RWTH AACHEN UNIVERSITY

NNPDF collaboration

NNPDF for experimentalists

Maria Ubiali

(RWTH University Aachen)

in collaboration with:

F. Cerutti, J.I. Latorre (Barcelona U)

R.D. Ball, L. Del Debbio (Edinburgh U)

A. Guffanti, V. Bertone (Freiburg U)

S. Forte, J. Rojo (Milan U)

7th July 2011

DESY, Hamburg --- Analysis center seminar

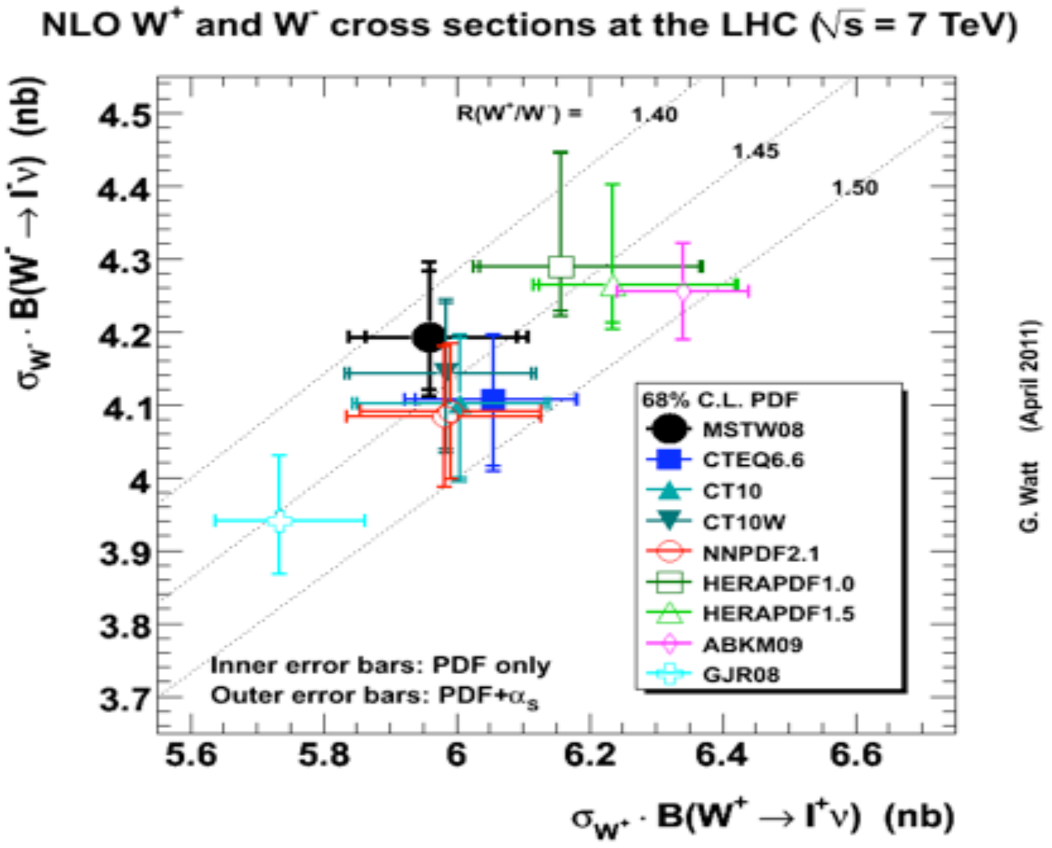
Outline

- PDFs for LHC
    - How we would like them to be
    - The NNPDF partons
    - Instructions for use and recent progress
    - Comparison with other PDF sets
- LHC for PDFs
    - Constraints from LHC
    - A collider-only analysis?
    - Reweighting techniques
    - Assessing impact of the LHC data without refitting
    - The W lepton asymmetry case
    - NNPDF2.2 set including LHC data
- Conclusion and outlook

PDFs for LHC

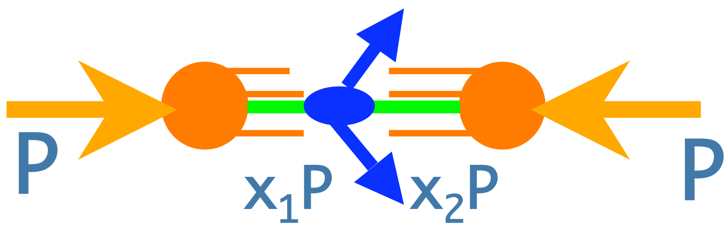
- How we would like them to be
- The NNPDF partons
- Instructions for use and recent progress
- Comparison with other PDF sets

LHC collisions
and PDF uncertainties



G. Watt, ArXiv: 1106.5788

For many standard candle processes at the LHC theoretical uncertainty is dominated by PDF uncertainties



A reliable understanding of PDFs and of their associated uncertainties plays a crucial role in relating theory to experimental observation at the LHC.

PDFs for LHC
Wish-list

❶ Precise: inclusion of as many info as possible and genuine statistical C.L.

- Correlation in experimental systematic
- Uncertainty treatment (normalization....)

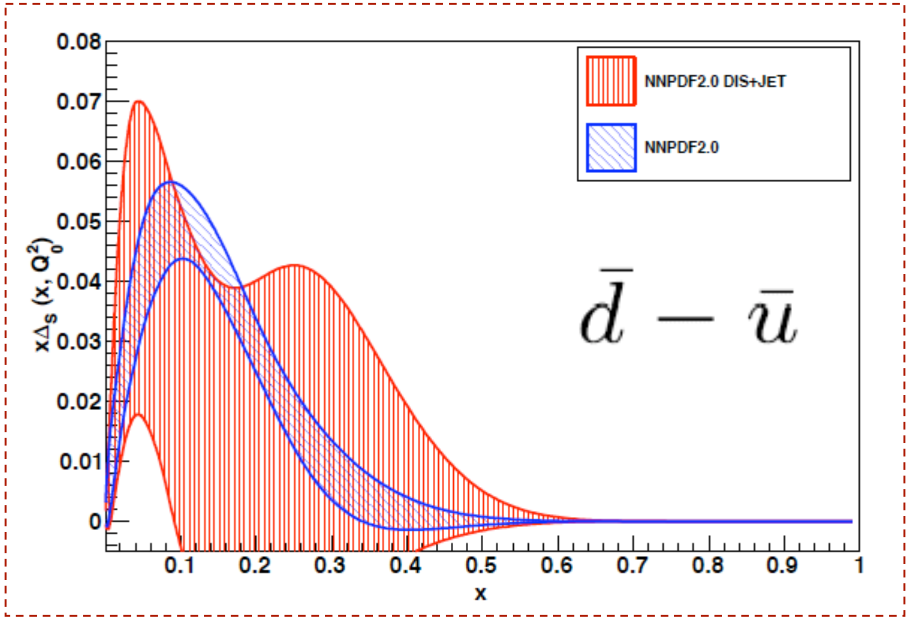
❷ No theoretical bias beyond the prior theory NLO(NNLO) QCD

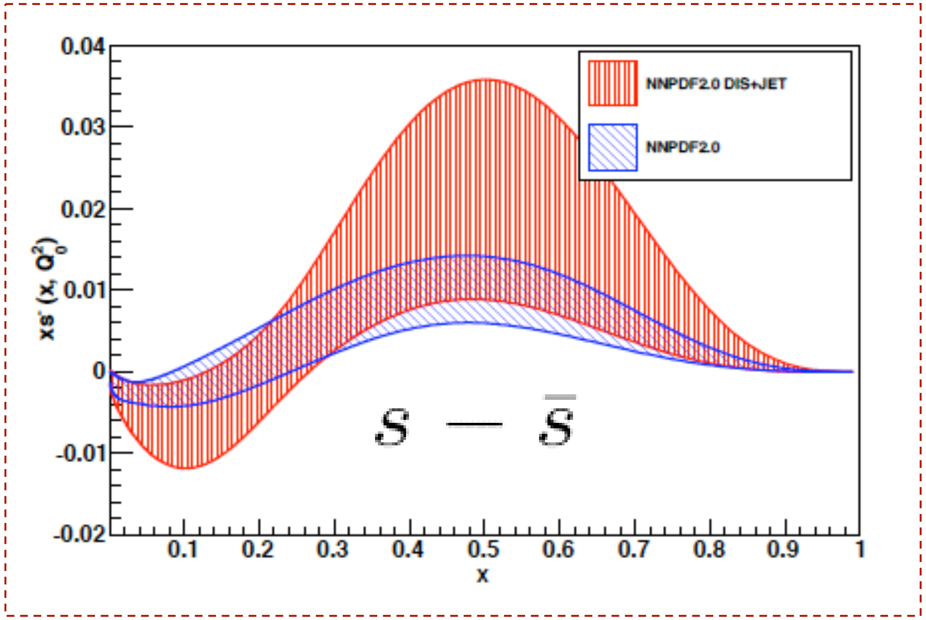
- Flexible parametrization
- Least flavor assumptions
- Statistically sound procedure for error determination

PDFs for LHC
Wish-list

❶ Precise: inclusion of as many info as possible and genuine statistical C.L.

DIS data alone do not constrain all PDF combination / kinematical regions



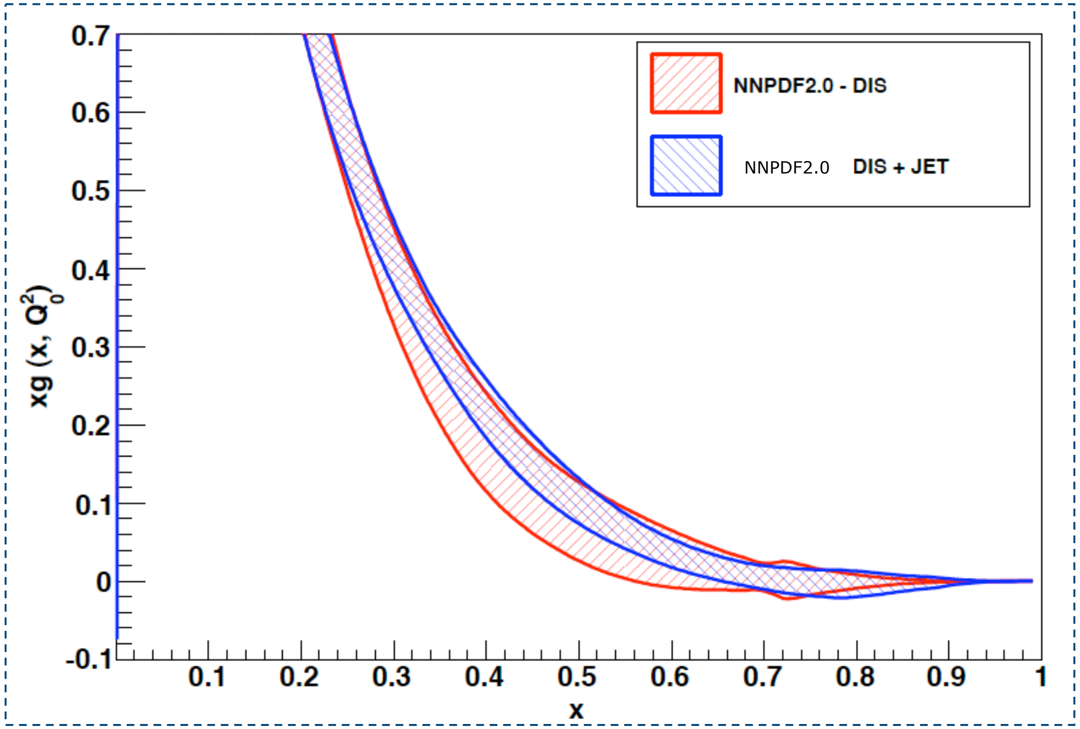


Drell Yan data ➔ Huge improvement in sea asymmetry and flavor separation

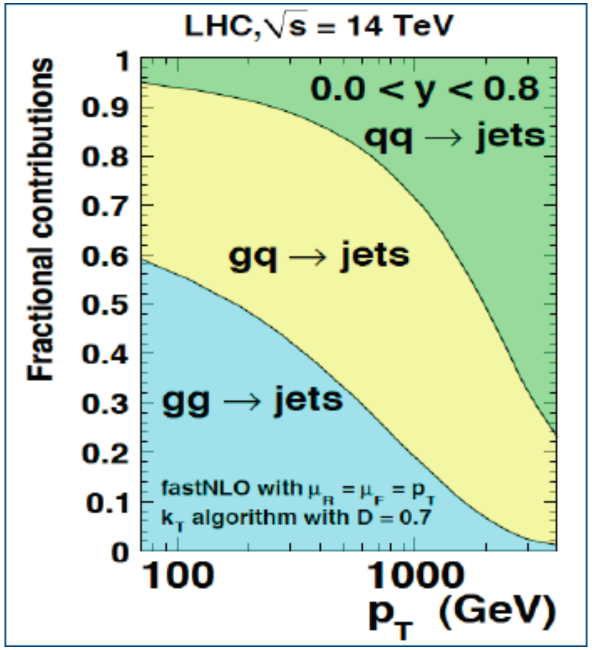
PDFs for LHC
Wish-list

❶ Precise: inclusion of as many info as possible and genuine statistical C.L.

DIS data alone do not constrain all PDF combination / kinematical regions



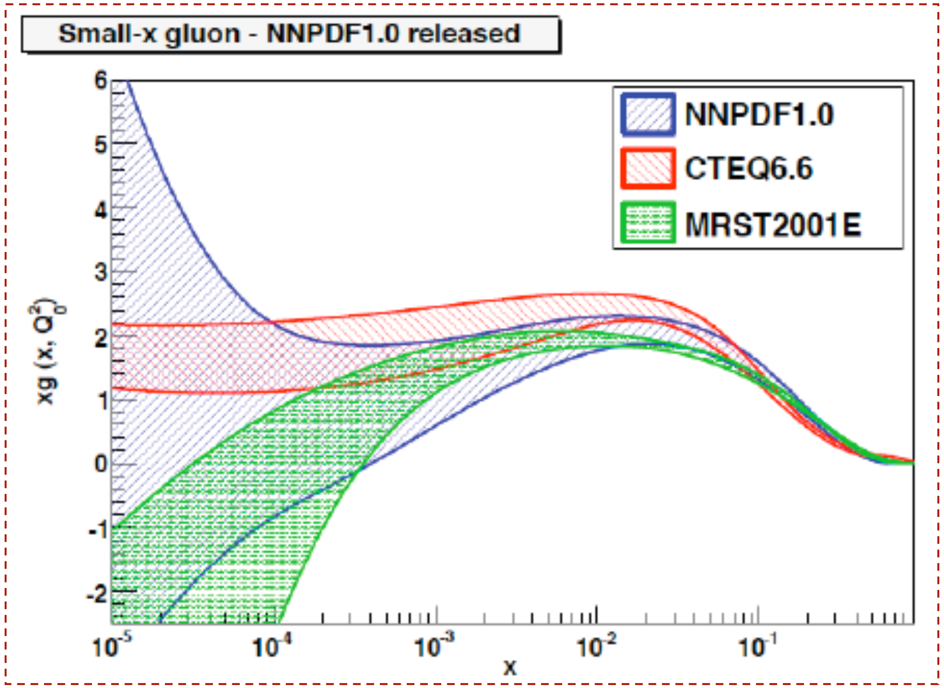
MSTW, ArXiv 0901.0002

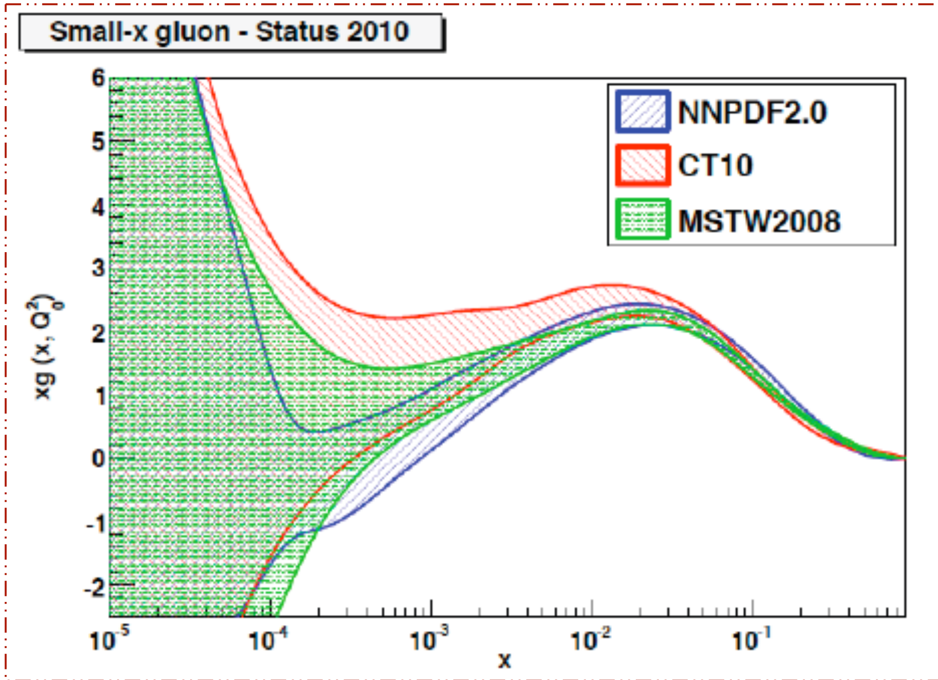


Jet data ➔ Constrain large x gluon

PDFs for LHC
Wish-list

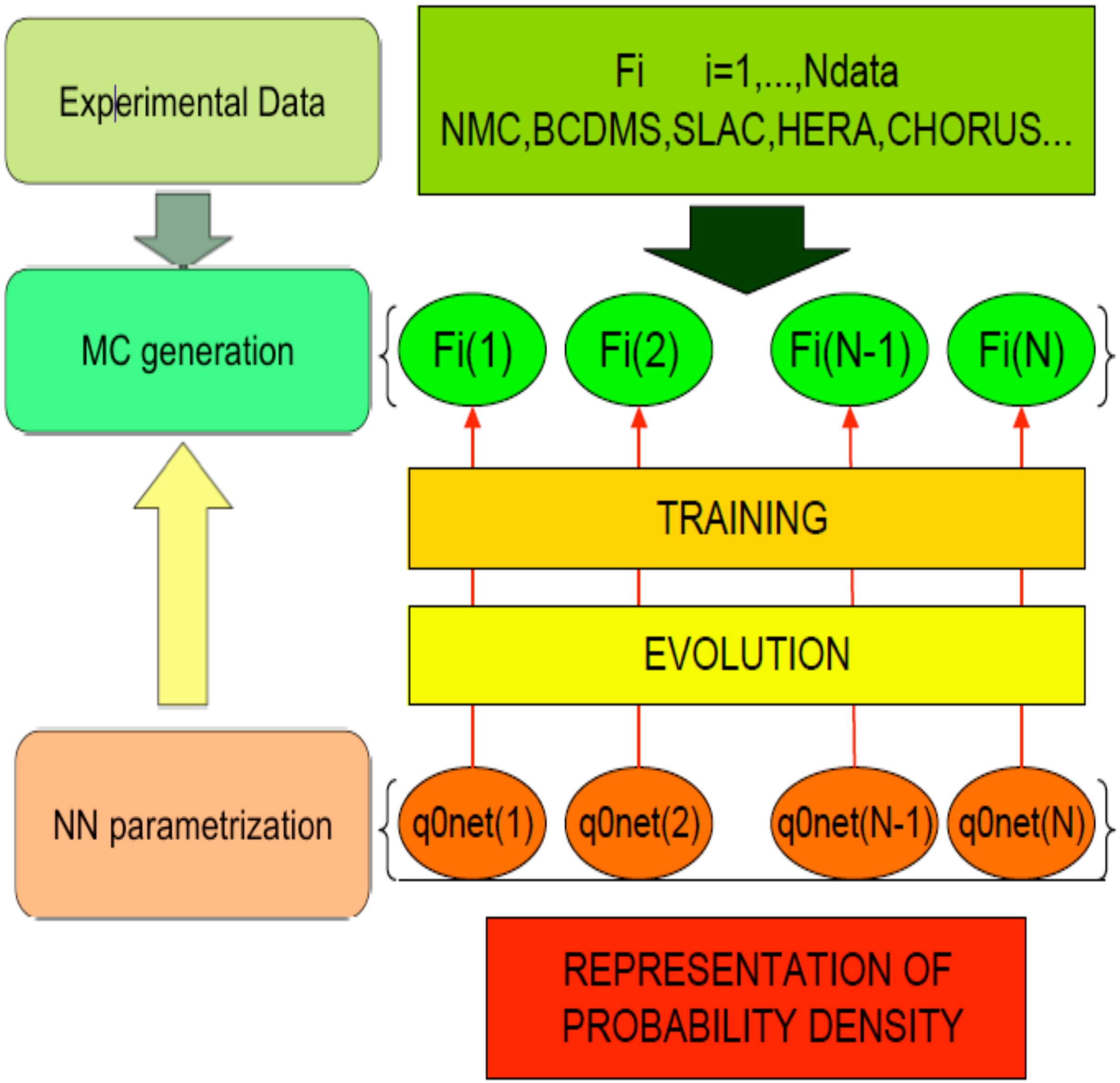
❷ No theoretical bias beyond the prior theory NLO(NNLO) QCD

