PDF benchmarking
…on the road to PDF4LHC21

J. Huston

Michigan State University

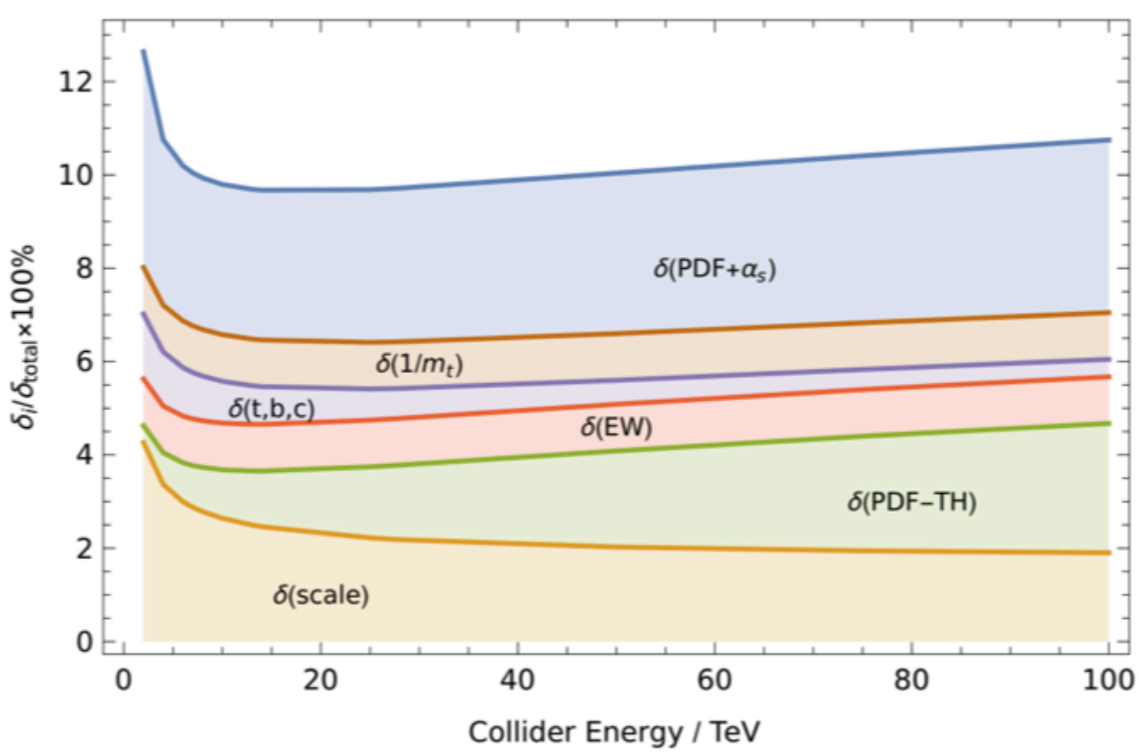
for PDF4LHC benchmarking group

For more details, see the talk of Tom Cridge at DIS21.
Thanks to Xiaoian Jing for new figures.

Precision physics

- Precision physics at the LHC requires precise determinations of PDFs and of $\alpha_s(m_Z)$ PDG 2019

