Update on NNPDF parton distributions

Nathan Hartland
The University of Edinburgh

The NNPDF Collaboration:
R. D. Ball, V. Bertone, S. Carrazza,
F. Cerutti, C. Deans, L. Del Debbio, S Forte,
A Guffanti, N.H, J.I. Latorre, J. Rojo and M. Ubiali.

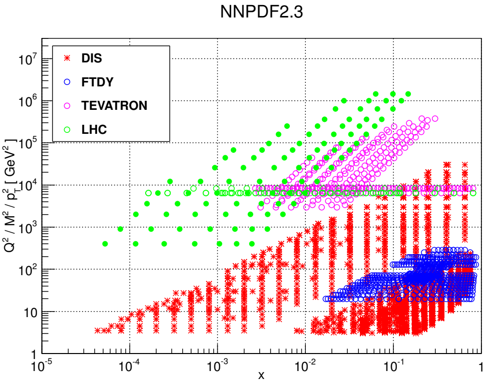
DIS 2013, Marseille
Tuesday 23nd April 2013

Current status of NNPDF determinations

- Most recent update: **NNPDF2.3** includes constraints from LHC data

New data in 2.3

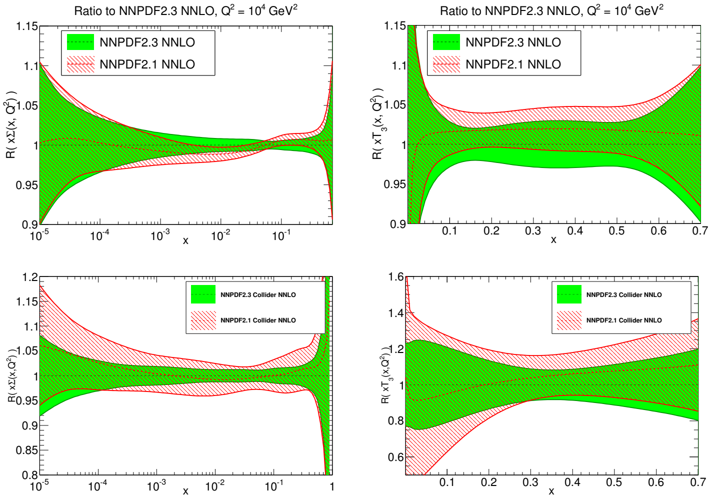
- **ATLAS 2010**
  Inclusive Jets,
  $W^{\pm}/Z$ rapidity distributions.
- **LHCb 2010**
  $W^{\pm}$ rapidity distributions.
- **CMS 2011**
  W lepton asymmetry.



NNPDF2.3 Family

NNPDF2.3 - global dataset including LHC data.
NNPDF2.3 noLHC - global dataset without LHC data.
NNPDF2.3 Collider - HERA, Tevatron and LHC data only.

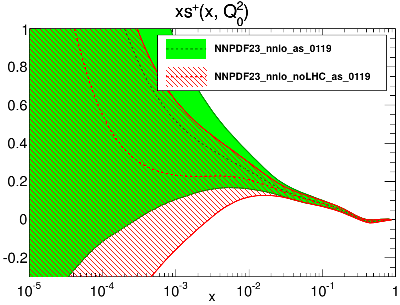
Constraints from LHC data



- LHC data generally demonstrates good consistency with the global dataset.
- Provides particularly large constraint for collider only PDFs.

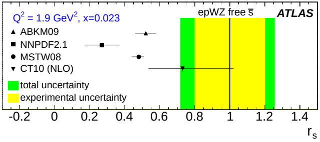
The strange content of the proton.

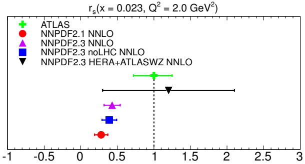
Strange distributions suffer from generally large uncertainties.