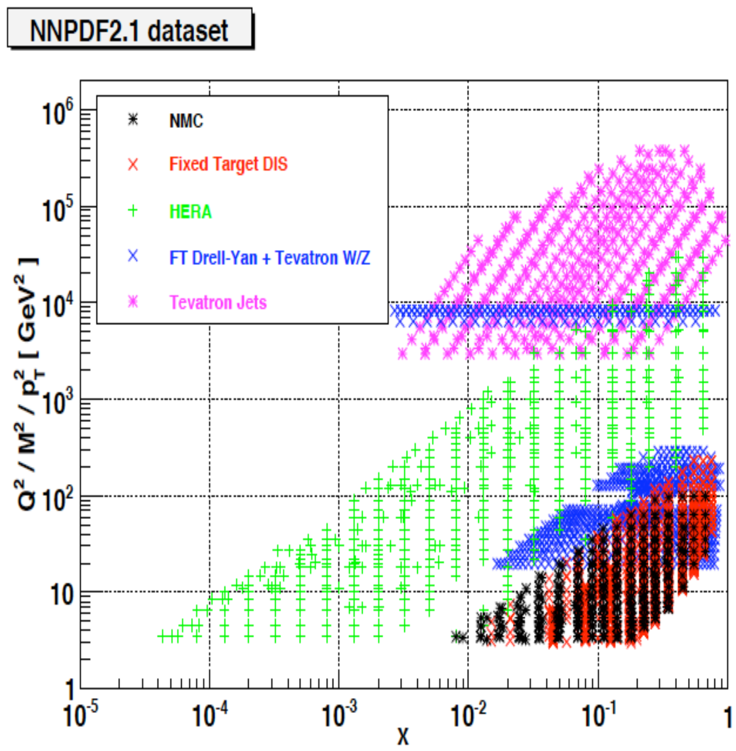


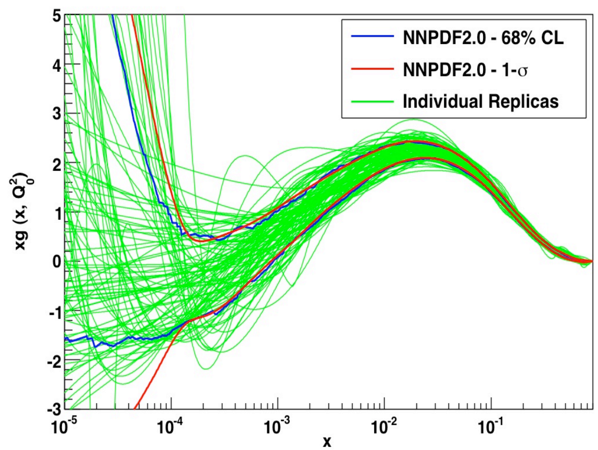


FIXED FUNCTIONAL FORM:
As new data are added, parametrization is increased to accommodate fit to new data or avoid unrealistic small uncertainties where data do not constrain PDFs.
What's the error associated to a given parametrization? (HERAPDF)

PDFs for the LHC
The NNPDF proposal





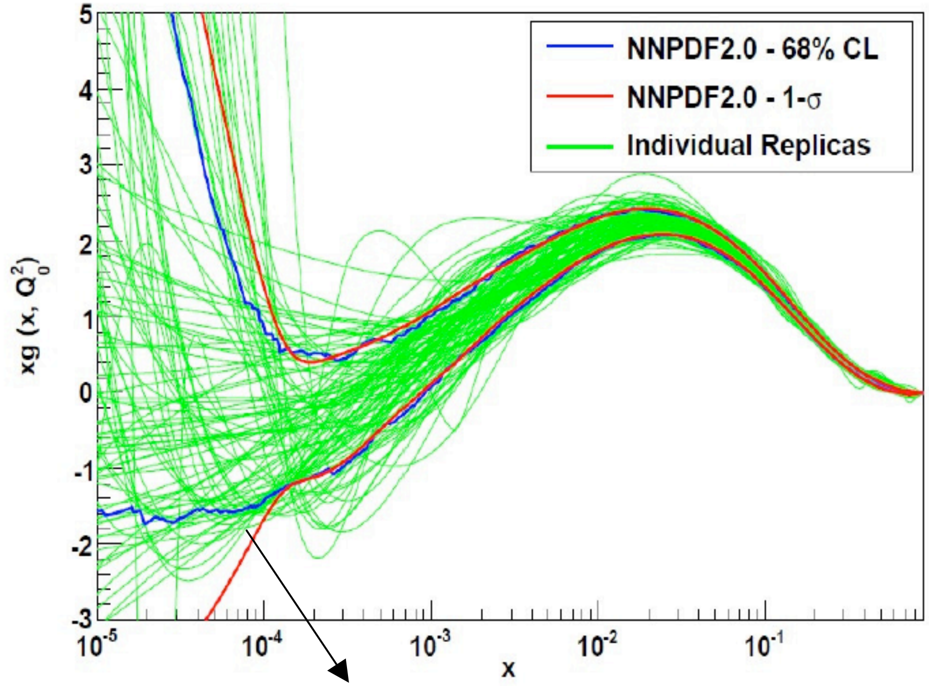


PDFs for the LHC
The NNPDF proposal

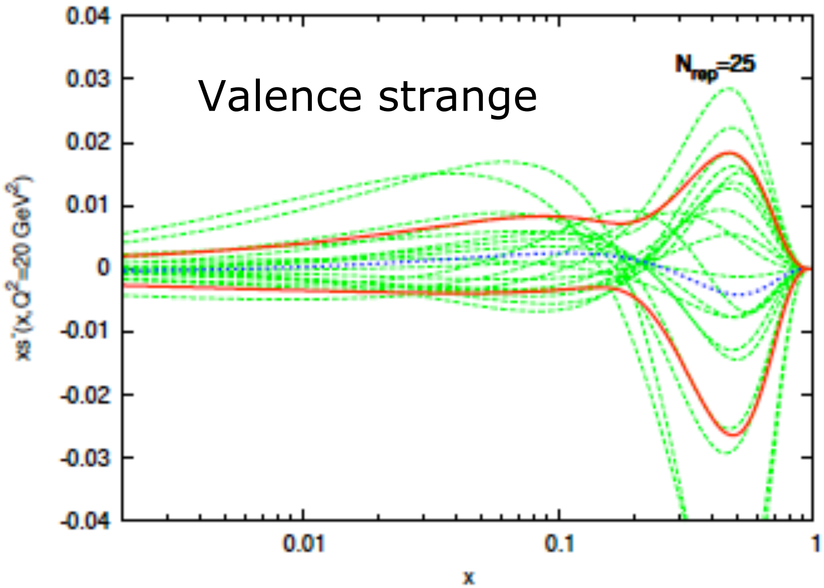
- Monte Carlo determination of errors
  - ✓ No need to rely on linear propagation of error
  - ✓ Keep into account all correlations
  - ✓ Possibility to test for non Gaussian behavior in fitted PDFs (1σ versus 68% C.L.)
- PDFs parametrized by Neural Networks
  - ✓ Provide an unbiased parametrization (NN: ~250 params versus ~ 30 params)
  - ✓ Can check independence of parametrization
- Stopping based on Cross-Validation
  - ✓ Ensures proper fitting avoiding over-learning

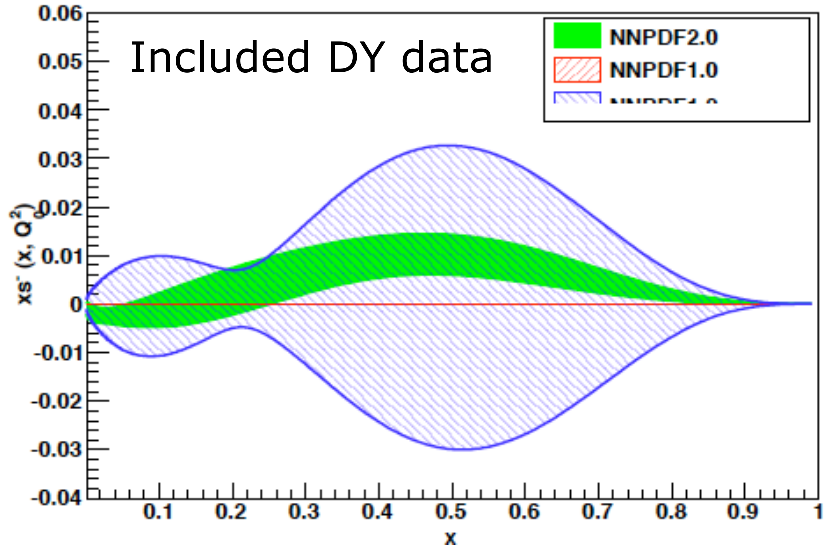
PDFs for the LHC
The NNPDF proposal

- ✓ Expected behavior upon addition of new data
- ✓ Use the same parametrization for all fits (> 250 param): can check own much subsets of data constrain PDFs
- ✓ Can determine 68% CL as well as 1σ



non Gaussian behavior in extrapolation region when positivity constraints on xsec are imposed





PDFs for the LHC

A brief historical overview

	NNPDF1.0 (2008)	NNPDF1.2 (2009)	NNPDF2.0 (2010)	NNPDF2.1 (2011)	NNPDF2.1 nnlo and lo	NNPDF2.2
DIS (FT and HERA)	✔	✔*	✔	✔	✔	✔
HERA combined	✘	✘	✔	✔	✔	✔
Drell-Yan (FT and Tevatron)	✘	✘	✔	✔	✔	✔
Jet data	✘	✘	✔	✔	✔	✔
LHC data	✘	✘	✘	✘	✘	✔
NNLO	✘	✘	✘	✘	✔	✔
Heavy Quark mass	✘	✘	✘	✔	✔	✔

* added neutrino and freed strange

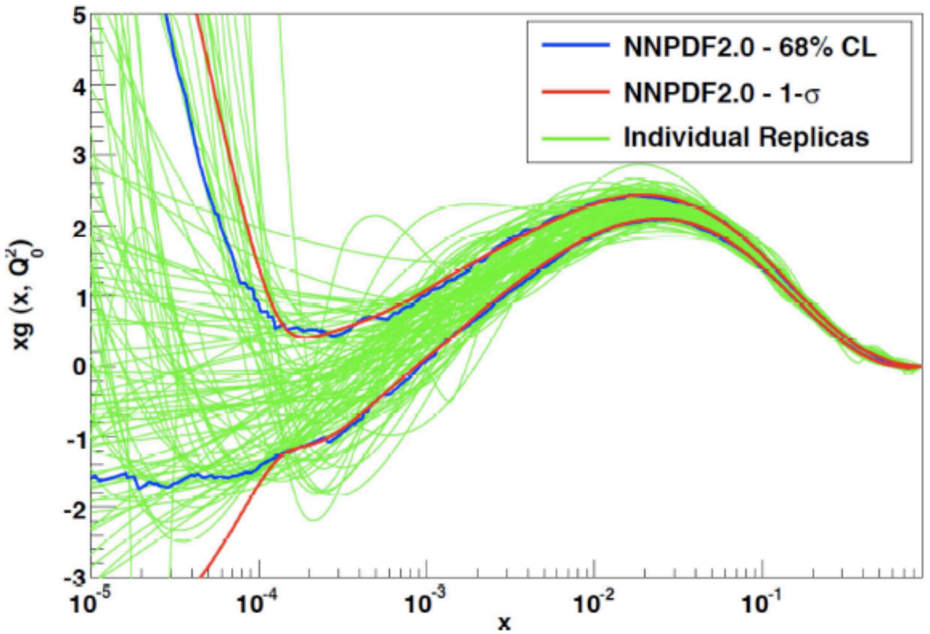
NNPDF for the LHC
Instruction for use

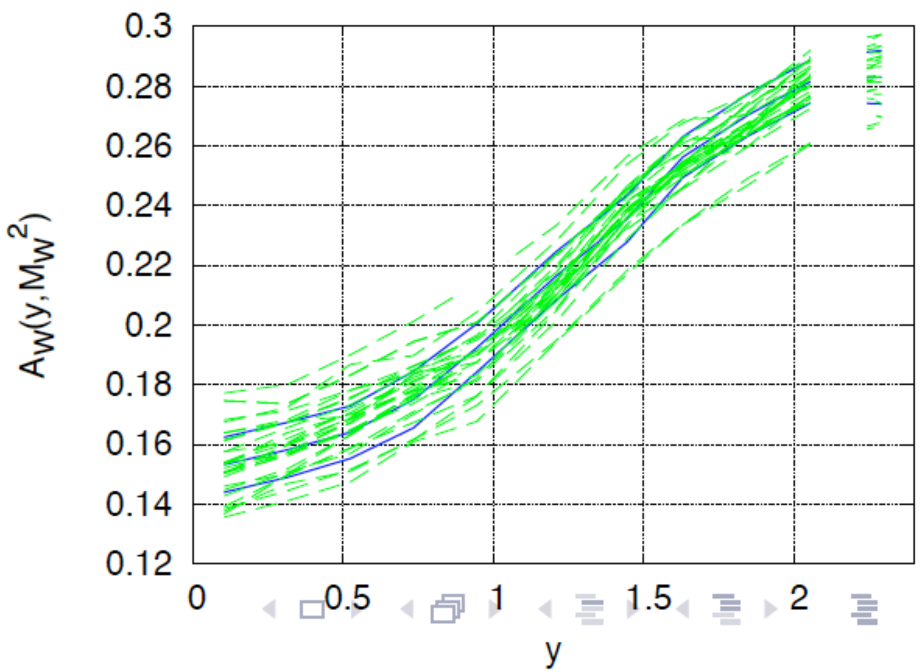
a) Download a set of NNPDF from LHAPDF http://projects.hepforge.org/lhapdf/

b) Each set contains an ensemble of N replicas (N = 100 or N = 1000)

c) Each replica is a set of PDFs given as a grid in (x,Q^2)

d) Each replica is a equally probable candidate PDF (uniform sampling)





NNPDF for the LHC
Instruction for use

$$f_0 = \langle f \rangle = \frac{1}{N}\sum_{k=1}^{N} f_k$$

set=0 in LHAPDF

set=1,..,N_{rep}

$$\mathrm{Var}[f] = \left\langle (f - \langle f \rangle)^2 \right\rangle = \frac{1}{N}\sum_{k=1}^{N} (f_k - f_0)^2$$

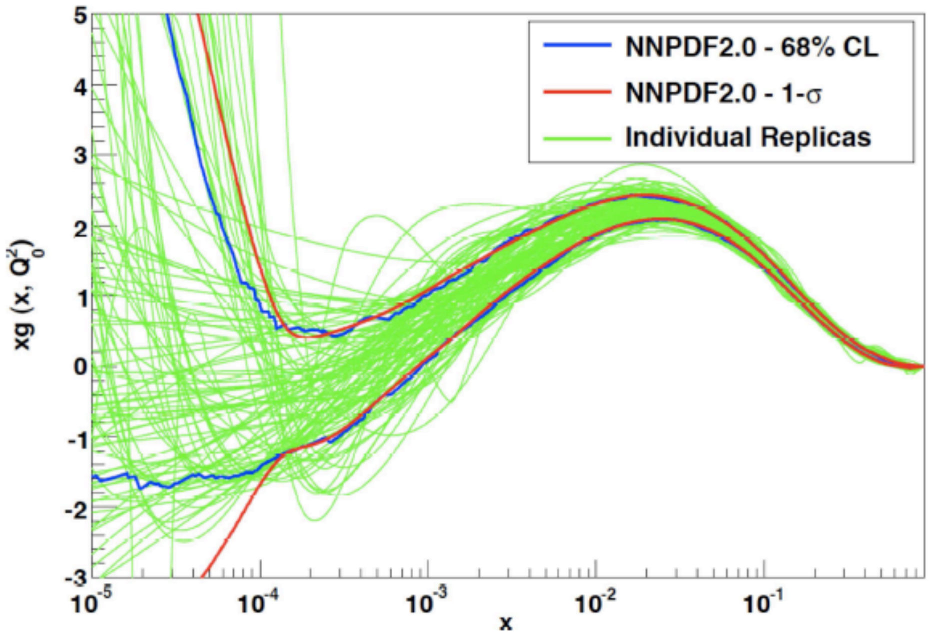
$$E[\sigma] = \langle \sigma[f] \rangle = \frac{1}{N}\sum_{k=1}^{N} \sigma[f_k] = \sigma[f_0]$$

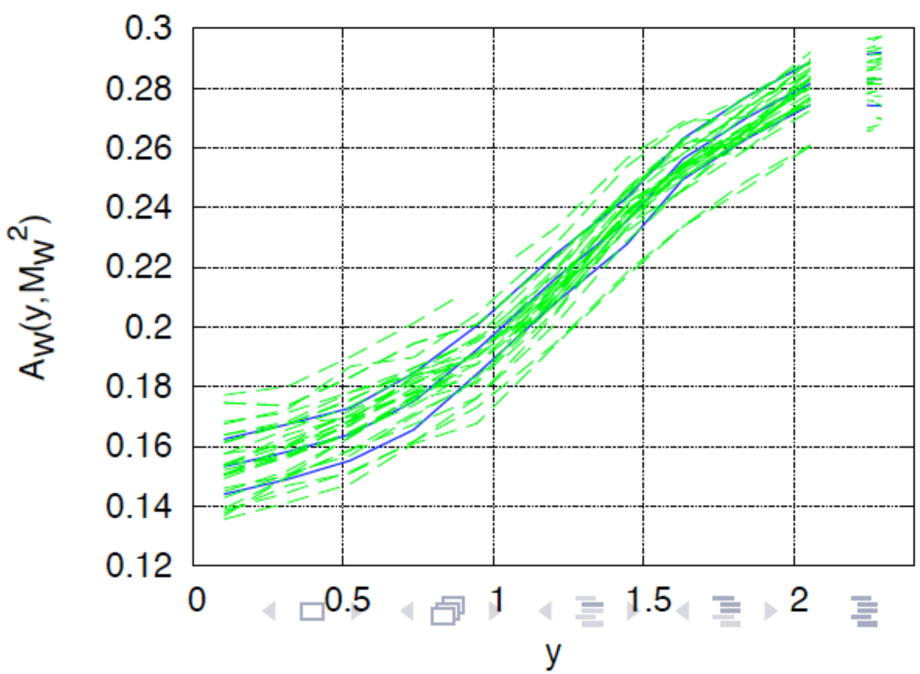
can be a xsec or any observable

$$\mathrm{Var}[\sigma] = \left\langle (\sigma[f] - \langle \sigma[f] \rangle)^2 \right\rangle = \frac{1}{N}\sum_{k=1}^{N} (\sigma[f_k] - \sigma[f_0])^2$$

PDFs

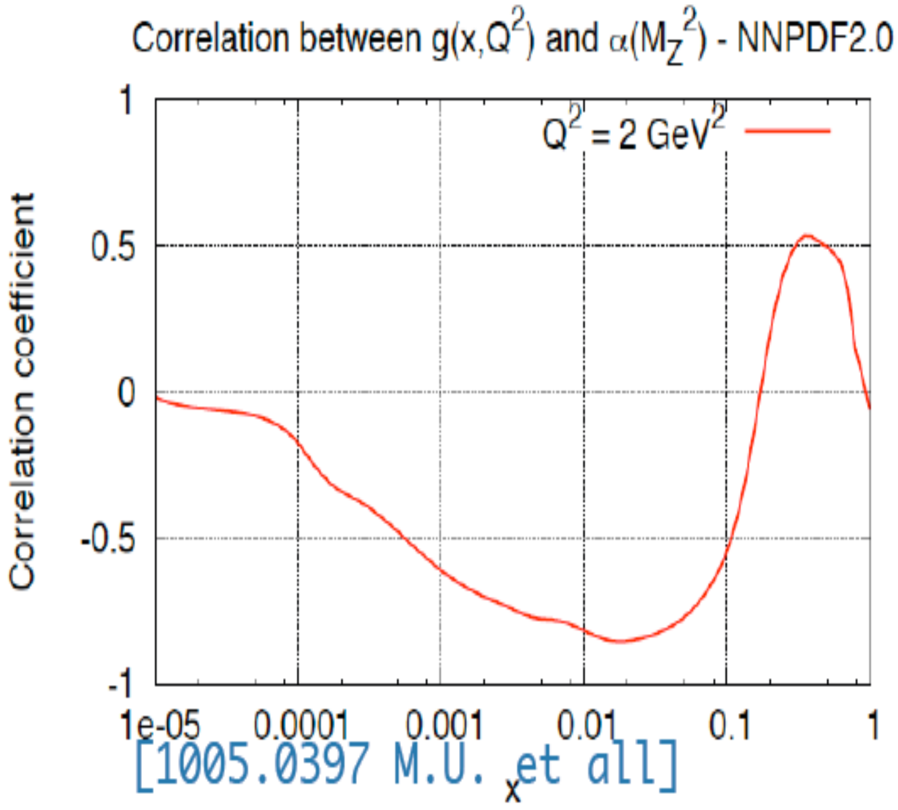
OBS



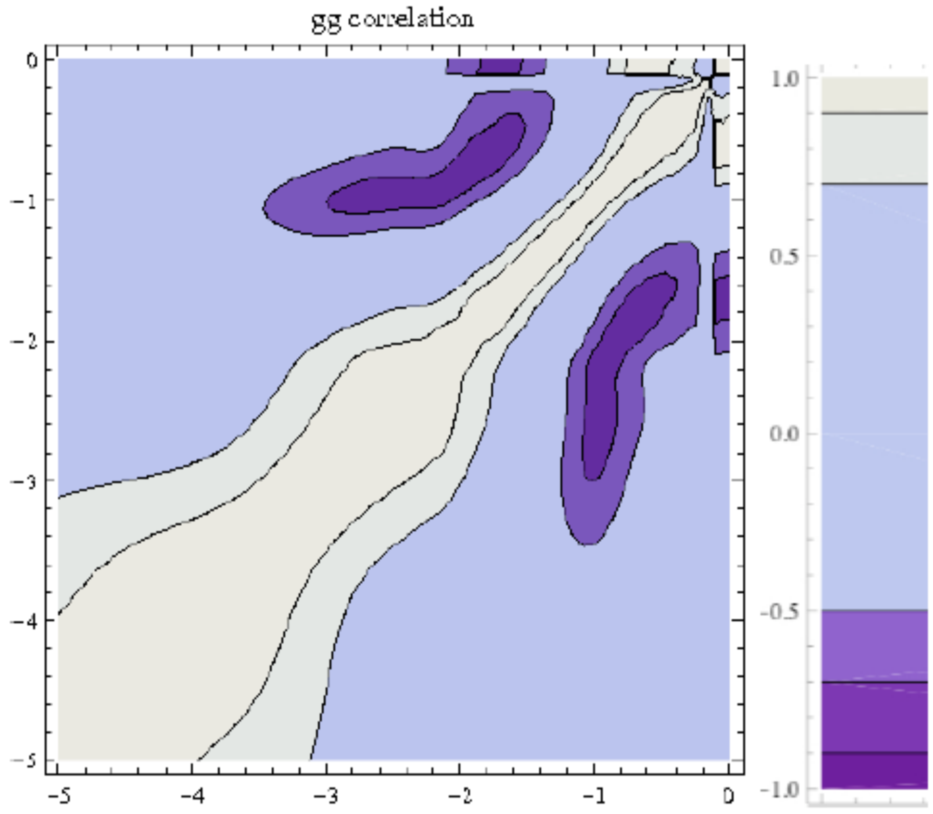


NNPDF for the LHC
Instruction for use

$$\mathrm{Corr}[f,f'] = \langle (f - \langle f \rangle)(f' - \langle f' \rangle) \rangle = \frac{1}{N}\sum_{k=1}^{N}(f_k - f_0)(f'_k - f'_0)$$