Gluon-gluon fusion into Higgs



$$\alpha_s(M_Z^2) = 0.1179 \pm 0.0010\,. \qquad (9.25)$$

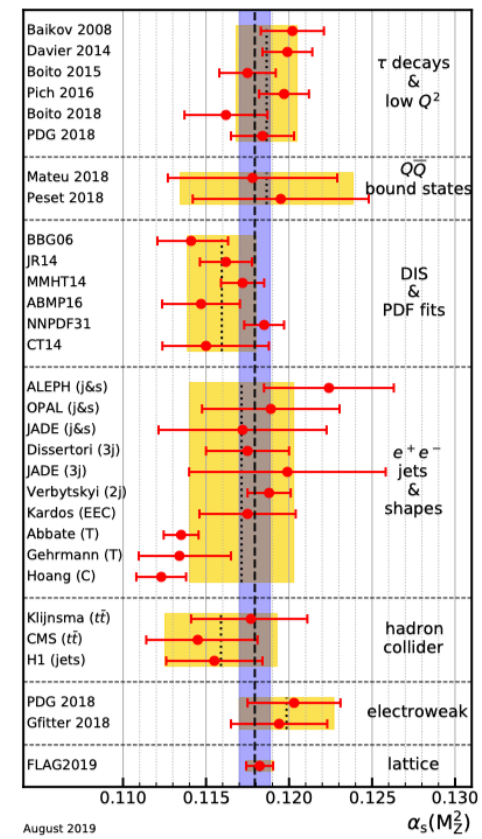


Figure 9.2: Summary of determinations of $\alpha_s(M_Z^2)$ from the seven sub-fields discussed in the text. The yellow (light shaded) bands and dotted lines indicate the pre-average values of each sub-field. The dashed line and blue (dark shaded) band represent the final world average value of $\alpha_s(M_Z^2)$.

PDFs

- Determined from global fits to data from a wide variety of processes, both from fixed target and collider experiments, with an increasing contribution from the LHC itself
- The 3 groups are CTEQ-TEA (CT), MSHT (new acronym) and NNPDF
- Each uses on order of 4000 data points to determine the best fit PDFs and their uncertainties
  - with CT and MSHT using a Hessian formalism and NNPDF using a neural net formalism
- Each group provides regularly updated sets of PDFs

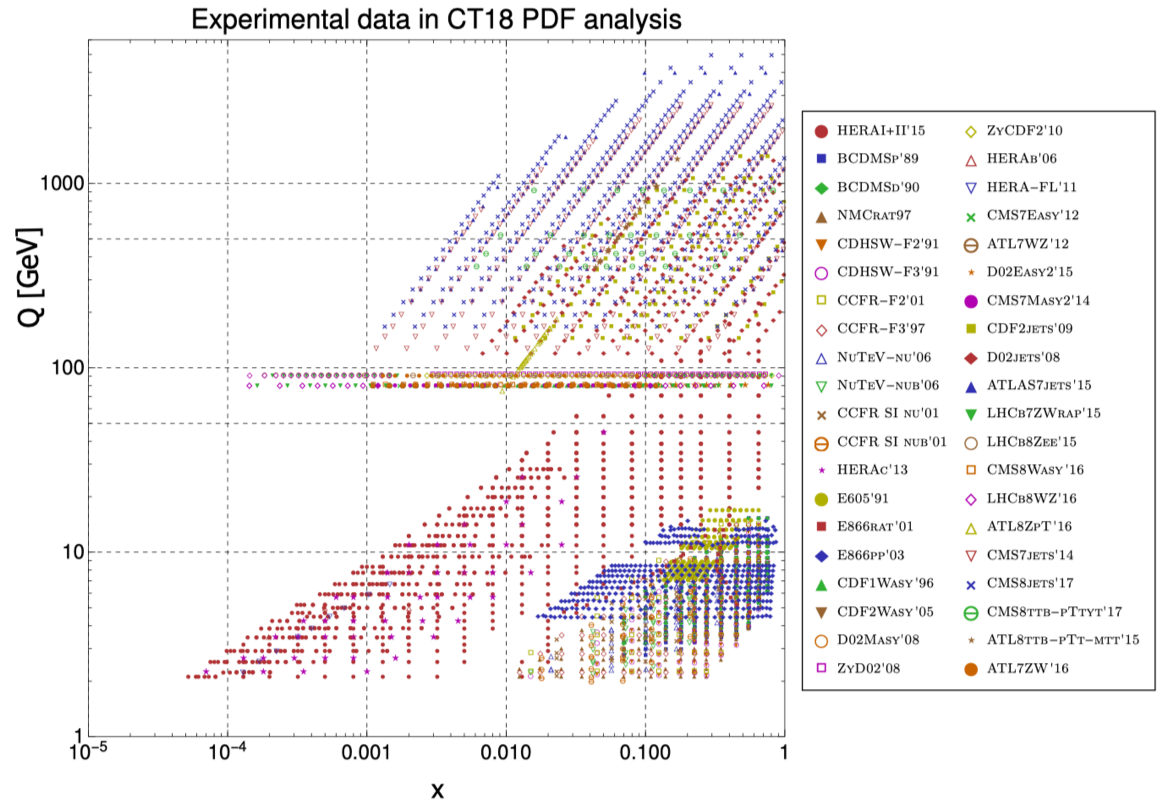


FIG. 1: The CT18 data set, represented in a space of partonic (x, Q), based on Born-level kinematical matchings, $(x, Q) = (x_B, Q)$, in DIS, *etc.*. The matching conventions used here are described in Ref. [20]. Also shown are the ATLAS 7 TeV W/Z production data (ID=248), labeled ATL7WZ'12, fitted in CT18Z.

