{p,N}^{(k)} F_{I,p}^{(\mathrm{exp})} \left(1 + \sum_{l=1}^{N_c} r_{p,l}^{(k)} \sigma_{p,l} + r_p^{(k)} \sigma_{p,s} \right) \ , \ k = 1, \ldots, N_{\mathrm{rep}} \ ,$$

where

$$S_{p,N}^{(k)} = \prod_{n=1}^{N_a} \left(1 + r_{p,n}^{(k)} \sigma_{p,n} \right) \prod_{n=1}^{N_r} \sqrt{1 + r_{p,n}^{(k)} \sigma_{p,n}} .$$

Basis set

- Each independent PDF at the initial scale $Q_0^2 = 2\mathrm{GeV}^2$ is parameterized by an individual NN.
- Little constraint on strange $\longrightarrow$ 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)
  - Intrinsic heavy quarks contributions neglected.
- Parametrization of **(4+1)** combinations of PDFs at $Q_0^2 = 2$ GeV^2:

Singlet : $\Sigma(x)$	$\longmapsto \mathrm{NN}_\Sigma(x)$	2-5-3-1 37 pars
Gluon : $g(x)$	$\longmapsto \mathrm{NN}_g(x)$	2-5-3-1 37 pars
Total valence : $V(x) \equiv u_V(x) + d_V(x)$	$\longmapsto \mathrm{NN}_V(x)$	2-5-3-1 37 pars
Non-singlet triplet : $T_3(x)$	$\longmapsto \mathrm{NN}_{T3}(x)$	2-5-3-1 37 pars
Sea asymmetry : $\Delta_S(x) \equiv \bar{d}(x) - \bar{u}(x)$	$\longmapsto \mathrm{NN}_\Delta(x)$	2-5-3-1 37 pars

185 parameters

Flavour assumptions

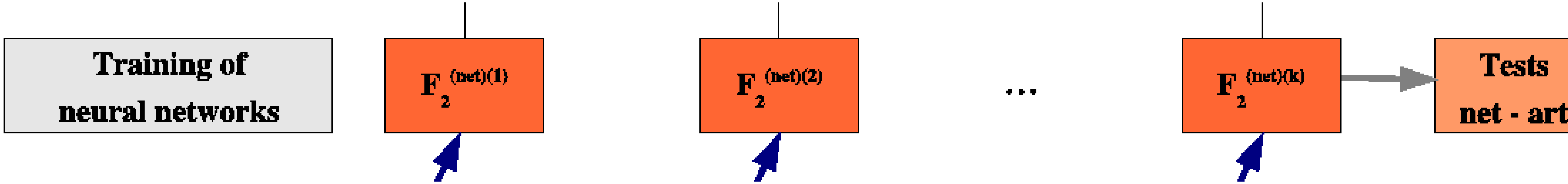
NNPDF1.0

$$
\begin{aligned}
\Sigma(x,Q_0^2) &= (1-x)^{m_\Sigma} x^{-n_\Sigma} \mathrm{NN}_\Sigma(x) \, , \\
V(x,Q_0^2) &= A_V (1-x)^{m_V} x^{-n_V} \mathrm{NN}_V(x) \, , \\
T_3(x,Q_0^2) &= (1-x)^{m_{T_3}} x^{-n_{T_3}} \mathrm{NN}_{T_3}(x) \, , \\
\Delta_S(x,Q_0^2) &= A_{\Delta_S} (1-x)^{m_{\Delta_S}} x^{-n_{\Delta_S}} \mathrm{NN}_{\Delta_S}(x) \, , \\
g(x,Q_0^2) &= A_g (1-x)^{m_g} x^{-n_g} \mathrm{NN}_g(x) \, , \\
s(x,Q_0^2) = \bar{s}(x,Q_0^2) &= C_s/2(\bar{u}(x,Q_0^2) + \bar{d}(x,Q_0^2))
\end{aligned}
$$

Normalization factors $\longrightarrow$ fixed by valence and momentum sum rules

$$
\begin{aligned}
\int_0^1 dx \, x \, (\Sigma(x) + g(x)) &= 1 \\
\int_0^1 dx \, (u(x) - \bar{u}(x)) &= 2 \\
\int_0^1 dx \, (d(x) - \bar{d}(x)) &= 1
\end{aligned}
$$