[1005.0397 M.U. et all]



NNPDF for the LHC

Instruction for use: some tips

If observable quadratic in PDFs take replica 0

- Central values:

$$E[\sigma_h] = \langle \sigma_h[f,f] \rangle = \frac{1}{N}\sum_{k=1}^{N} \sigma_h[f_k, f_k] \approx \sigma_h[f_0, f_0]$$

- Variances:

$$\begin{aligned} \mathrm{Var}[\sigma_h] &= \left\langle \left(\sigma_h[f,f] - \langle \sigma_h[f,f] \rangle\right)^2 \right\rangle \\ &= \frac{1}{N}\sum_{k=1}^{N}\left(\sigma_h[f_k,f_k] - \frac{1}{N}\sum_{k=1}^{N}\sigma_h[f_k,f_k]\right)^2 \\ &\approx \frac{1}{N}\sum_{k=1}^{N}\left(\sigma_h[f_k,f_0] + \sigma_h[f_0,f_k] - 2\sigma_h[f_0,f_0]\right)^2 \end{aligned}$$

If observable quadratic in PDFs and Gaussian approximate:
this way you can use less replicas

NNPDF for the LHC
Instruction for use: some F.A.Q.

Q: how many replicas N do I need?

A: depends on required accuracy: fluctuations fall as $1/\sqrt{N}$

typically use f_0 for central values, ~ 25-100 f_k for variances etc

Q: which replicas should I use?

A: any random selection! – all replicas are equally probable

Q: for hadronic xsecs, should I use the exact or approx formulae?

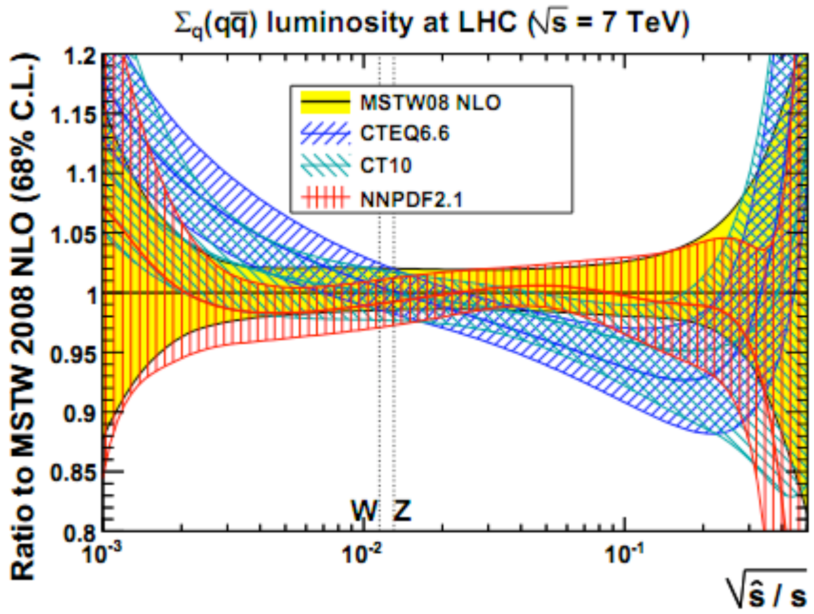
A: error from using approx is $O(Var/E^2)$

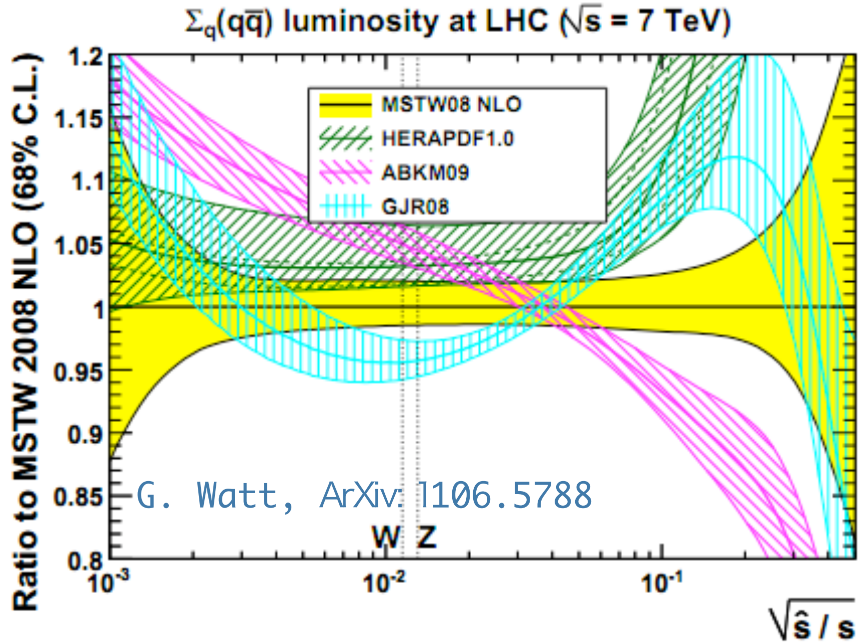
(so for typical 10% uncertainty, error is O(1%))

and variance formula neglects nonGaussian errors

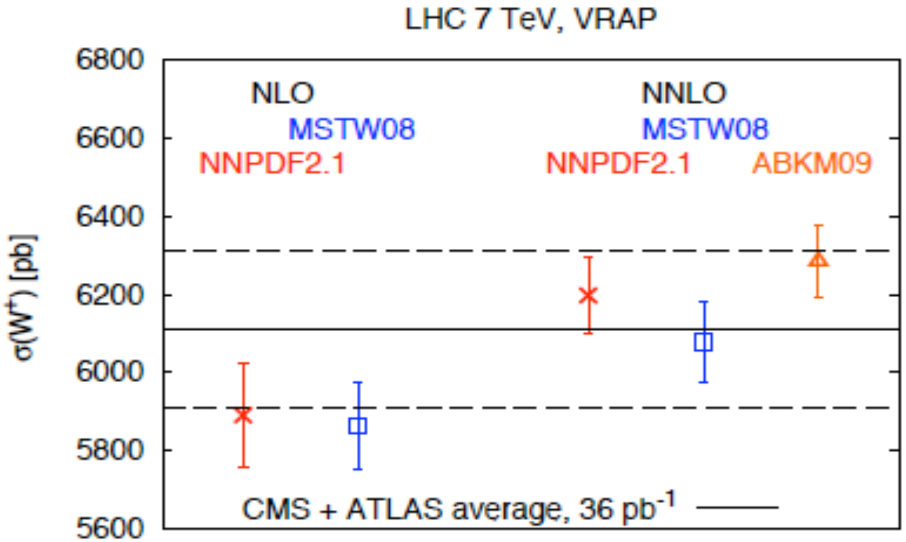
PDFs for the LHC

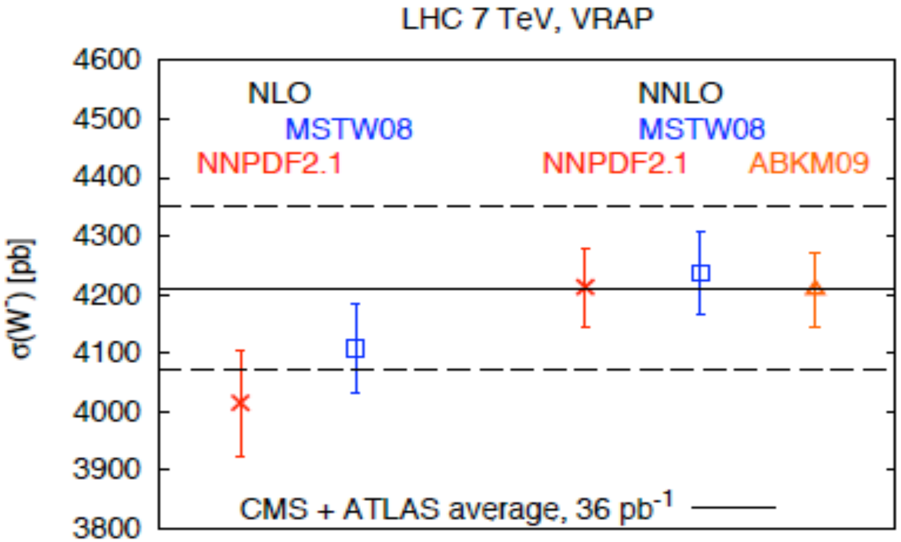
Some comparisons





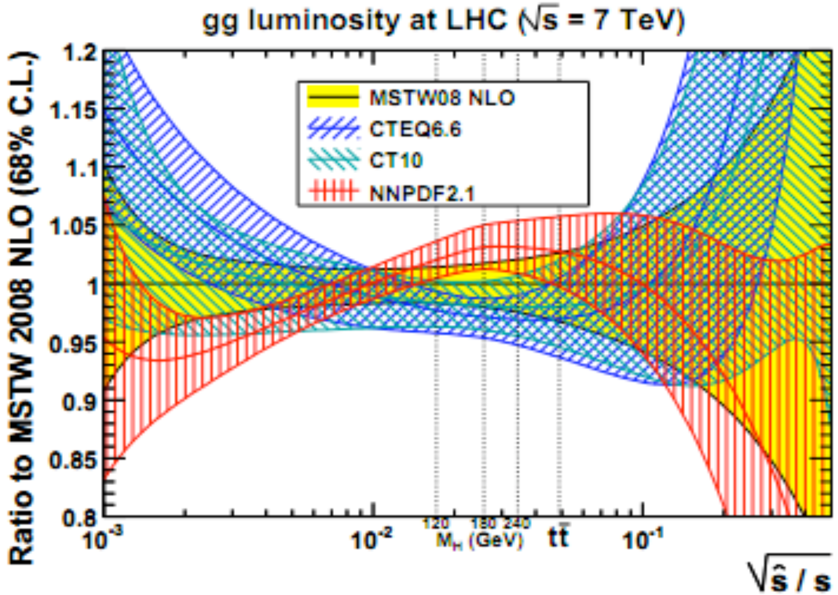
G. Watt, ArXiv 1106.5788

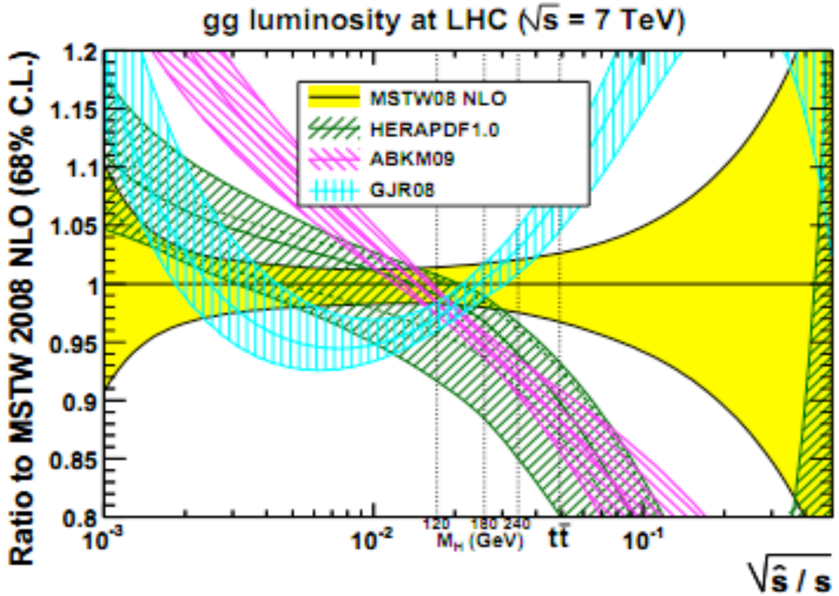


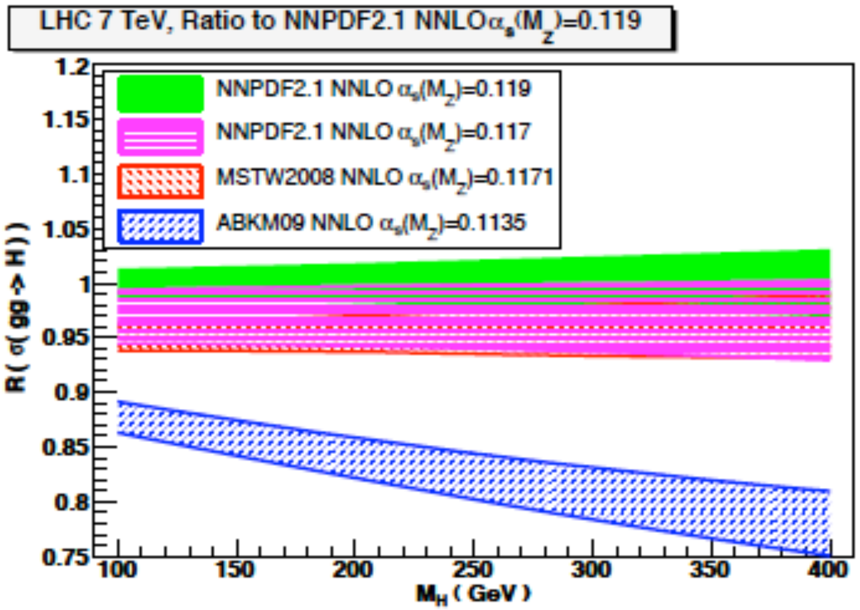


PDFs for the LHC

Some comparisons





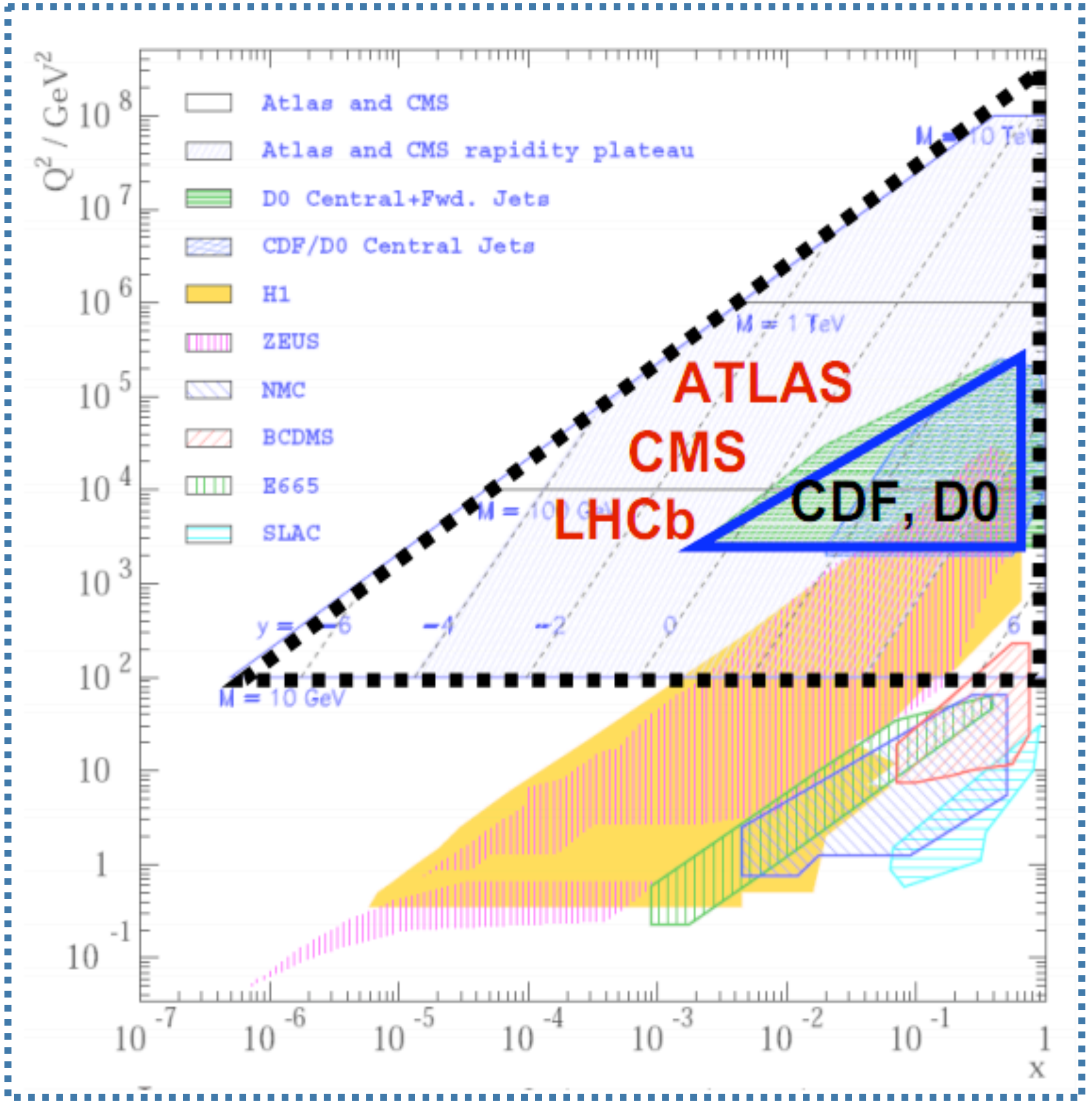


- Towards increasing agreement among global fits
- Choice of parameters such as α_S (and heavy quark masses) are increasingly relevant
- As LHC data become more precise they will be able to discriminate between different PDF sets

LHC for PDFs

- Constraints from LHC
- A collider-only analysis?
- Reweighting techniques
- Assessing impact of the LHC data
- The W lepton asymmetry case
- NNPDF2.2 set including LHC data

LHC collisions
and PDF uncertainties

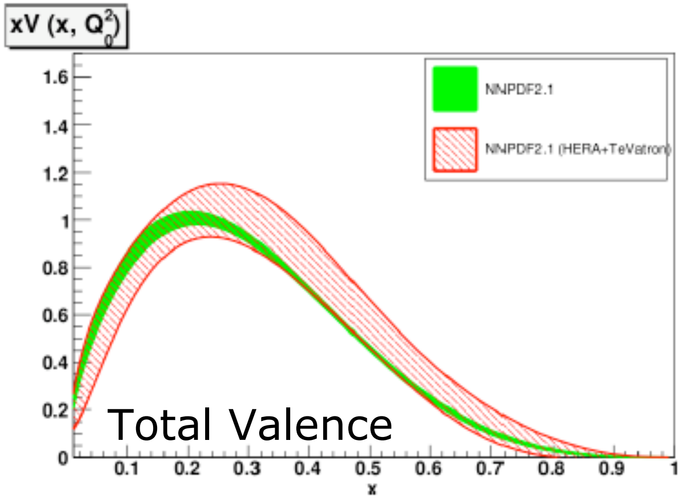


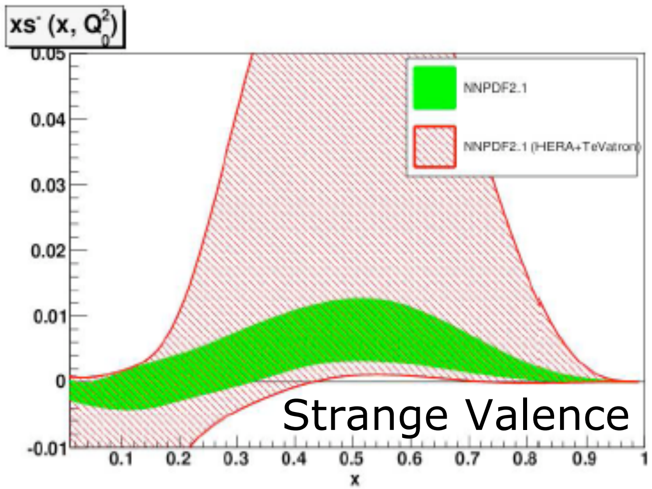
V. Radescu, DIS2011

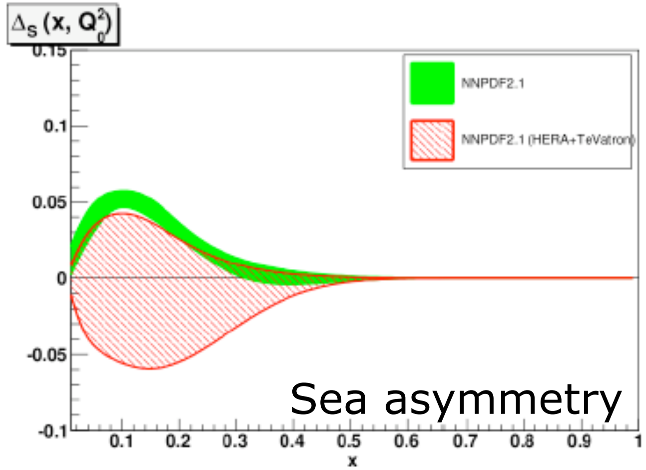
For many standard candle processes at the LHC theoretical uncertainty is dominated by PDF uncertainties

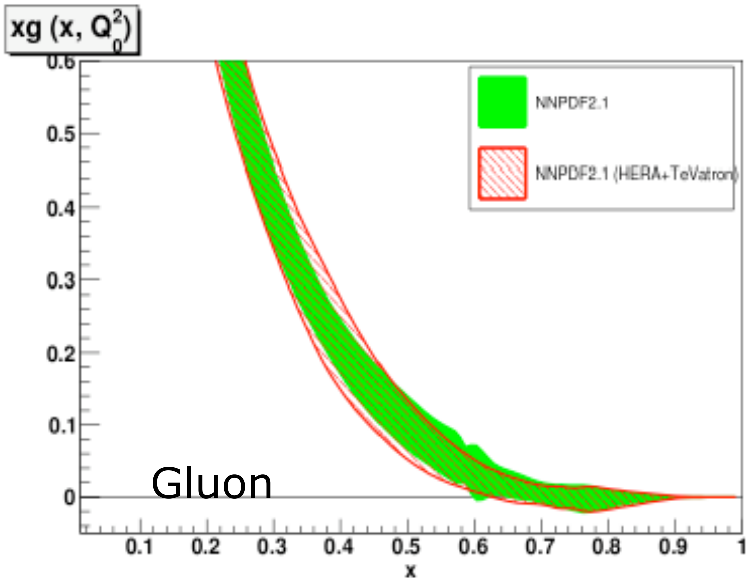
Can LHC data help in this direction by providing more info on PDFs and thereby reduce their uncertainties & discriminate between PDF sets?