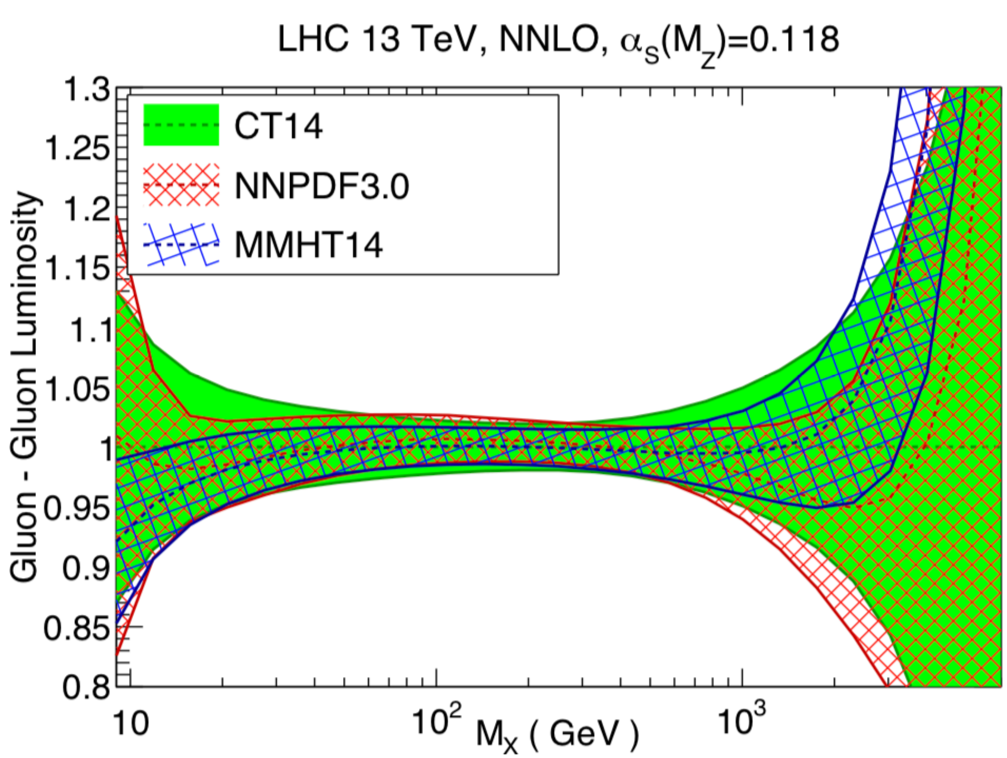
→ to better understand similarities and differences, it is useful to periodically perform benchmarking exercises