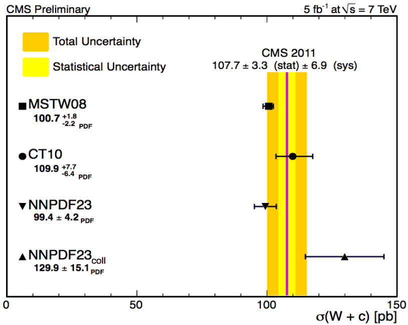


- LHC Electroweak measurements in NNPDF2.3 offers substantial constraint. Particularly in the collider only determination which previously suffered from a lack of data targeting strangeness.
- Data in 2.3 Collider only determination prefers larger values for total strangeness, although uncertainties remain large.

The strange content of the proton.







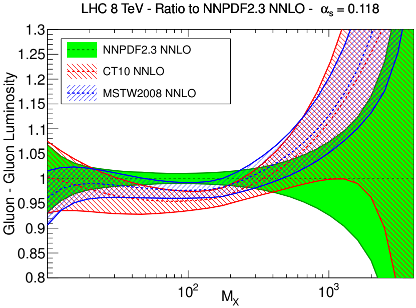
- NNPDF fit to HERA and ATLAS-WZ data finds central value consistent with ATLAS[1] determination of $r_s(x) = (s(x) + \bar{s}(x))/2d(x)$ within a large uncertainty.
- Recent CMS[2] measurement of $W + c$ consistent with strangeness in global fits. Slightly disfavours the larger strange sea in NNPDF2.3 Collider only, but consistent within uncertainties.

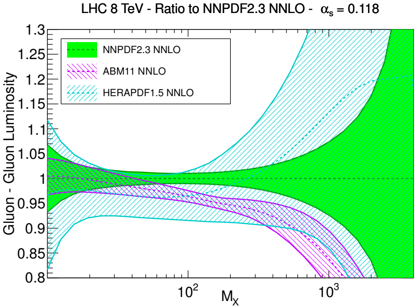
[1] arXiv:1203.4051

[2] CMS-SMP-12-002

PDF Benchmarking

[arXiv:1211.5142] - Benchmark study of different PDF determinations.
Detailed comparison at common α_S of the most up to date NNLO fits from the ABM, CT, HERAPDF, MSTW and NNPDF collaborations.

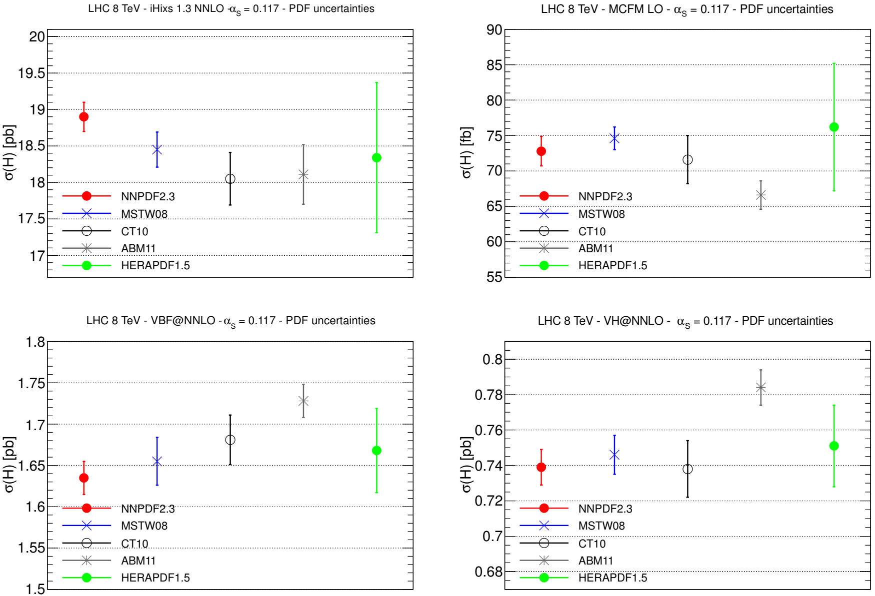




Reasonable agreement was found between CT, MSTW, NNPDF.

ABM softer large-x gluon and harder quarks.