LHC for PDFs
Towards a collider only fit?



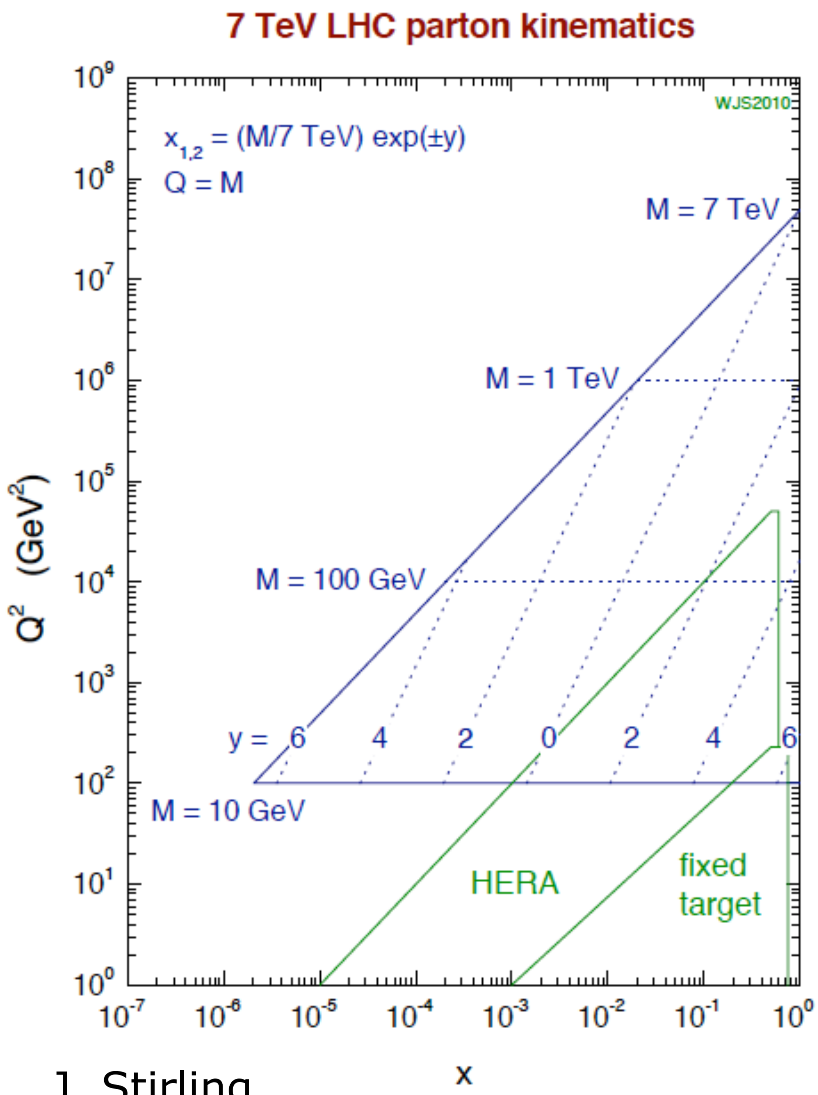






- No fixed target data ➔ No low energy troubles (nuclear corrections, higher twists...)
- HERA + Tevatron:
  - Good accuracy for gluon
  - Loss of accuracy for flavor separation and strange
  - What about HERA + Tevatron + LHC?

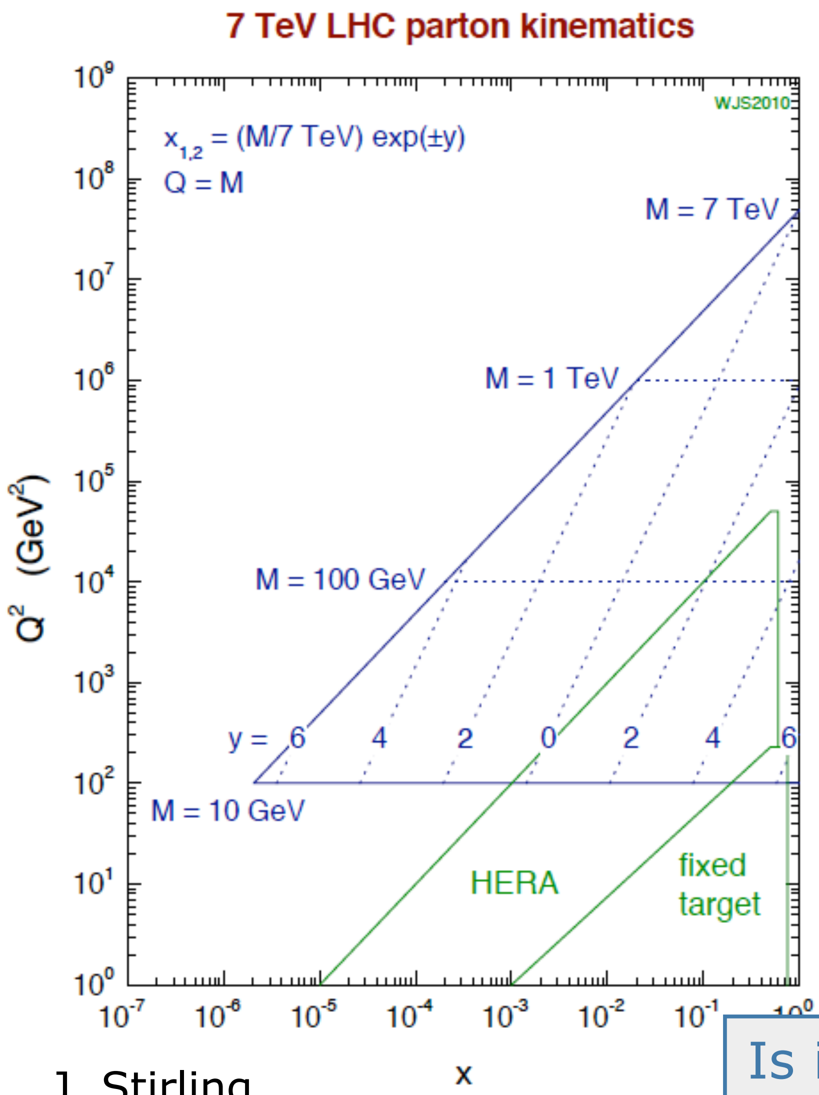
LHC for PDFs
Towards a collider only fit?



J. Stirling

- Medium and large x gluon
  - ✓ Prompt photons
  - ✓ Precision jets
- Light flavors and flavor separation at medium and small x
  - ✓ Low mass DY
  - ✓ Z rapidity distributions
  - ✓ W asymmetry
- Strangeness and heavy flavors
  - ✓ W+c
  - ✓ Z+c
  - ✓ Z+b

LHC for PDFs
Towards a collider only fit?



J. Stirling

- Medium and large x gluon
  - ✓ Prompt photons
  - ✓ Precision jets
- Light flavors and flavor separation at medium and small x
  - ✓ Low mass DY
  - ✓ Z rapidity distributions
  - ✓ W asymmetry
- Strangeness and heavy flavors
  - ✓ W+c
  - ✓ Z+c
  - ✓ Z+b

Is it possible to assess the impact of the new LHC data as they come out without necessarily refitting?

LHC for PDFs
A tool: Bayesian reweighting

R.D.Ball et al. ArXiv:1012.0836

- Inspired by Giele and Keller [hep-ph/9803393]

- The N_{rep} of a Monte Carlo fit give the probability density in the space of PDFs
- Expectation values are MC integrals, same for errors, correlations...
- We can assess the impact of including new data in the fit by updating the probability density distribution

- It can be done also in Hessian parton fit [LHCb De Lorenzi - McNulty studies]

Refitting:

✗ Whenever add new data, need to do full refitting, tune parametrization and statistic treatment

✗ Can be done only by PDF fitting collaborations themselves.

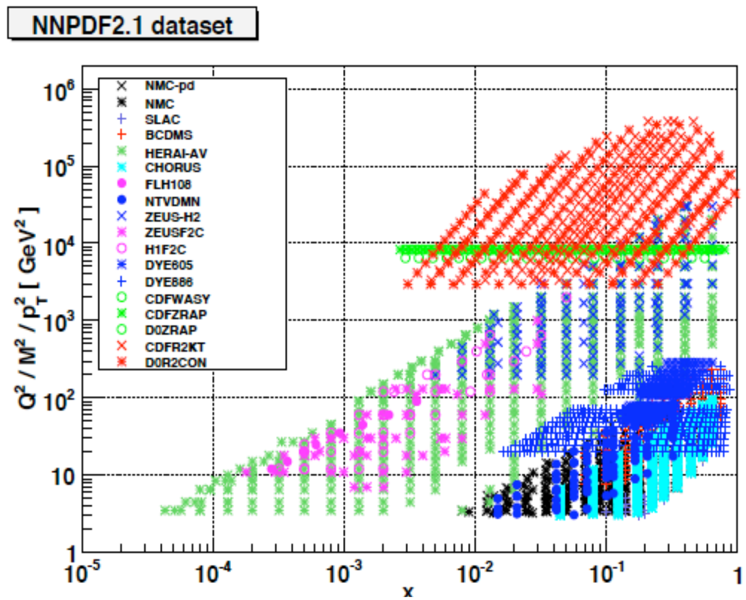
Reweighting:

✓ Immediate: no need to refit

✓ Anybody can do it just evaluating weights with each replica of a PDF set and producing a new PDF set through unweighting

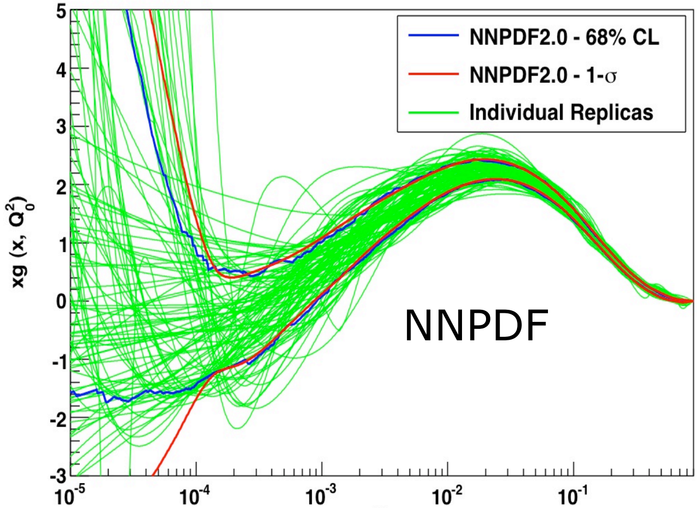
LHC for PDFs

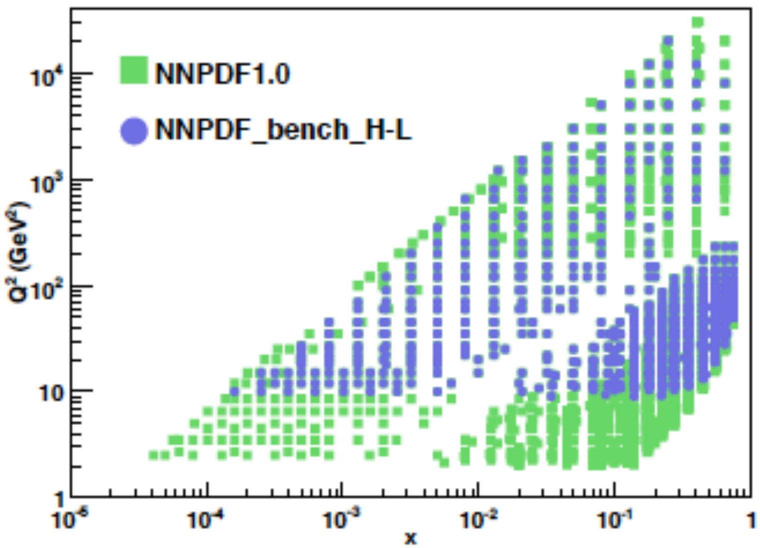
Reweighting: how does it work?

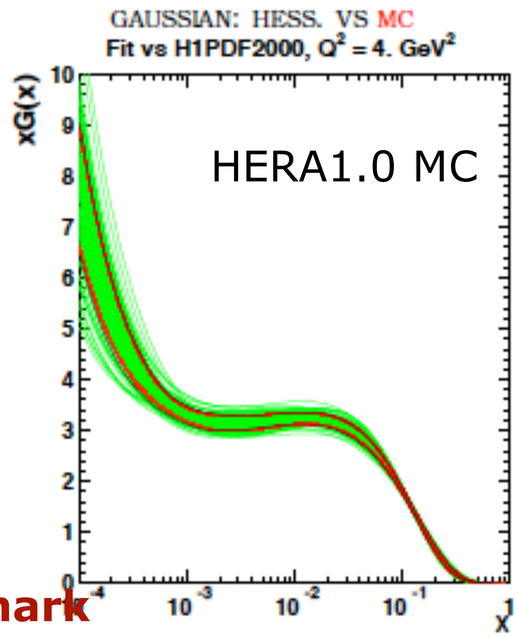


1) Uniform sampling

$w_k = 1$

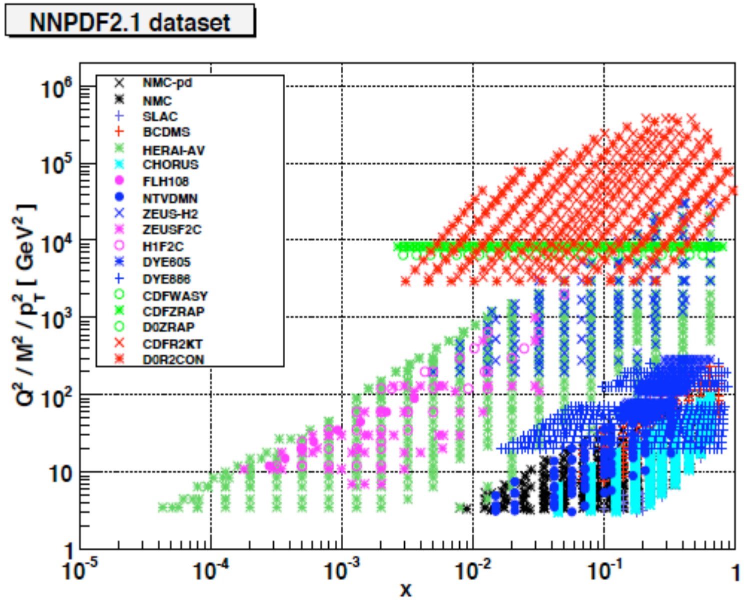






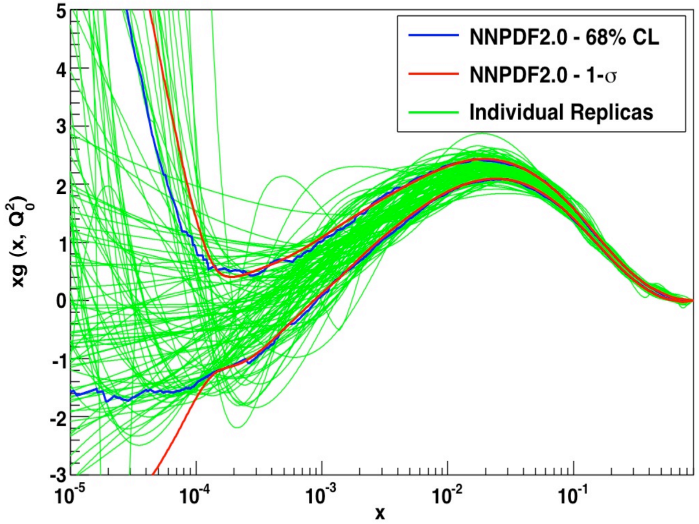
HERA-LHC 2008 benchmark

LHC for PDFs
Reweighting: how does it work?

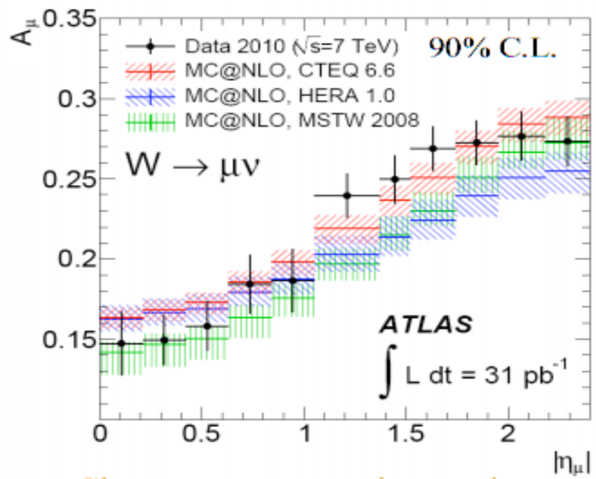


1) From data of global analysis: uniform sampling

$w_k = 1$

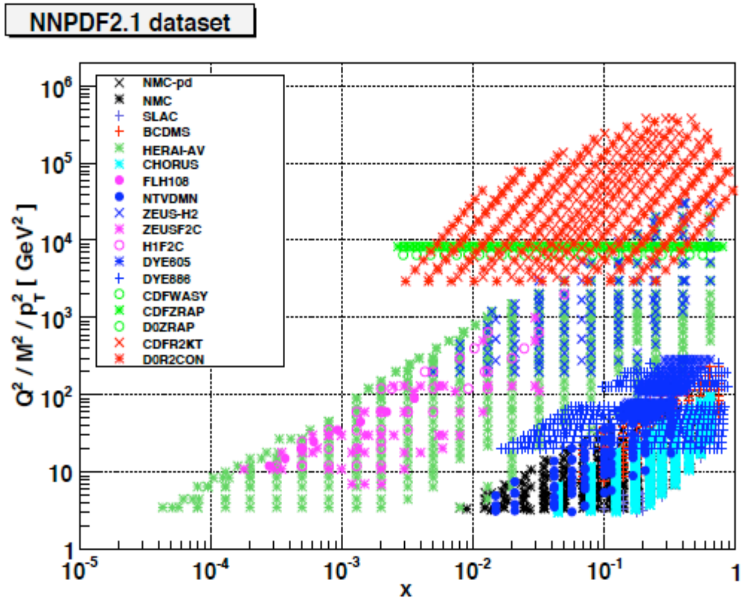


2) Suppose we want to check effect of ATLAS W asymmetry data on PDFs



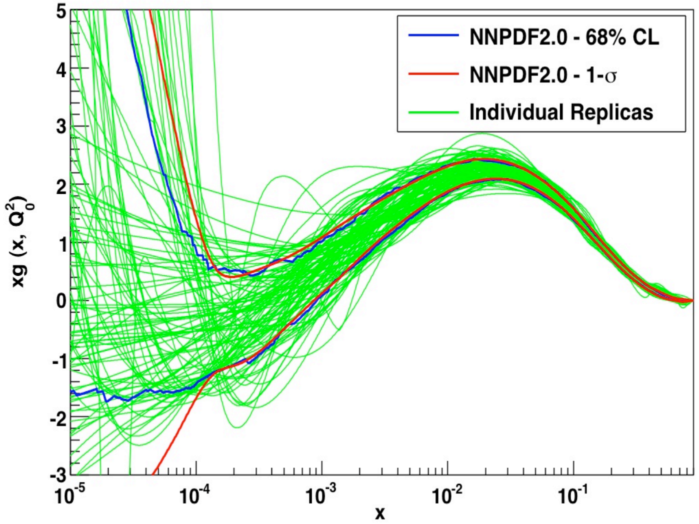
What should we do?