Training strategy

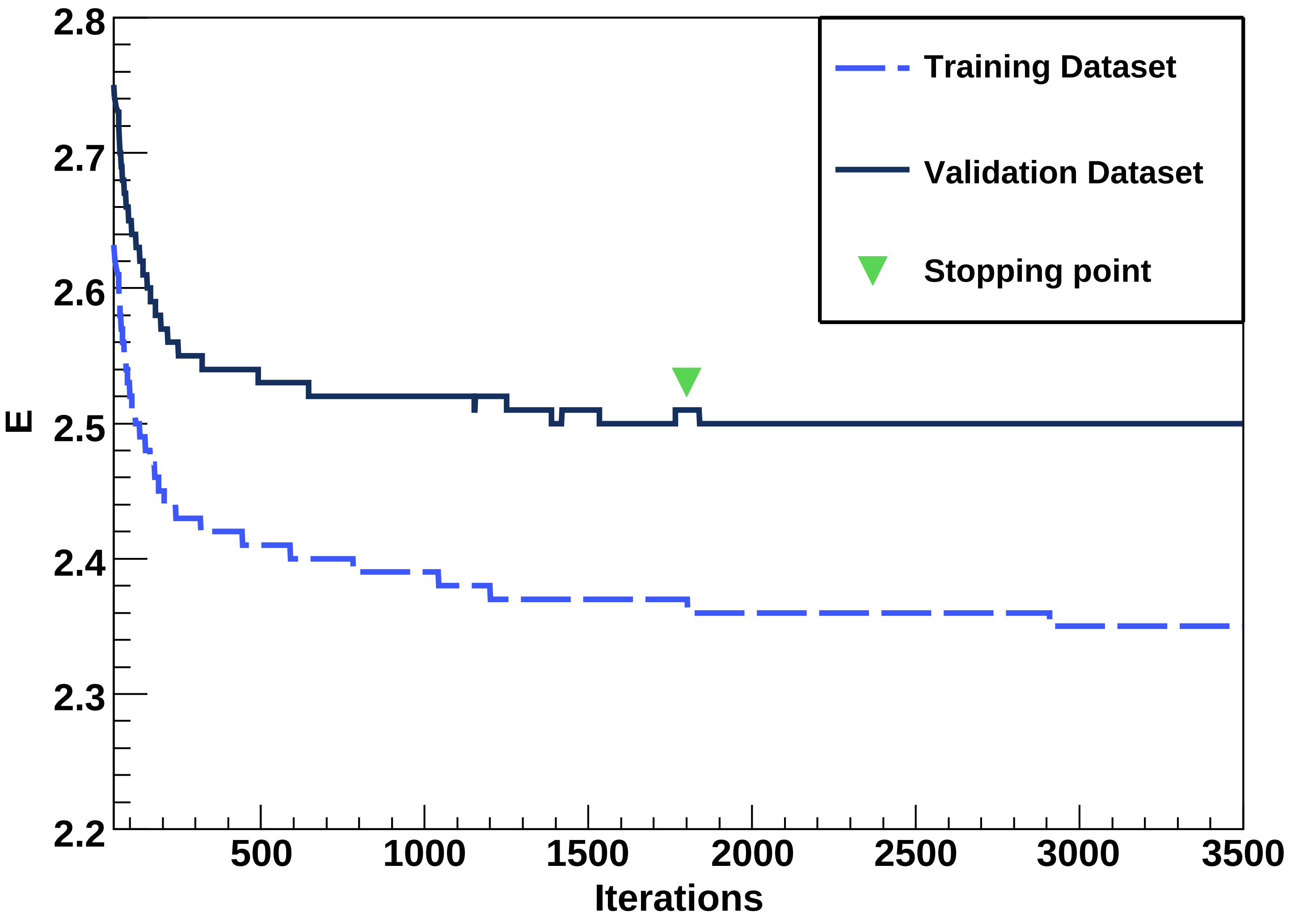


- unbiased basis of functions, parametrized by a **large** number of parameters
- genetic algorithms for minimization

- might accomodate statistical fluctutations of the data
- optimal training, beyond which the fit is just adjusting to statistical fluctutations

- dynamical stopping by cross validation
- for each replica divide the data randomly into **training** and **validation**
- minimization performed on the training set **only**
- when the training χ^2 still decreases while the validation χ^2 stops decreasing $\longrightarrow$ STOP

Dynamical stopping



Distance between MC ensembles.