Central values of HERAPDF1.5 NNLO agree with global fits, larger uncertainties due to reduced dataset.

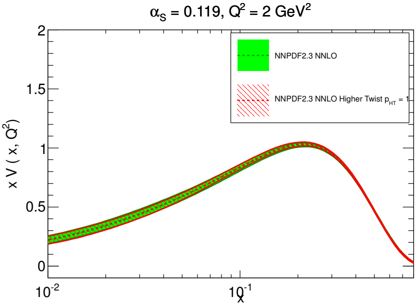


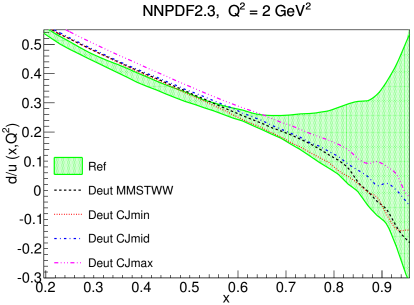
LHC 8 TeV - iHixs 1.3 NNLO -αS = 0.117 - PDF uncertainties
σ(H) [pb]
20
19.5
19
18.5
18
17.5
17
LHC 8 TeV - MCFM LO - αS = 0.117 - PDF uncertainties
σ(H) [fb]
90
85
80
75
70
65
60
55
LHC 8 TeV - VBF@NNLO -αS = 0.117 - PDF uncertainties
σ(H) [pb]
1.8
1.75
1.7
1.65
1.6
1.55
1.5
LHC 8 TeV - VH@NNLO - αS = 0.117 - PDF uncertainties
σ(H) [pb]
0.8
0.78
0.76
0.74
0.72
0.7
0.68
NNPDF2.3
MSTW08
CT10
ABM11
HERAPDF1.5

Theoretical uncertainties in PDF determination.

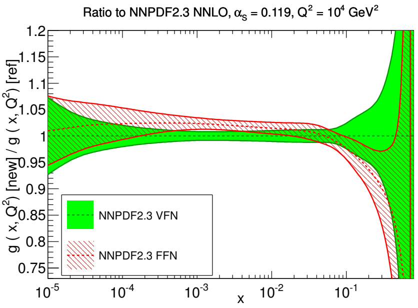
[arXiv:1303.1189] - NNPDF study of contributions to theoretical uncertainty.

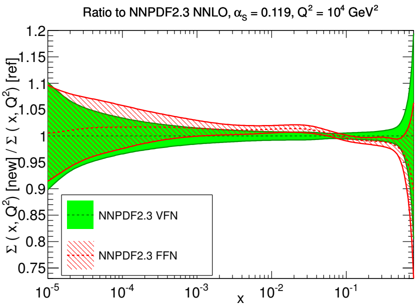
- **Dynamical Higher Twist**
  NNPDF Fit with higher twist corrections (from ABM determination) indicates modest impact upon PDFs
- **Deuterium nuclear corrections**
  Potentially affects down quark, fitted from deuterium data. Impact is limited to the d/u ratio in the range $0.1 \leq x \leq 0.5$





Impact of Flavour Number Scheme choice.

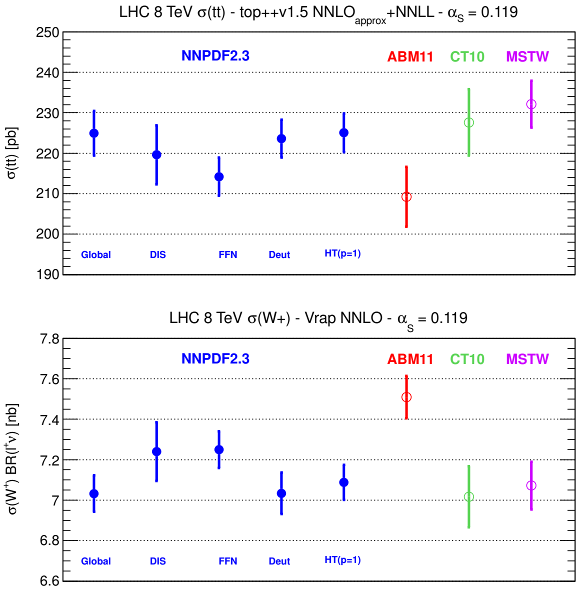




- **GM-VFNS vs FFNS**
  NNPDF fit with a fixed flavour number scheme treatment for DIS observables.
  Substantial impact observed in PDF central values.