LHC for PDFs
Reweighting: how does it work?

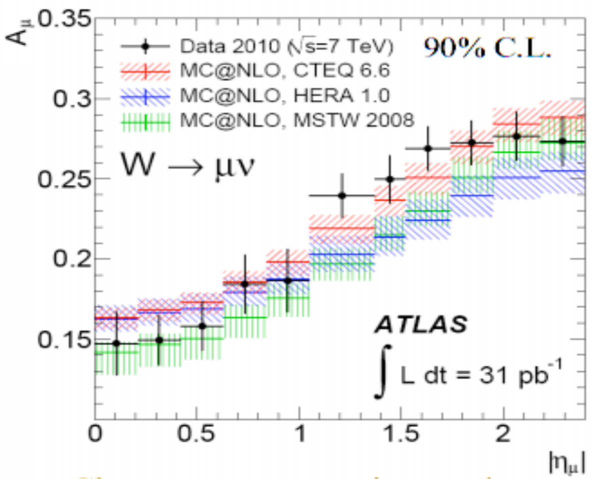


1) From data of global analysis: uniform sampling

$w_k = 1$



2) Suppose we want to check effect of ATLAS W asymmetry data on PDFs



☑ Compute χ^2_k d.o.f. to the NEW data for each error set (or replica) k of the initial ensemble with any of your favorite code (MCFM, DYNNLO, APPLGRID …)

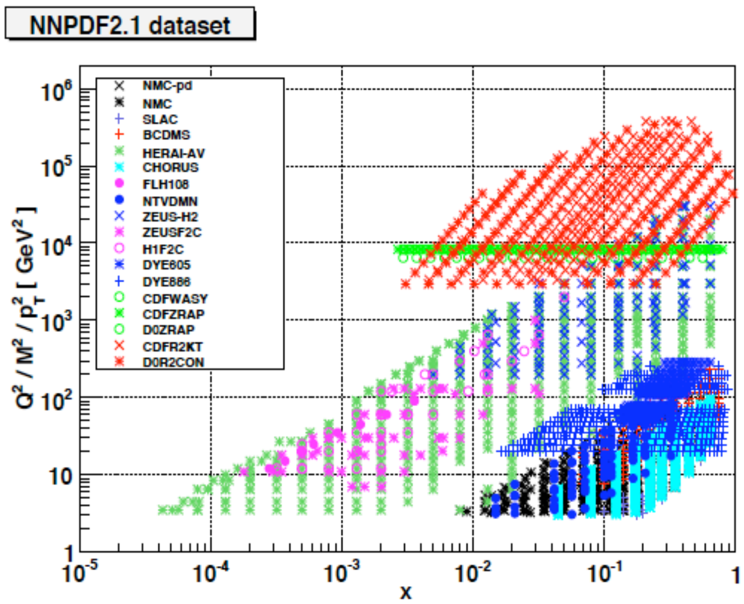
☑ Assign a weight w_k

$$w_k \propto (\chi^2_k)^{\frac{1}{2}(n-1)} e^{-\frac{1}{2}\chi^2_k}$$

higher weight for replicas with $\chi^2 \sim 1$

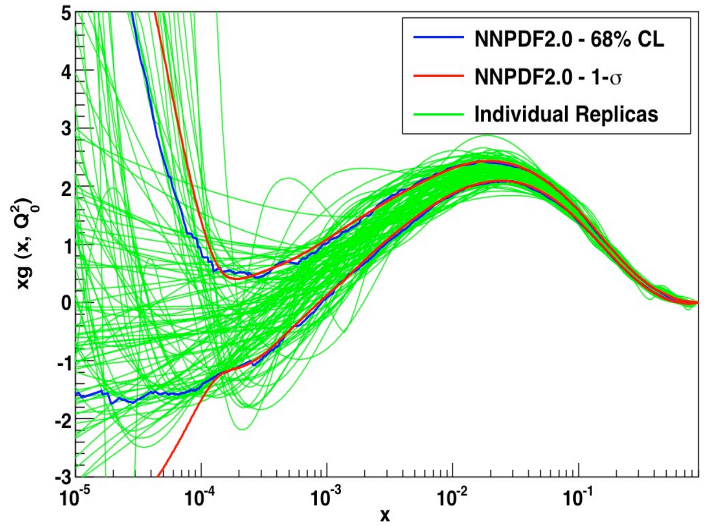
LHC for PDFs

Reweighting: how does it work?

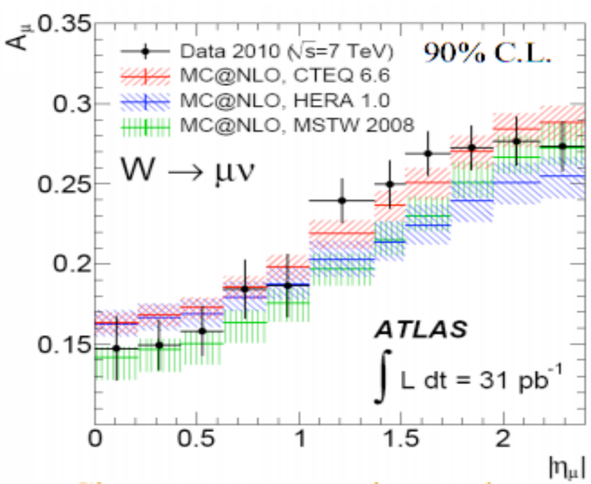


1) From data of global analysis: uniform sampling

$w_k = 1$



2) Suppose we want to check effect of ATLAS W asymmetry data on PDFs

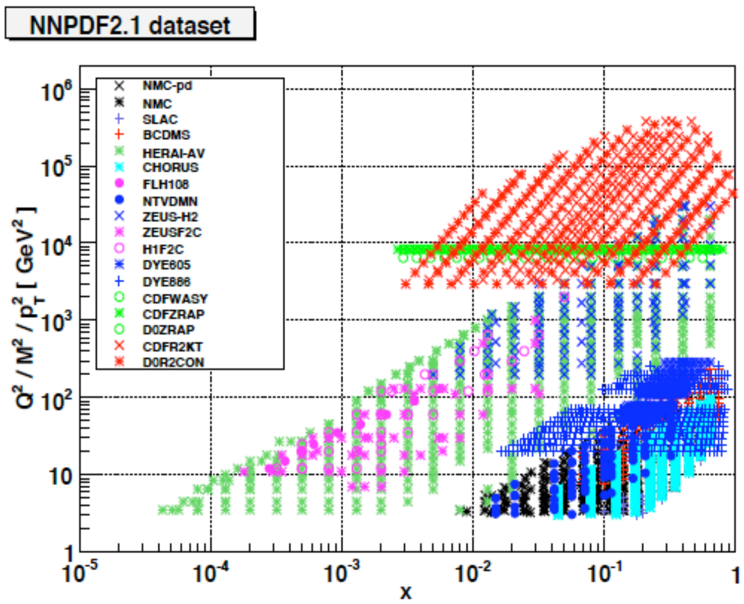


From uniform sample to weighted sample!

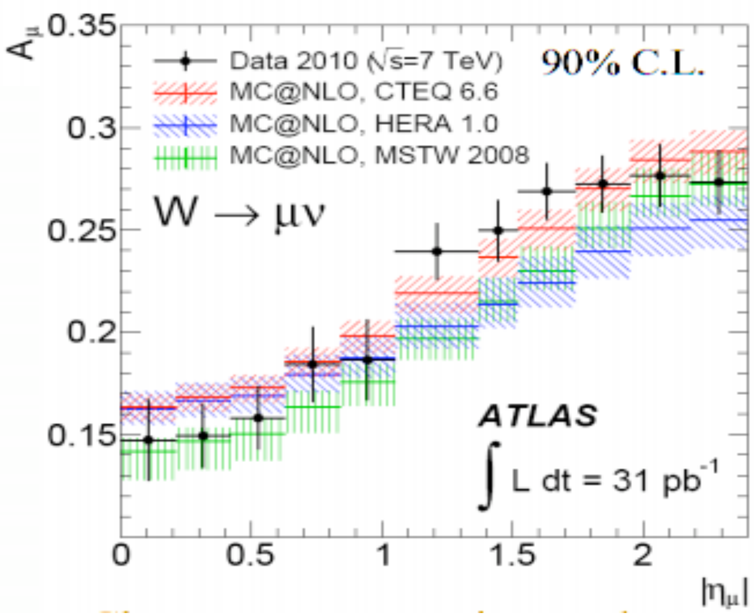
$$\langle \mathcal{O} \rangle = \int \mathcal{O}[f]\,\mathcal{P}(f)\,Df = \frac{1}{N}\sum_{k=1}^{N} \mathcal{O}[f^{(k)}]$$

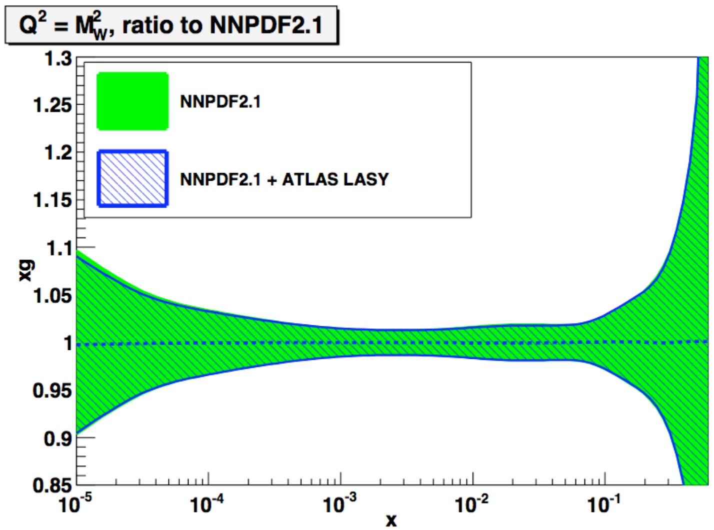
$$\langle \mathcal{O} \rangle_{\rm new} = \int \mathcal{O}[f]\mathcal{P}_{\rm new}(f)\mathcal{D}f = \frac{1}{N}\sum_{k=1}^{N} w_k \mathcal{O}[f^{(k)}]$$

LHC for PDFs
Reweighting: how does it work?



+





Can

✔ check if and how PDFs, their uncertainties and predictions for observable change upon addition of new data

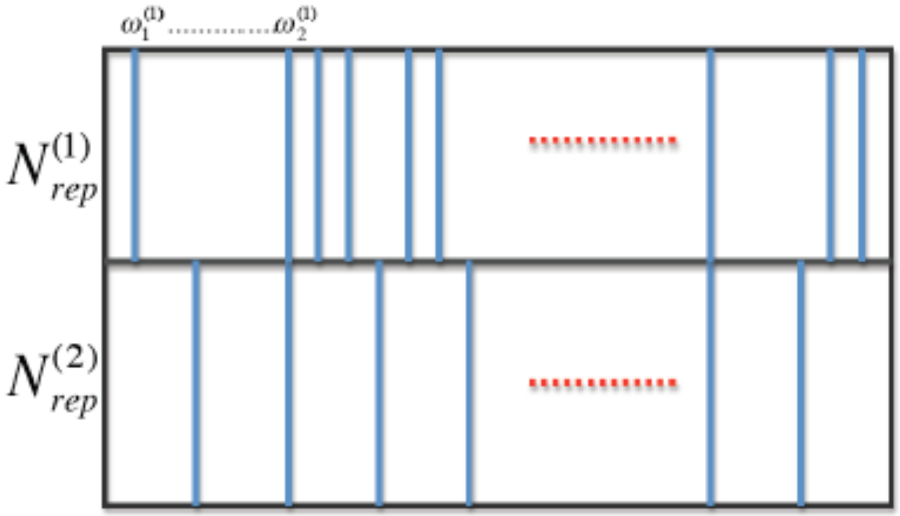
✔ assess how much constraining are the new data by computing N_{eff} replicas: if N_{eff} << N new data are constraining or incompatible

LHC for PDFs
Reweighting: how does it work?

Last step: unweighting

From reweighted ensemble to unweighted one for usual LHAPDF interface

- From reweighted set construct unweighted "normal" set
- Reweighted ensemble and uniform ensemble obtained by unweighting are statistically equivalent if N_{eff} is large enough
- NNPDF2.2 is the first PDF analysis which will include LHC data



$$W_k \equiv W_{k-1} + \frac{w_k}{N_{\rm rep}^{(1)}} = \sum_{j=0}^{k} \frac{w_j}{N_{\rm rep}^{(1)}}$$

$$\tilde{w}_k = \sum_{j=1}^{N_{\rm rep}^{(2)}} \Theta\left(\frac{j}{N_{\rm rep}^{(2)}} - W_{k-1}\right)\Theta\left(W_k - \frac{j}{N_{\rm rep}^{(2)}}\right)$$

$$\sum_k \tilde{w}_k = N_{\rm rep}^{(2)}$$

↲ number of effective replicas

- Each w_k can be either $0,1,\ldots, N^{(2)}_{rep}$
- Take w_1 copies of rep 1, w_2 copies of rep 2...
- End up with uniform sampling of $N^{(2)}_{rep}$

LHC for PDFs

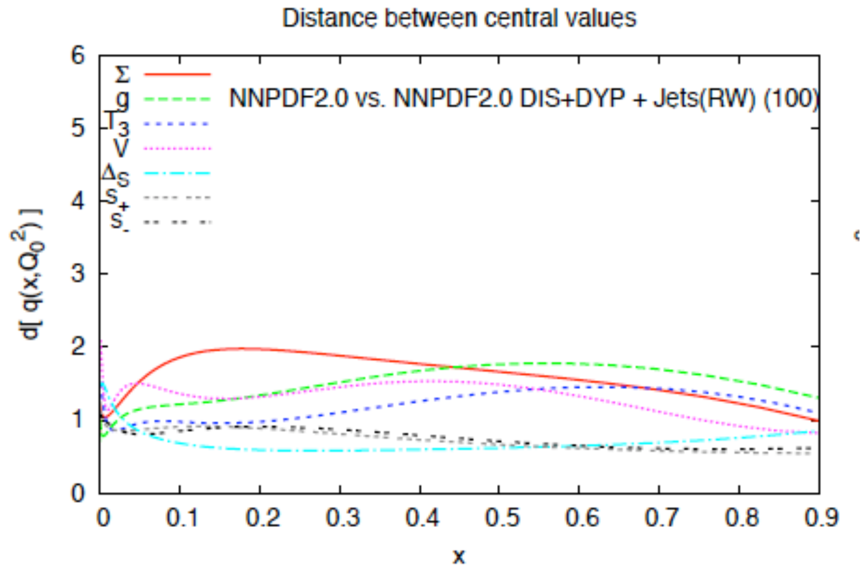
Reweighting: does it actually work?

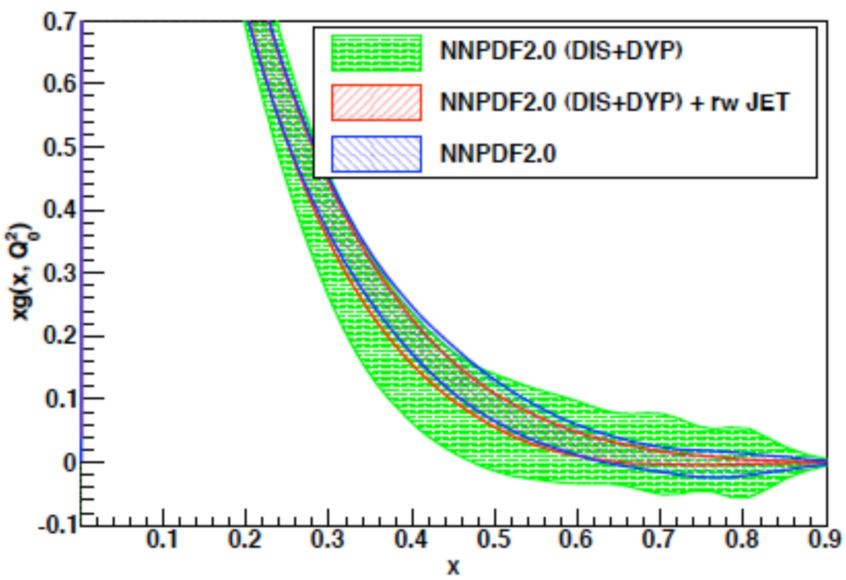
Test #1:

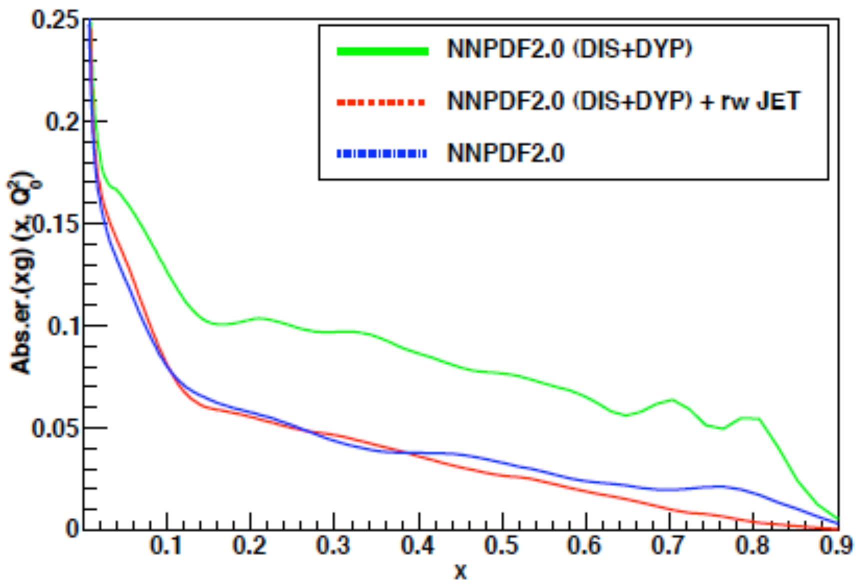
Check whether it is equivalent to refitting in a well-known case

P_{old} : NNPDF2.0 analysis based on DIS+DY data only

P_{new} : Add Tevatron Run II inclusive JET data through RW and refitting





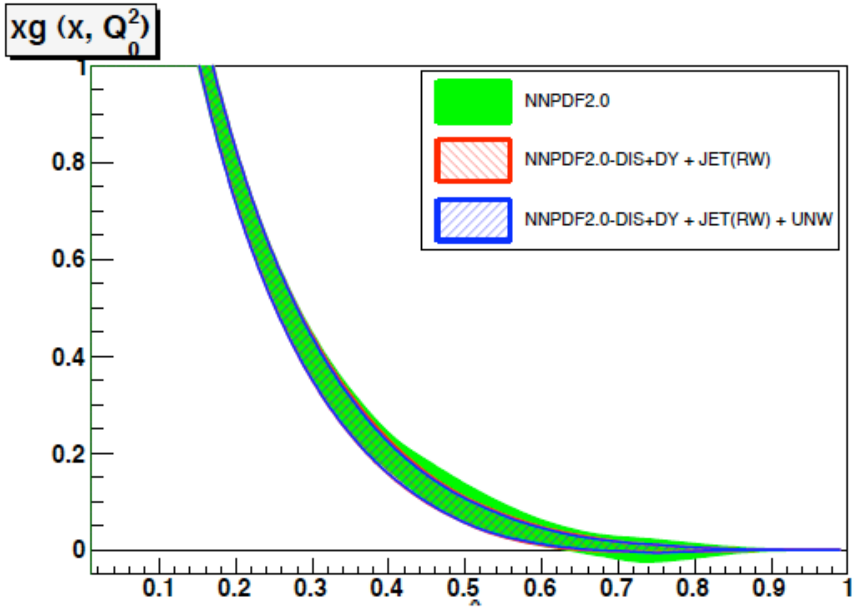


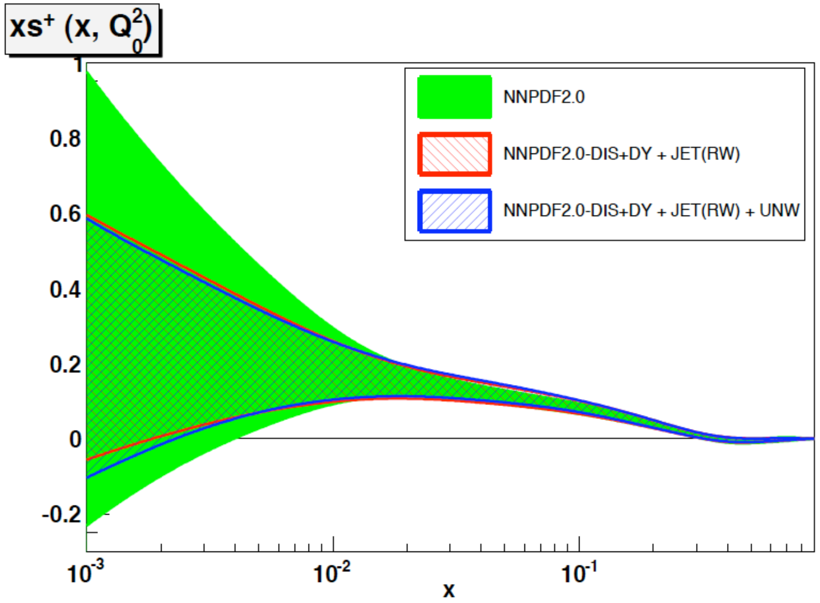
LHC for PDFs
Reweighting: does it actually work?