- Stability of the NNPDF parton set can be assessed by using standard statistical tools.
- Distances between two probability distributions: $\left\{ f_{ik}^{(1)} = f_k^{(1)}(x_i, Q_0^2) \right\}$

$$\langle d[f] \rangle = \sqrt{ \left\langle \frac{ \left(\langle f_i \rangle_{(1)} - \langle f_i \rangle_{(2)} \right)^2 }{ \sigma^2[f_i^{(1)}] + \sigma^2[f_i^{(2)}] } \right\rangle_{\rm pts} }$$

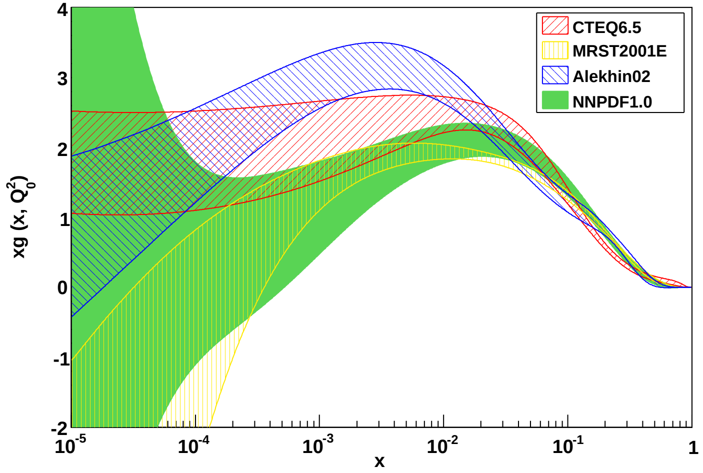
- where:

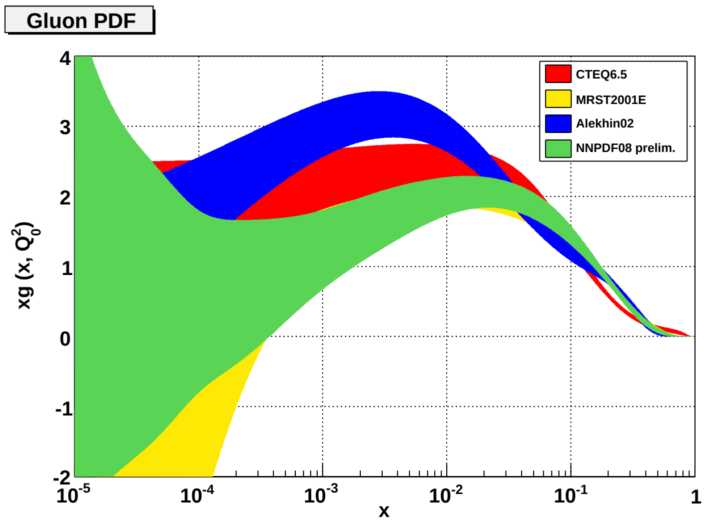
$$\langle f_i \rangle_{(1)} \equiv \frac{1}{N_{\rm rep}^{(1)}} \sum_{k=1}^{N_{\rm rep}^{(1)}} f_{ik}^{(1)} \, ,$$

$$\sigma^2[f_i^{(1)}] \equiv \frac{1}{N_{\rm rep}^{(1)}(N_{\rm rep}^{(1)} - 1)} \sum_{k=1}^{N_{\rm rep}^{(1)}} \left(f_{ik}^{(1)} - \langle f_i \rangle_{(1)} \right)^2$$

- For statistically equivalent PDF sets: $\langle d[f] \rangle \sim \langle d[\sigma_f] \rangle \sim 1$

Stability under variation of the parametrization





	Data	Extrapolation
$\Sigma(x, Q_0^2)$	$5\,10^{-4} \leq x \leq 0.1$	$10^{-5} \leq x \leq 10^{-4}$
$\langle d[f] \rangle$	0.98	1.25
$\langle d[\sigma] \rangle$	1.14	1.34
$g(x, Q_0^2)$	$5\,10^{-4} \leq x \leq 0.1$	$10^{-5} \leq x \leq 10^{-4}$
$\langle d[f] \rangle$	1.52	1.15
$\langle d[\sigma] \rangle$	1.16	1.07
$T_3(x, Q_0^2)$	$0.05 \leq x \leq 0.75$	$10^{-3} \leq x \leq 10^{-2}$
$\langle d[f] \rangle$	1.00	1.11
$\langle d[\sigma] \rangle$	1.76	2.27
$V(x, Q_0^2)$	$0.1 \leq x \leq 0.6$	$3\,10^{-3} \leq x \leq 3\,10^{-2}$
$\langle d[f] \rangle$	1.30	0.90
$\langle d[\sigma] \rangle$	1.10	0.98
$\Delta_S(x, Q_0^2)$	$0.1 \leq x \leq 0.6$	$3\,10^{-3} \leq x \leq 3\,10^{-2}$
$\langle d[f] \rangle$	1.04	1.91
$\langle d[\sigma] \rangle$	1.44	1.80

- Stability under change of architecture of the nets:
  37 pars $\longrightarrow$ **31 pars**
- Independence on the parametrization!