  The FFN quarks increase at medium to low-x
  The FFN gluon becomes softer at high-x.

Impact of FFN Scheme



Towards NNPDF3.0

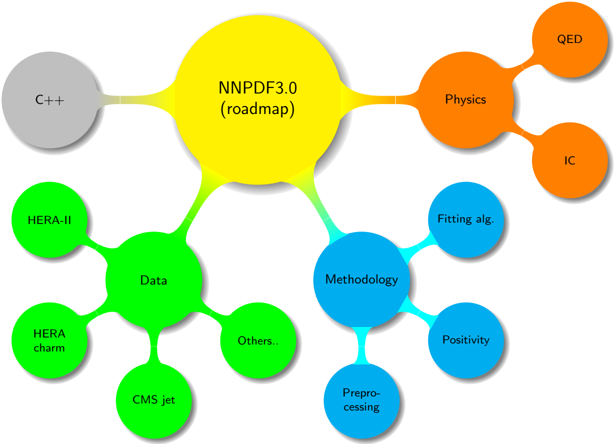
Challenges

- **Predictions for LHC data are computationally rather expensive.**
  - A fast fitting framework is required.
- **New datasets from the LHC and HERA are becoming available.**
  - The ability to rapidly implement and assess the impact of new data is vital.
- **Fitting methodology should be analysed in the light of new data.**
  - Is our current methodology doing the best job it can?

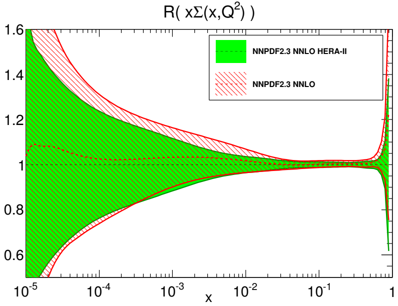
A new NNPDF code is under development, built from scratch to support rapid iteration.

- Faster fits means more aggressive minimisation, more scope for Methodological surveys.
- Code is written in modern C++ to aid flexibility and maintainability.
- Rewrite provides an in-depth cross check of the implementation.

Towards NNPDF3.0



New data - HERA-II DIS

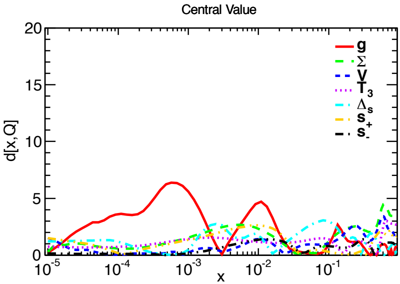


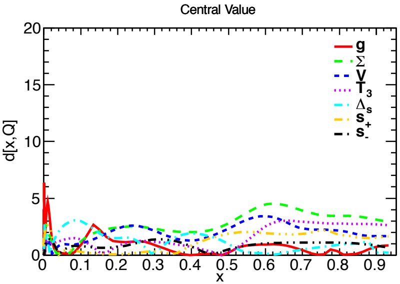
Preliminary Fits to HERA-II Inclusive DIS Datasets.

Large new DIS datasets from the ZEUS and H1 collaborations.

Over **600** new data points demonstrate excellent consistency in global fit, providing constraint particularly for the singlet PDF.

PDF distances demonstrate this consistency. A distance of $d = 10$ implies a shift by 1-σ.