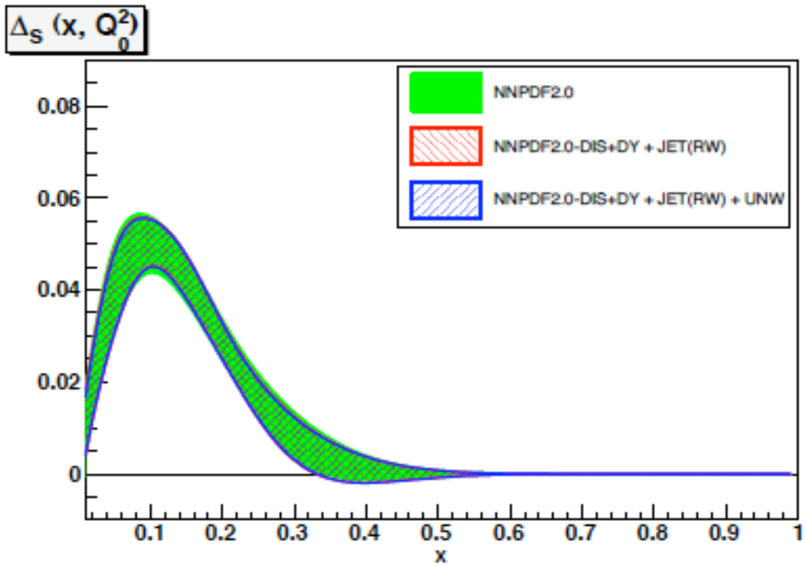
Test ♯2:

Check the unweighting procedure

1) Start from NNPDF2.0 DIS+DY data only
2) Reweight jets data
   NNPDF2.0 DIS+DY+JET(RW)
3) Unweight the reweighted ensemble
   NNPDF2.0 DIS+DY+JET(RW)+UNW





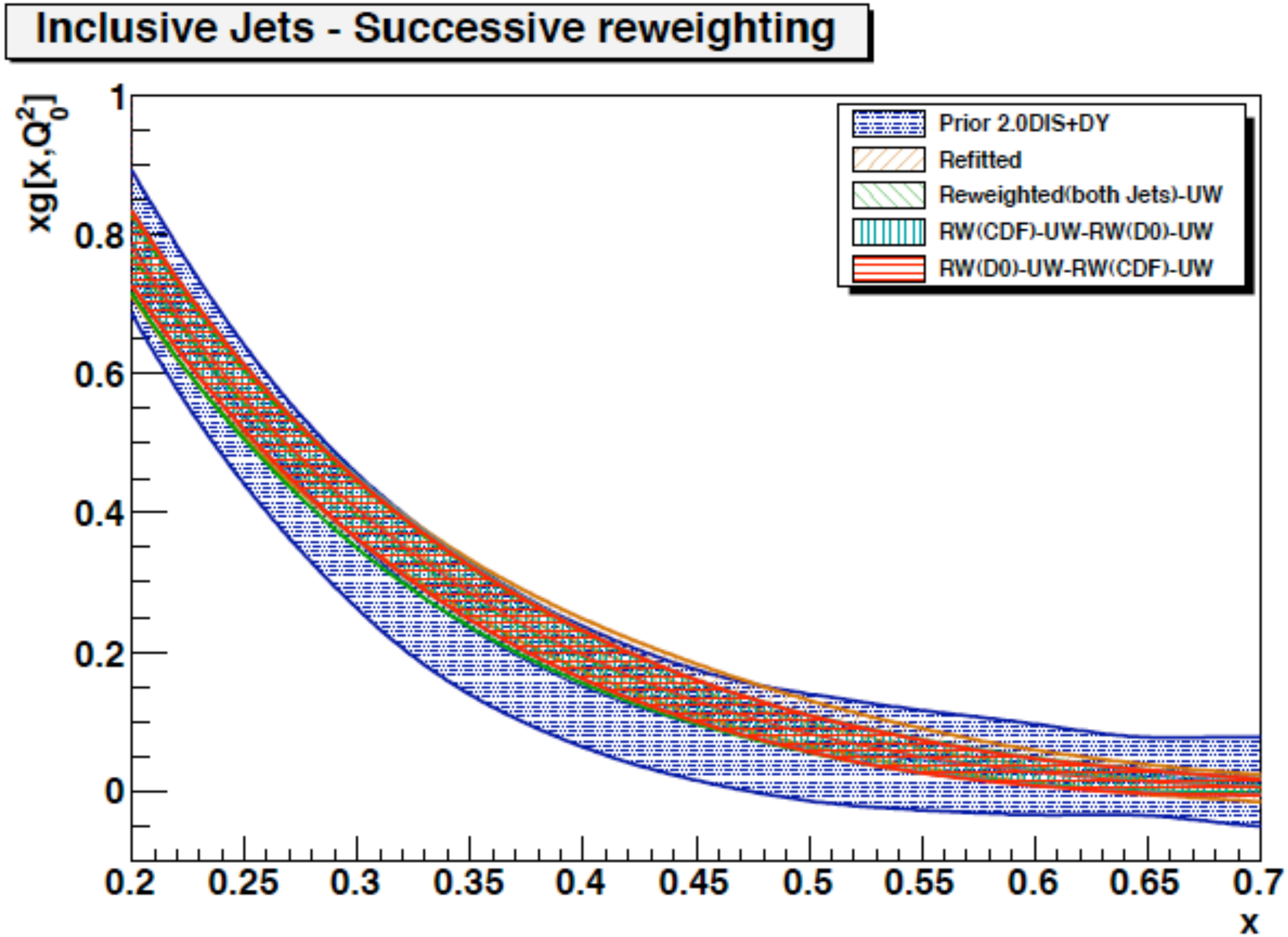


LHC for PDFs
Reweighting: does it actually work?

Test ＃3:

Check commutativity

1) Start from NNPDF2.0 DIS+DY only fit
2) Reweight CDF and D0 jet data at the same time and unweight
3) Rewight CDF data, unweight, reweight D0 data on unweighted set and unweight again
4) Do it the other way round



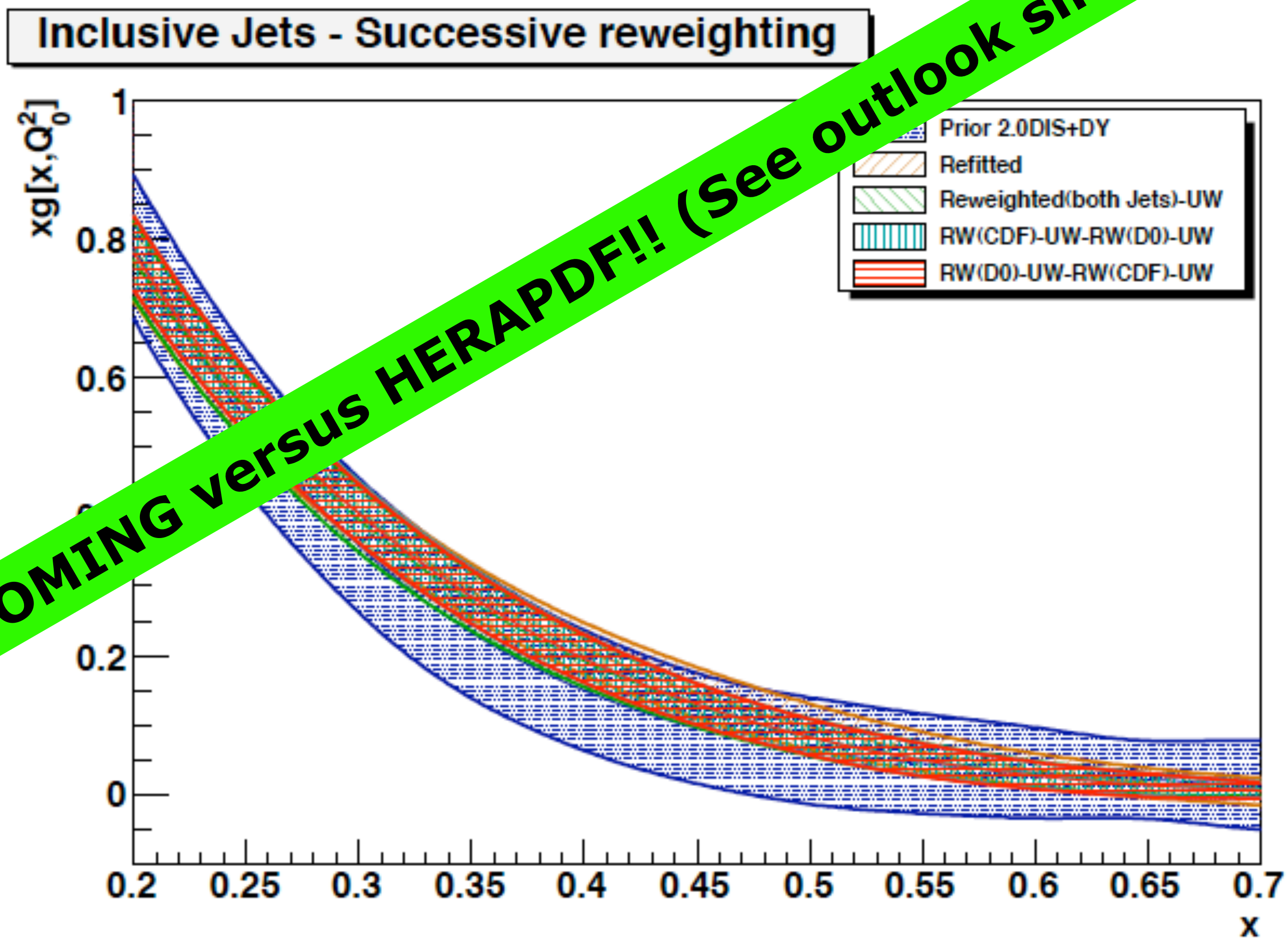
LHC for PDFs

Reweighting: does it actually work?

Test ♯3:

Check commutativity

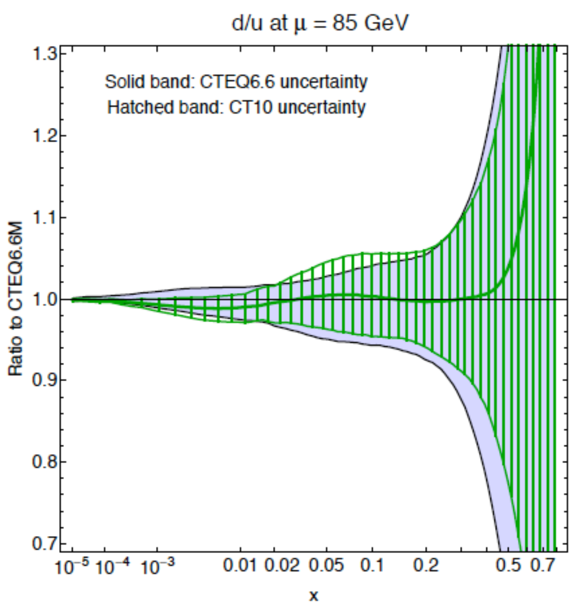
1) Start from NNPDF2.0 DIS+DY only fit
2) Reweight CDF and D0 jet data at the same time and unweight
3) Rewight CDF data, unweight, reweight D0 data on unweighted set and unweig again
4) Do i ther way



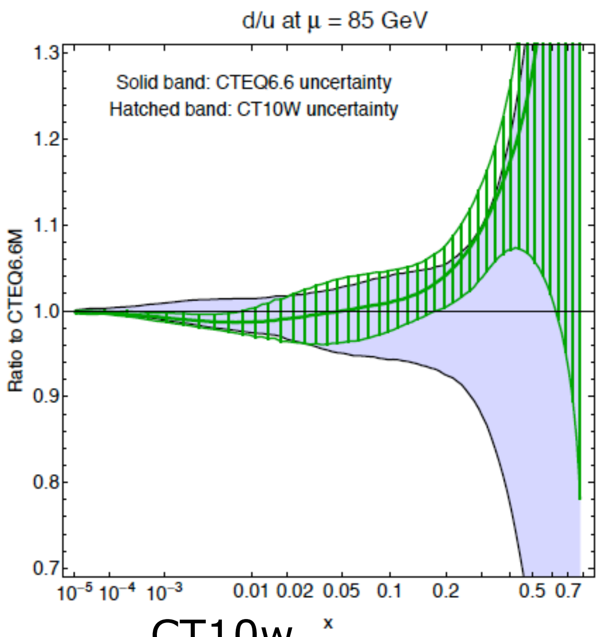
MORE TESTS UPCOMING versus HERAPDF!! (See outlook slide)

Some examples
Reweighting Tevatron W lepton asymmetry

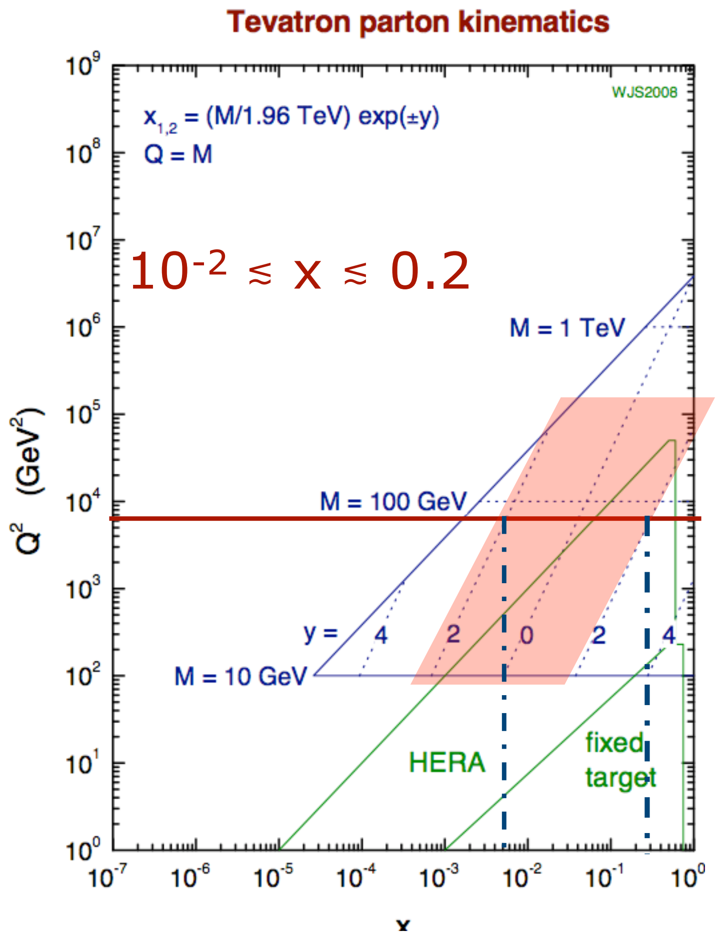
- ✓ CDF W charge asymmetry is fitted in NNPDF2.1
  [ArXiv:0901.2169]
- ➔ D0 muon charge asymmetry in single p_T^μ bin
  [ArXiv:0709.4254]
- ➔ D0 electron charge asymmetry combined and separated p_T^e bins :
  [ArXiv: 0807.3367]



CT10



CT10w



$$\frac{u(x_1)d(x_2) - d(x_1)u(x_2)}{u(x_1)d(x_2) + d(x_1)u(x_2)}$$

Some examples
Reweighting Tevatron W charge asymmetry

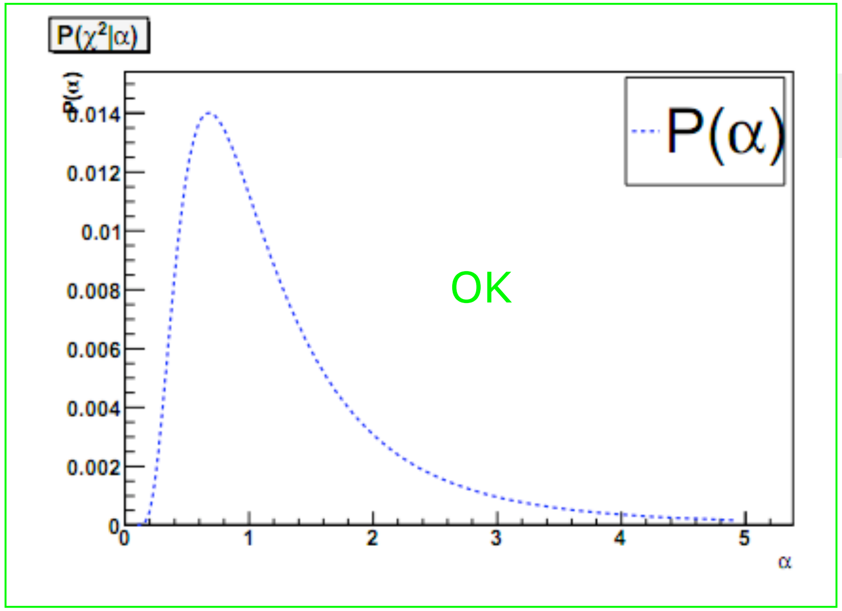
Reweighting analysis
[NNPDF, ArXiv:1012.0836]

✓ It is possible to include D0 lepton asymmetry inclusive data in global analyses: no need of producing separate sets

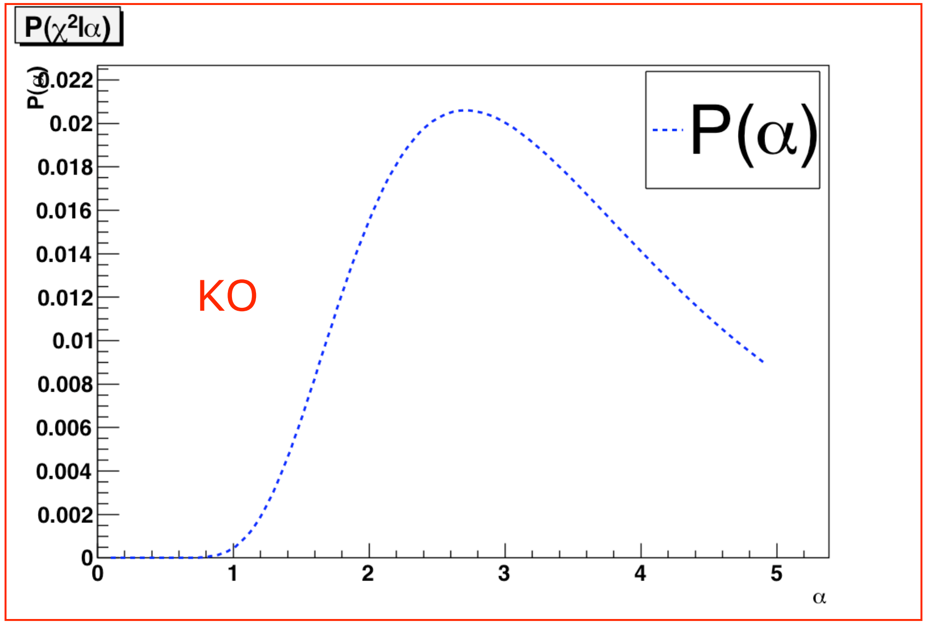
✓ Issues with exclusive bins

- Need to encourage use of least sensitive observable – either a very inclusive lepton asymmetry or the W asymmetry itself

H. Schellman, DIS 2011



Inclusive bins in p^T_{muon}



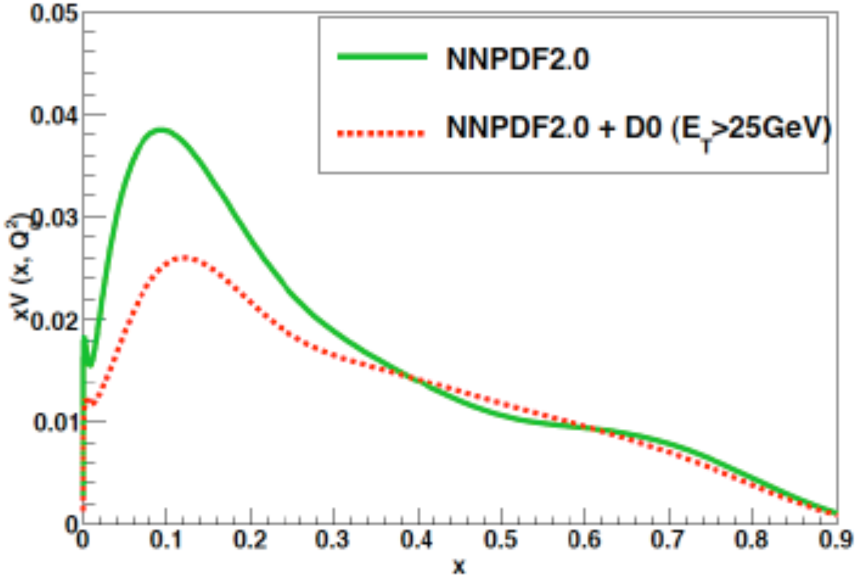
Exclusive bins in p^T_{el}

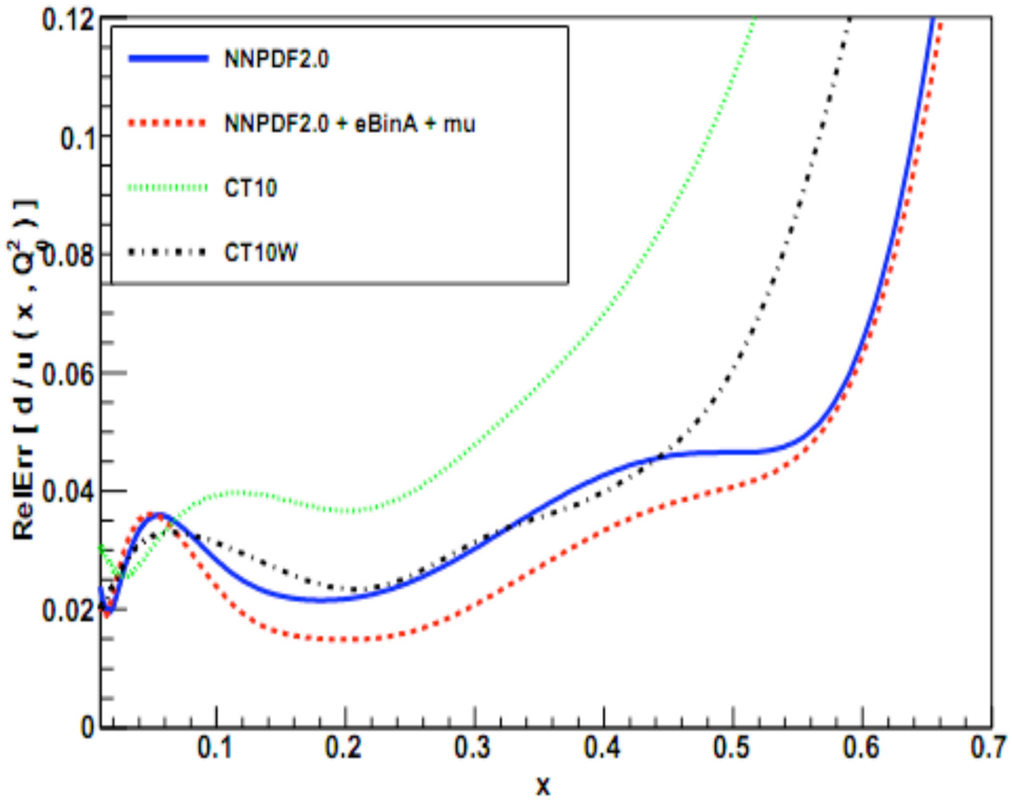
Some examples
Reweighting Tevatron W charge asymmetry

Reweighting analysis
[NNPDF, ArXiv:1012.0836]

✓ It is possible to include D0 lepton asymmetry inclusive data in global analyses: no need of producing separate sets

✓ Issues with exclusive bins





Reduction of the valence PDFs uncertainty and on the d/u ratio!

Some examples
Reweighting the LHC W charge asymmetry

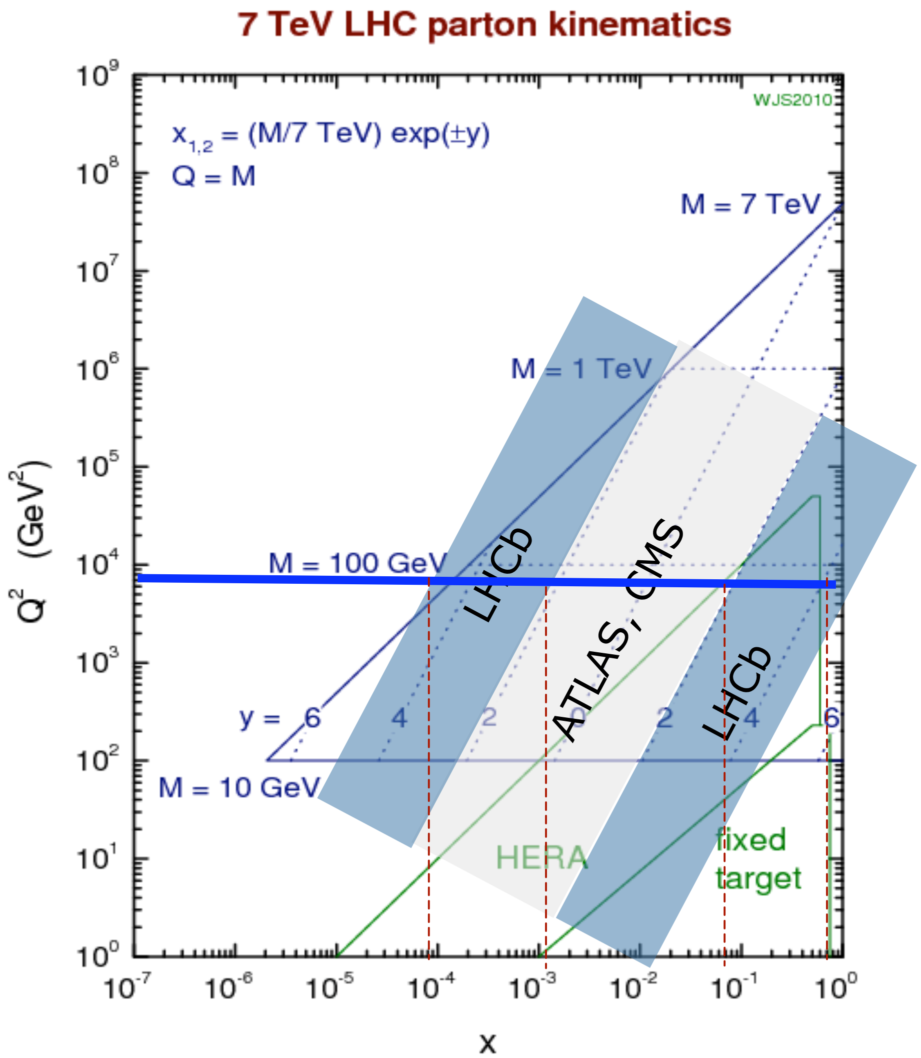
$$A_l = \frac{d\sigma(l^+)/dy(l^+) - d\sigma(l^-)/dy(l^-)}{d\sigma(l^+)/dy(l^+) + d\sigma(l^-)/dy(l^-)}$$

$$\frac{u(x_1)\bar{d}(x_2) - d(x_1)\bar{u}(x_2)}{u(x_1)\bar{d}(x_2) + d(x_1)\bar{u}(x_2)}$$

✓ **ATLAS:**
W muon charge asymmetry (31pb^{-1})
ArXiv: 1103.2929

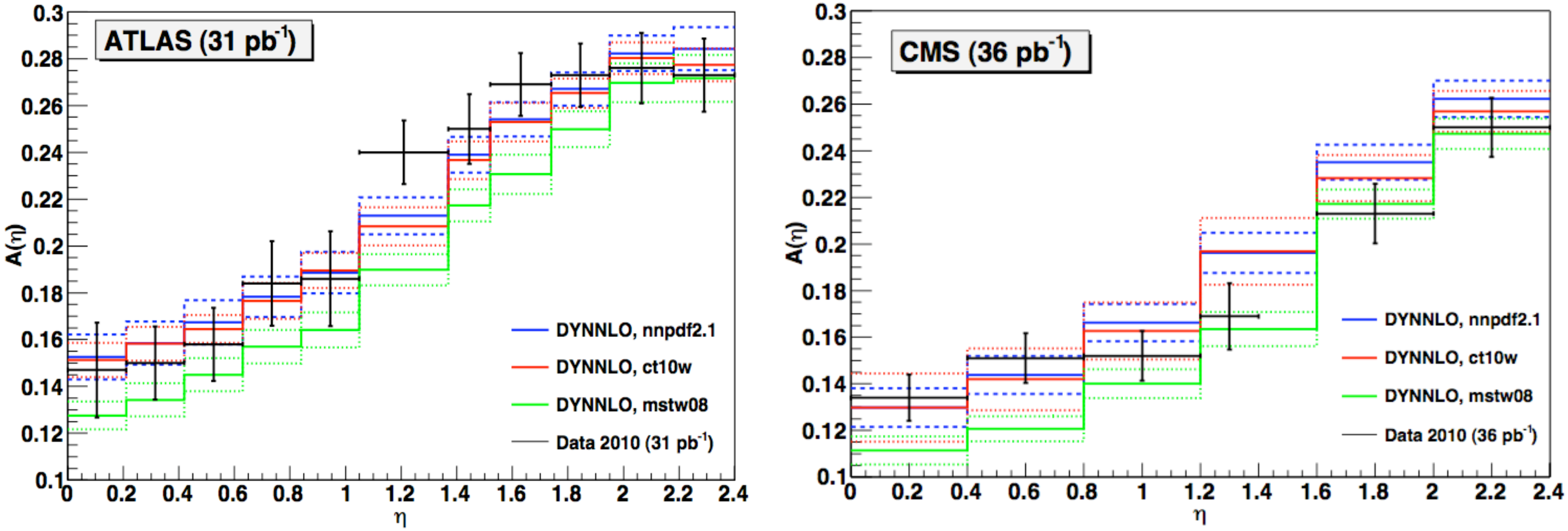
✓ **CMS:**
W muon and electron charge asymmetry (36pb^{-1})
ArXiv: 1103.3470

× **LHCb:**
Preliminary forward W muon charge asymmetry (16.5pb^{-1})
Not corrected for FSR radiation



Some examples

Reweighting the LHC W charge asymmetry

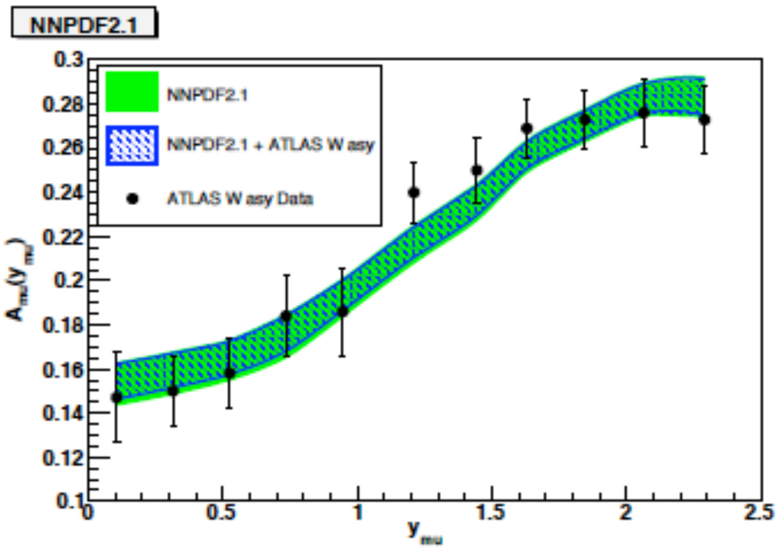


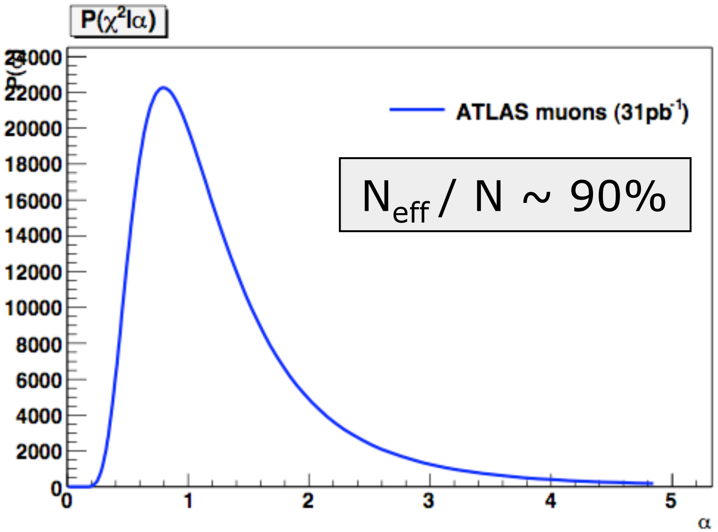
NLO theoretical predictions obtained with DYNNLO
[ArXiv: 0903.2120]

χ^2/d.o.f.	NNPDF2.1	CT10w	MSTW08
ATLAS(31pb^{-1}) muon $p_T > 20$ GeV	0.7	0.8	3.2
CMS(36pb^{-1}) electron $p_T > 25$ GeV	1.9	0.8	2.4
CMS(36pb^{-1}) electron $p_T > 30$ GeV	1.7	1.2	2.5
CMS(36pb^{-1}) muon $p_T > 25$ GeV	1.3	0.5	1.1
CMS(36pb^{-1}) muon $p_T > 30$ GeV	0.8	0.6	1.3

Some examples (preliminary!)

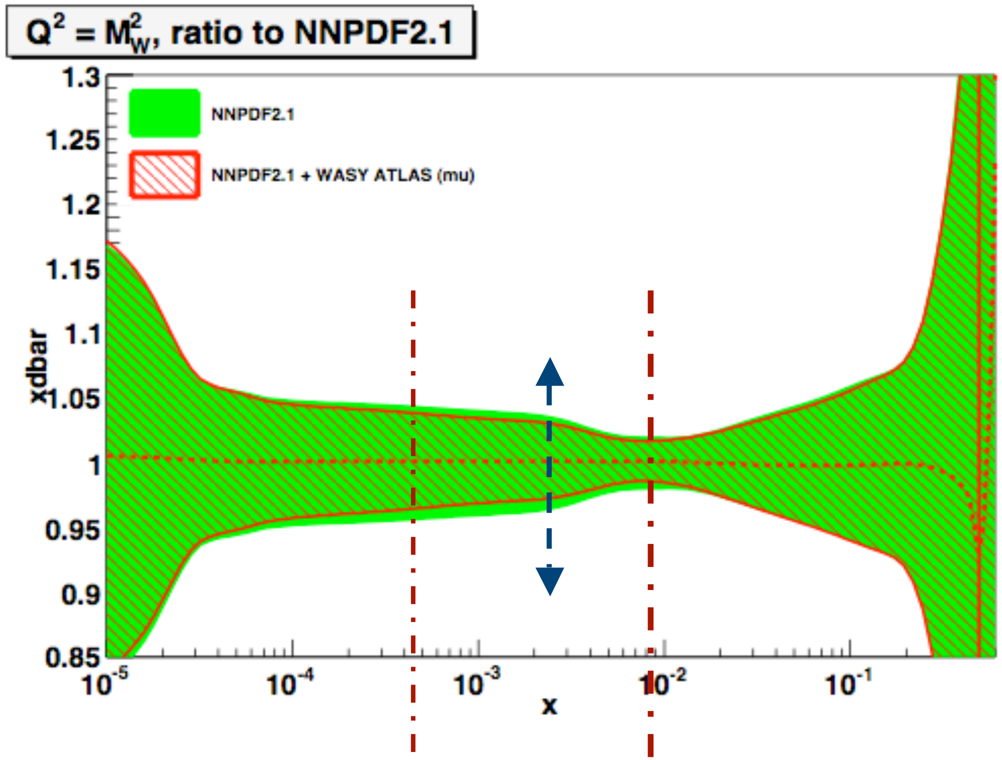
Reweighting the LHC W charge asymmetry





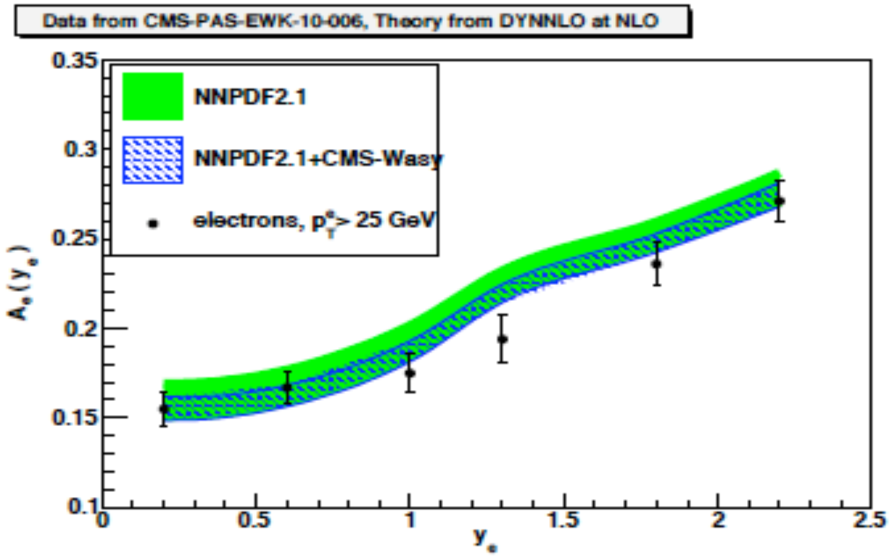
- Data compatible with data included in global PDF analysis
- Slight reduction of uncertainty at medium-small x for light (anti)quark

ATLAS



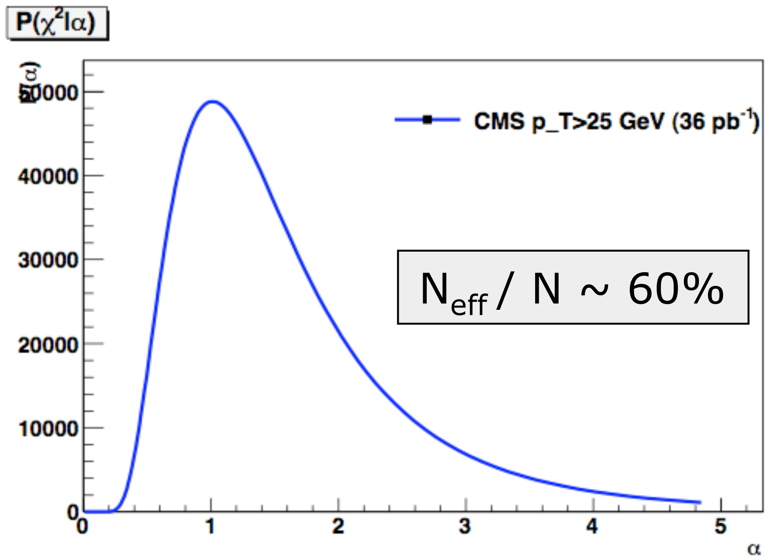
Some examples (preliminary!)