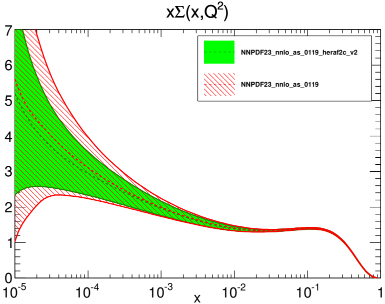


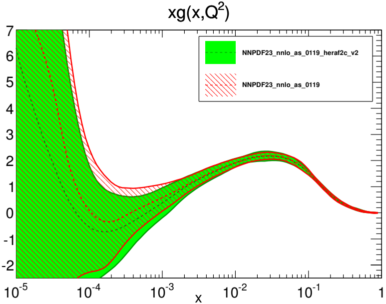


New data - HERA Combined F_c^2

[arXiv:1211.1182] - ZEUS and H1 Combination F_c^2 data.

Supersedes separate datasets used in previous NNPDF fits, providing additional information on data correlations.





Introducing the extra correlations has a generally modest impact, restricted to the singlet and gluon distributions.

Methodology: Closure tests

How do we ensure that our fit minimises *bias*?
Related studies by Thorne-Watt [arXiv:1205.4024]

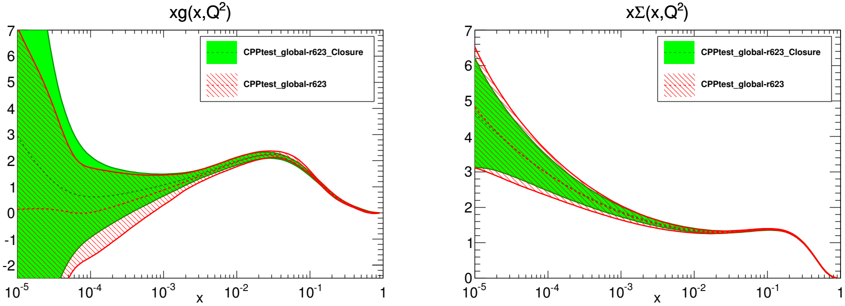
Perform a **Closure Test**:

- **Generate artificial pseudo-data based upon a known PDF distribution.**
  Pseudodata generated according to NLO pQCD. Dataset is therefore free of internal inconsistencies.
- **Simulate experimental noise in the pseudodata.**
  Data points perturbed according to multi-gaussian distribution defined by the experimental covariance matrix.
- **Perform a full PDF fit to the pseudo-dataset.**
  Closure fit should recover generating PDF up to the level of experimental uncertainty.

Methodology: Closure tests

Preliminary NNPDF closure test fits.

First closure tests performed with NNPDF C++ code to a NLO fit with a global dataset.



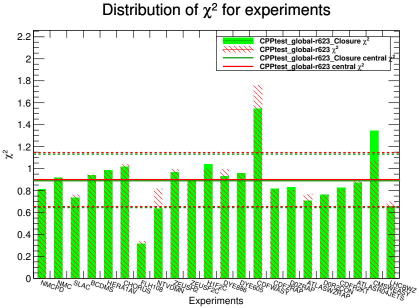
Early closure test fits demonstrate good reproduction of the generating function, accurate up to experimental error.

Reproduction of uncertainty is particularly interesting. Lack of data inconsistency does not appear to lead to significant change in the resulting uncertainties.

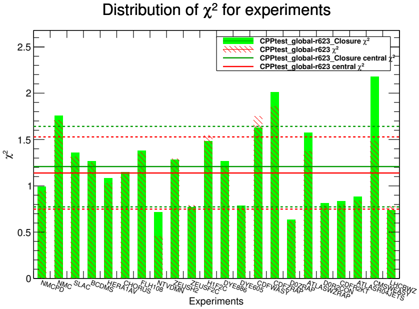
The NNPDF methodology can be consistently applied to experimental data and theoretically perfect pseudodata without the need for tolerance.

Preliminary closure tests - fit quality

Pseudodata



Experimental data



- The χ^2 value for each dataset is assessed both for the artificial pseudodata, and for the real experimental data.
- Closure test fit demonstrates self consistency of the NNPDF procedure.

Summary and Outlook

Current Status

- **NNPDF2.3**
  The NNPDF2.3 family of fits provide a determination of parton distribution functions with a global dataset, including a sizeable LHC contribution.

  Measurements from the LHC and HERA will provide interesting further constraints upon PDFs, particularly in collider only determinations.

Looking Forward

A great deal of progress in NNPDF determinations across many fronts.

- **Progress towards the next global NNPDF set**
  - New fitting framework developed from scratch in C++.
  - Preliminary methodological studies by Closure Testing.
  - Impact of new HERA combinations upon NNPDF2.3 studied.
  - Plenty of new data to come (e.g CMS Inclusive Jets, $W + c$).
- **New Results for PDFs with QED Corrections, and polarised NNPDFs**
  see talks by S.Carrazza, E.Nocera.

BACKUPS

Including new experimental data - reweighting

How can we add new LHC data to an existing parton set?

- Reweight existing Monte Carlo parton set.

Each replica in the set is assigned a weight based upon it's χ^2 to the new data.

$$\langle \mathcal{O} \rangle_{\text{new}} = \frac{1}{N} \sum_{k=1}^{N} w_k \mathcal{O}[f_k], \qquad w_k \propto (\chi^2_k)^{(n-1)/2} e^{-\frac{1}{2}\chi^2_k}$$

- Application: NNPDF2.2 Parton Set [arXiv:1012.0836]
  LHC Electroweak data added by Bayesian Reweighting

However, reweighting method is impractical for large/constraining data sets. Number of effective replicas reduced after reweighting:

$$N_{\text{eff}} \equiv \exp\left(\frac{1}{N_{\text{rep}}} \sum_{k=1}^{N_{\text{rep}}} w_k \ln(N_{\text{rep}}/w_k) \right)$$

Including new experimental data - refitting

How can we efficiently include LHC data into a full refit?
Tools: APPLgrid/FastNLO projects

- Precompute and store MC Weights on an interpolation grid

$$\sigma = \sum_p \sum_{l=0}^{N_{\mathrm{sub}}} \sum_{\alpha,\beta}^{N_x} \sum_{\tau}^{N_Q} W_{\alpha\beta\tau}^{(p)(l)} \left(\frac{\alpha_s\left(Q_\tau^2\right)}{2\pi} \right)^p F^{(l)}\left(x_\alpha, x_\beta, Q_\tau^2\right) \quad (1)$$

PDF Evolution in the FastKernel method is a similar procedure,

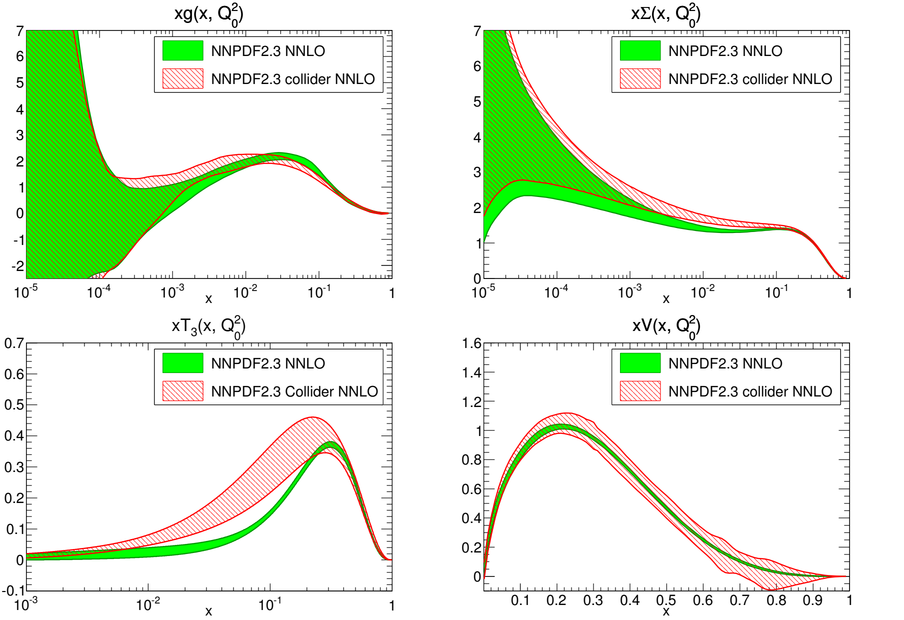
$$f_i(x_\alpha, Q_\tau^2) = \sum_\beta^{N_x} \sum_j^{N_{\mathrm{pdf}}} A_{\alpha\beta ij}^{\tau} N_j^0(x_\beta)$$