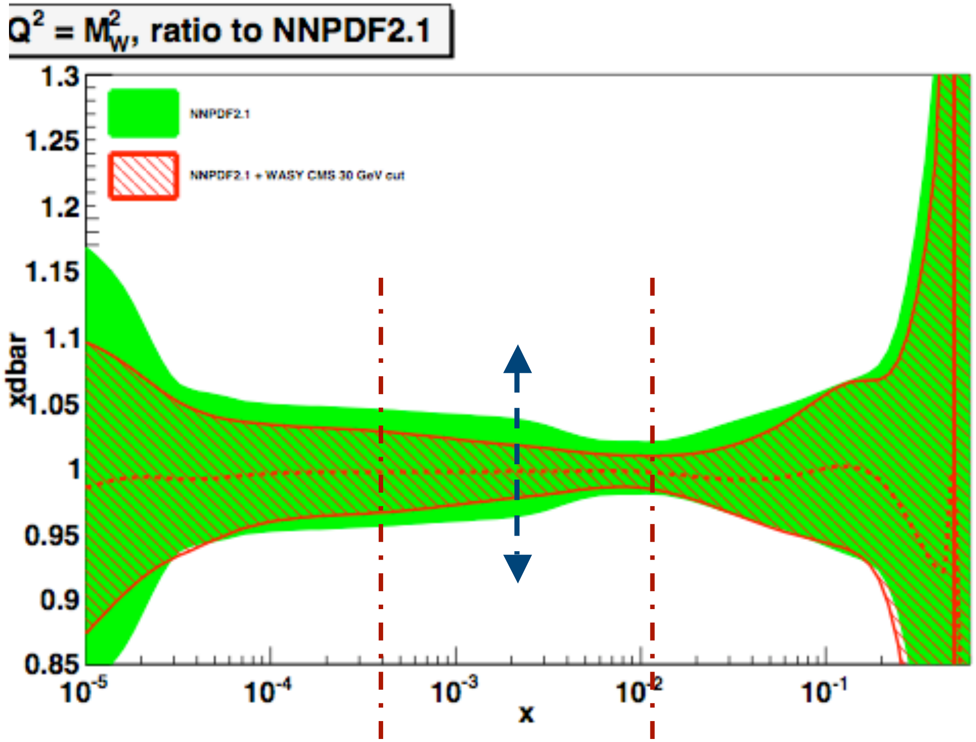
Reweighting the LHC W charge asymmetry



- Data compatible with data included in global PDF analysis
- Significant reduction of uncertainty at medium-small x for light (anti)quark

CMS

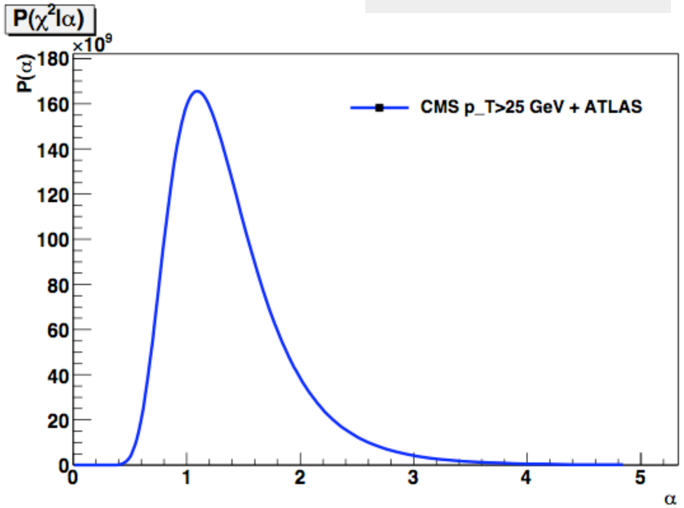


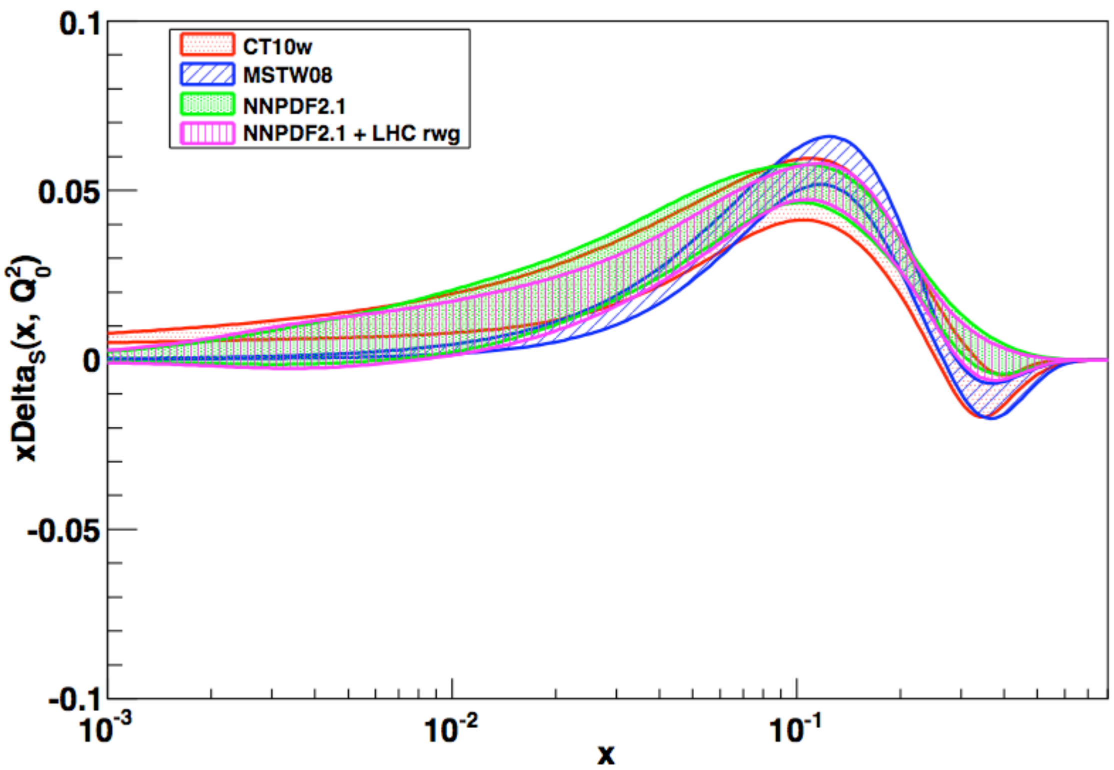


Some examples (preliminary!)
Reweighting the LHC W charge asymmetry

ATLAS+CMS

Data/χ^2	NNPDF2.1	NNPDF2.1 + rwt SET
ATLAS + CMS $p_T^l > 25$ GeV	1.0	0.9
ATLAS + CMS $p_T^l > 30$ GeV	1.2	1.0





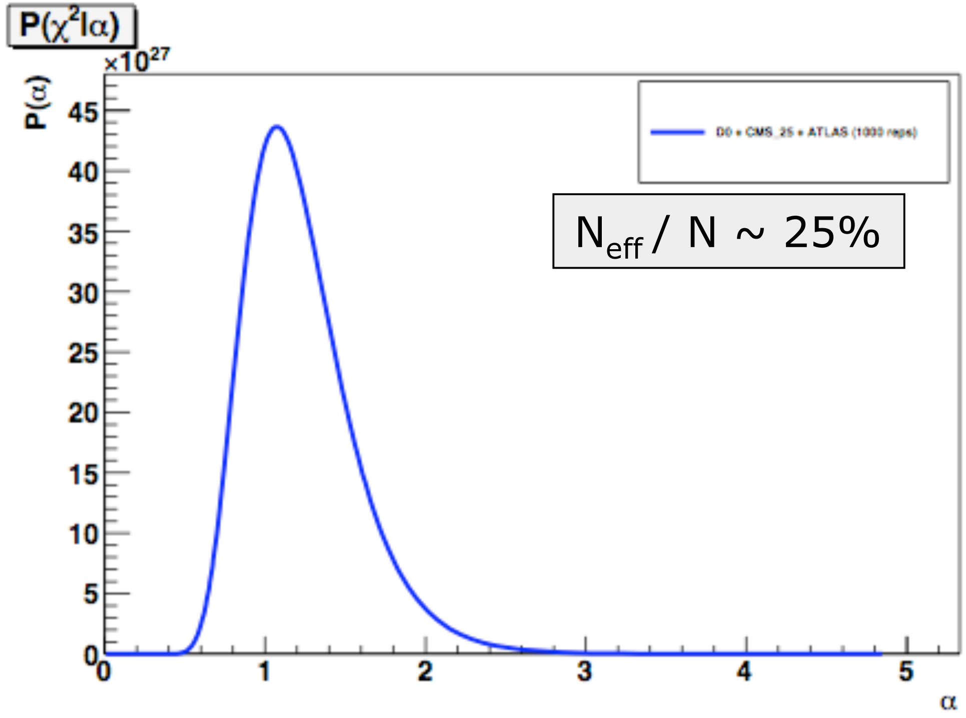
- ATLAS and CMS data can be added at the same time, no clear signs of tension
- Inclusion of data in PDFs fits reduces uncertainty in the small-medium x region for light (anti)quark PDFs

Some examples (preliminary!)

Reweighting TeV + LHC W charge asymmetry

	NNPDF2.1	NNPDF2.1 + TeV + LHC
χ^2	2.0	0.7

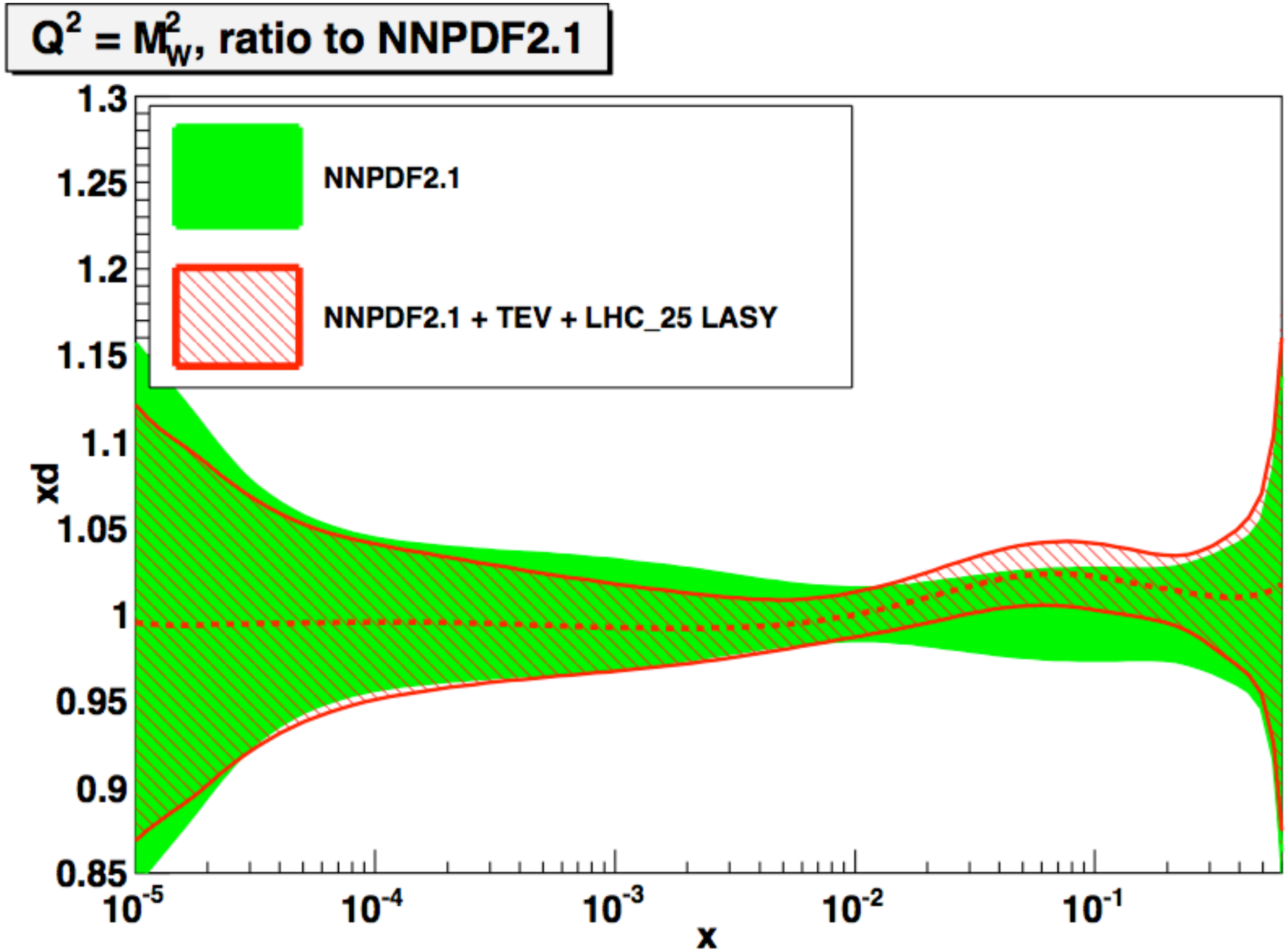
- ✓ YES data are very constraining on PDFs but compatible with data included in the global analyses
- ✓ After reweighting the reweighted set fits very well both Tevatron (D0 muon and D0 electron inclusive) and LHC (ATLAS and CMS) W lepton asymmetry data
- ✓ The description of W asymmetry from CDF does not deteriorate



What about the PDFs central values and PDFs uncertainty?

Some examples (preliminary!)

Reweighting Tev + LHC W charge asymmetry



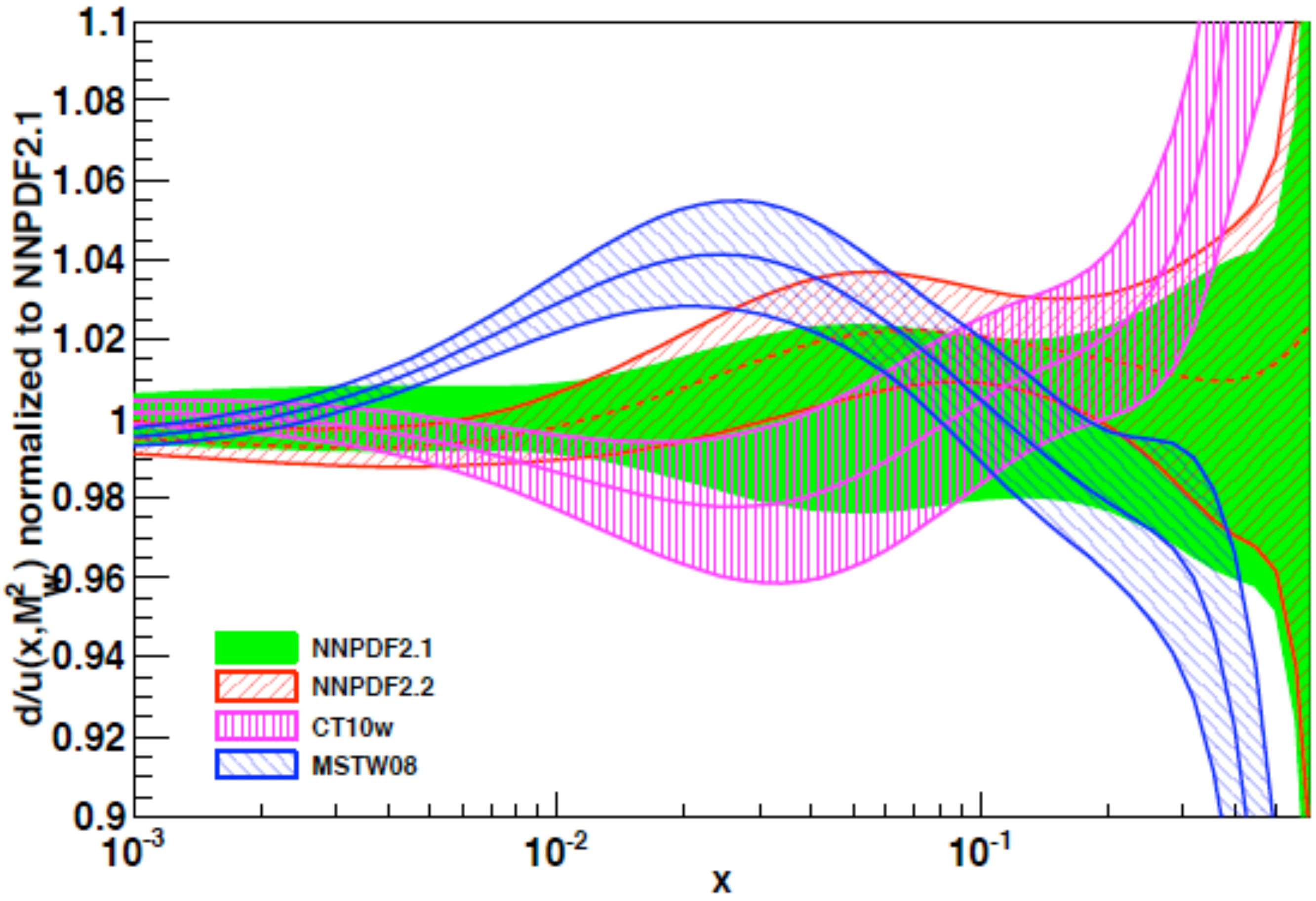
✓ Shift and reduction of uncertainty of light quark and anti-quark PDFs in small-medium x region is driven by LHC data.

✓Reduction of uncertainty in light quark PDFs driven by D0 Tevatron data.

NNPDF2.2 parton set including these data is going to be available soon on LHAPDF

Some examples (preliminary!)

Reweighting Tev + LHC W charge asymmetry

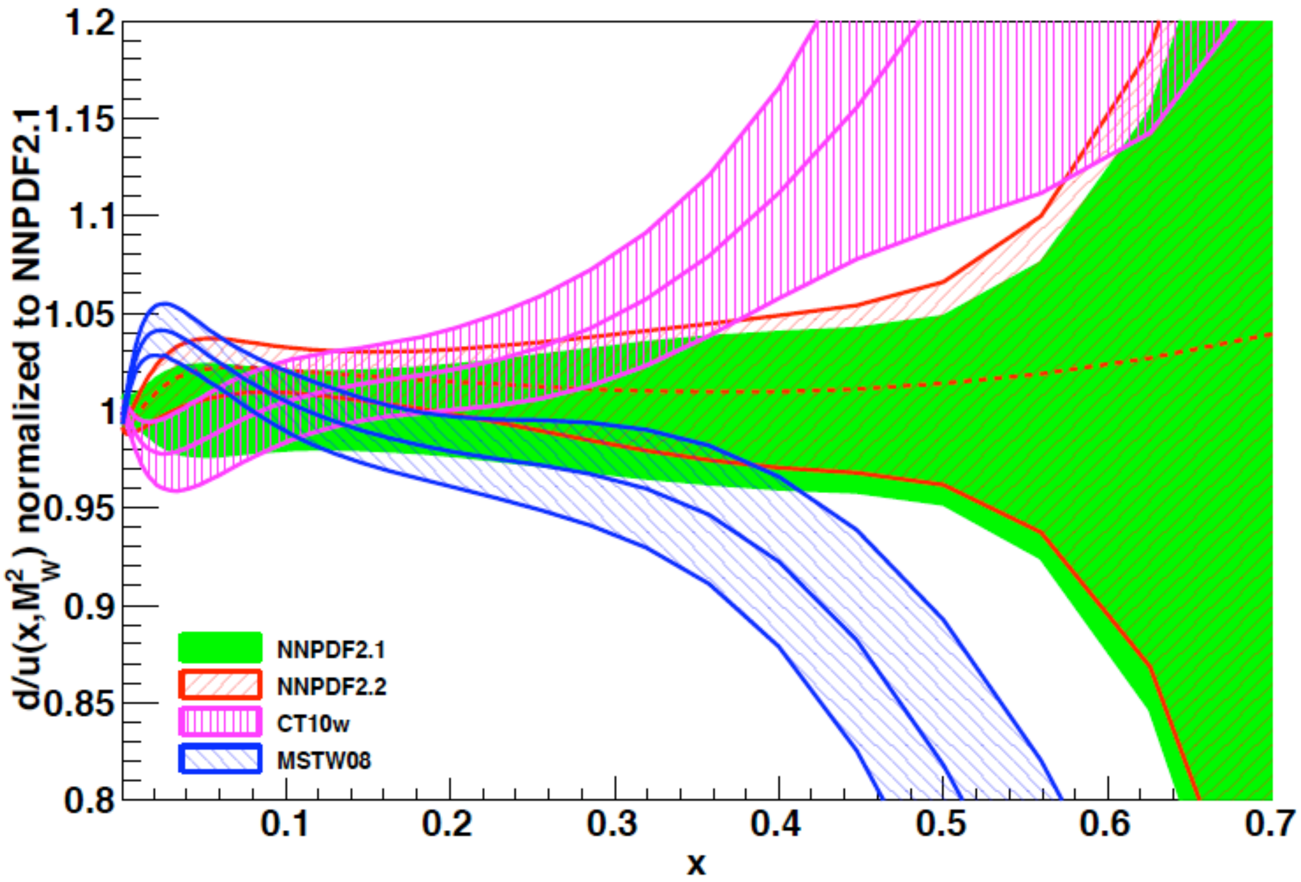


- ✓ Shift and reduction of uncertainty of light quark and anti-quark PDFs in small-medium x region is driven by LHC data.
- ✓ Reduction of uncertainty in light quark PDFs driven by D0 Tevatron data.

NNPDF2.2 parton set including these data is going to be available soon on LHAPDF

LOG scale: see mainly effect of LHC data

Some examples (preliminary!)
Reweighting Tev + LHC W charge asymmetry



LIN scale: see mainly effect of Tevatron data

✓ Shift and reduction of uncertainty of light quark and anti-quark PDFs in small-medium x region is driven by LHC data.

✓ Reduction of uncertainty in light quark PDFs driven by D0 Tevatron data.

NNPDF2.2 parton set including these data is going to be available soon on LHAPDF

Conclusion
and outlook

✓ NNPDF provides reliable and statistically sound global fit to PDFs good for LHC phenomenology

✓ As statistics increases PDF uncertainty can be further reduced and precise LHC data can discriminate among different PDF determinations

✓ Is NNPDF2.1 HERA + Tevatron + LHC (without fixed-target data) the future?

✓ Reweighting PDFs allows fast and efficient inclusion of LHC data as they come out

➔ It can be applied and tested on any Monte Carlo set
➔ HERA also provides Monte Carlo replicas for PDFs

➔ Test reweighting on different Monte Carlo set: take NNPDF2.1_HERAonly and HERAPDF fit with Monte Carlo replicas
➔ Compare results
➔ Test against refitting
➔ Work in progress for W asymmetry data
➔ Much more data to come...

Conclusion
and outlook

✓ NNPDF provides reliable and statistically sound global fit to PDFs good for LHC phenomenology

✓ As statistics increases PDF uncertainty can be further reduced and precise LHC data can discriminate among different PDF determinations

✓ Is NNPDF2.1 HERA + Tevatron + LHC (without fixed-target data) the future?

✓ Reweighting PDFs allows fast and efficient inclusion of LHC data as they come out

➔ It can be applied and tested on any Monte Carlo set
➔ HERA also provides Monte Carlo replicas for PDFs

➔ Test reweighting on different Monte Carlo set: take NNPDF2.1_HERAonly and HERAPDF fit with Monte Carlo replicas
➔ Compare results
➔ Test against refitting
➔ Work in progress for W asymmetry data
➔ Much more data to come... stay tuned!

**THANKS FOR
YOUR ATTENTION and
HOSPITALITY!!**