Idea: Combine weight grids with evolution grids

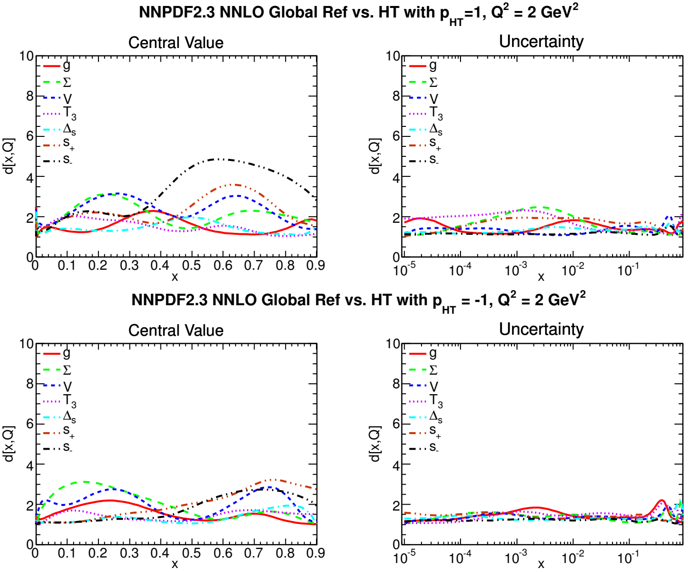
$$\sigma = \sum_{\alpha,\beta}^{N_x} \sum_{i,j}^{N_{\mathrm{pdf}}} \sigma_{\alpha\beta ij} N_i^0(x_\alpha) N_j^0(x_\beta)$$

- Precomputing all Q^2 dependence leads to extremely efficient calculations.

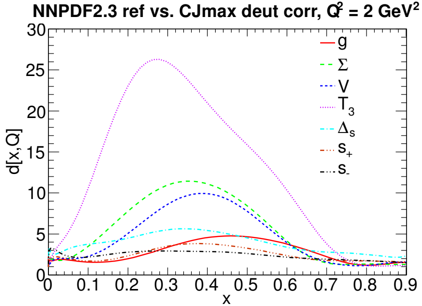
NNPDF2.3 Collider only vs NNPDF2.3

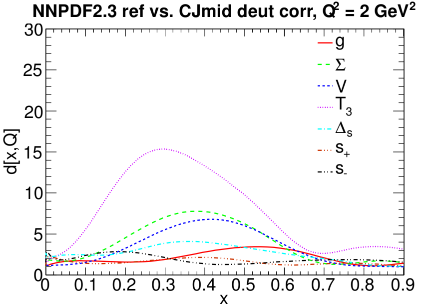


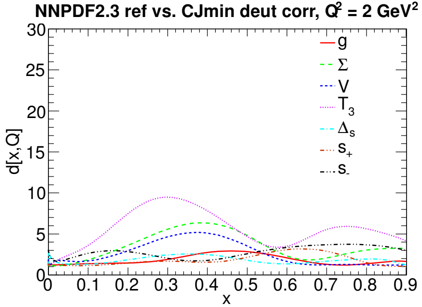
Impact of HT upon NNPDFs

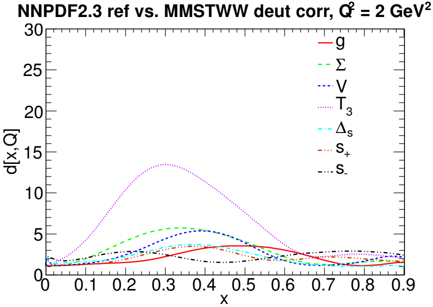


Impact of deuterium corrections upon NNPDFs









Impact of FFN Scheme