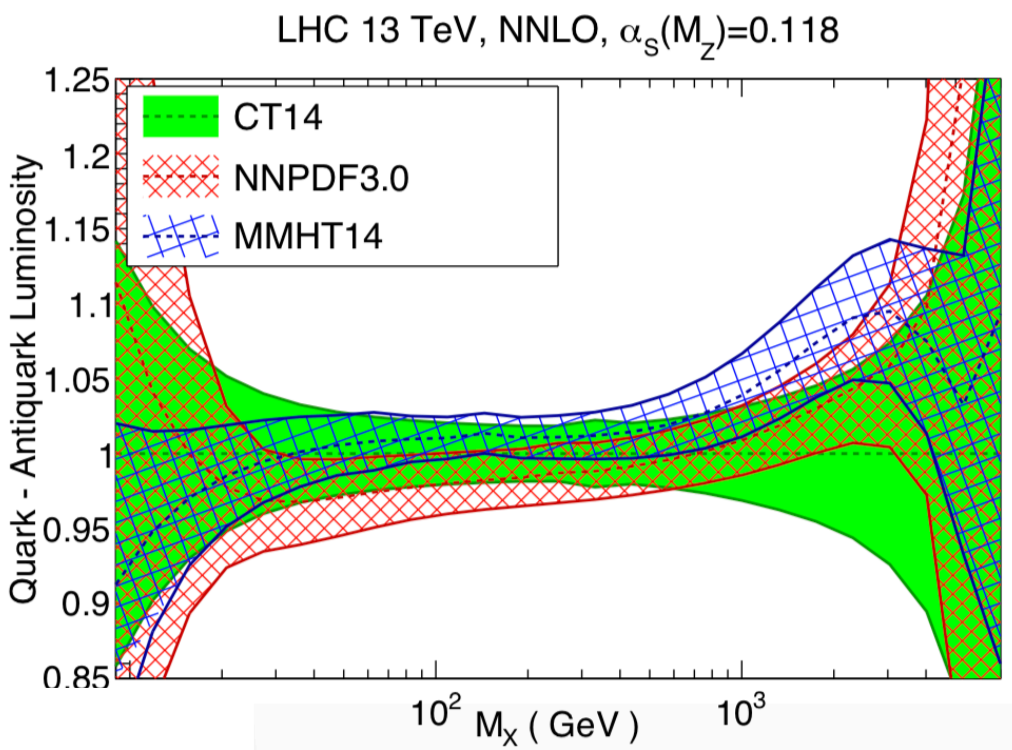
…as for example, PDF4LHC15

- combination of CT14, MMHT2014, NNPDF3.0

…over 1200 citations

  - 1 year benchmarking exercise comparison of above PDFs
  - comparing theory and treatment of experimental data from each group
  - 300 Monte Carlo replicas generated for each of the above PDFs
  - condensed to Hessian sets with from 30-100 members for distribution to users with central PDFs and error PDFs representing the three published PDFs
  - good (too good?) agreement for gluon-gluon luminosity



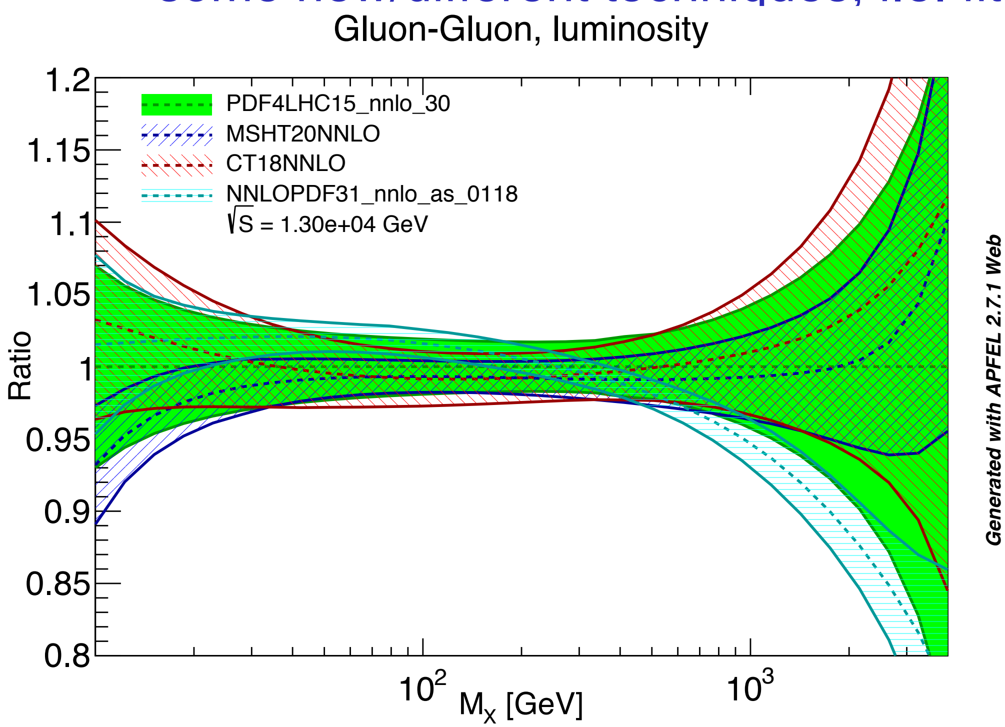


…in the meantime

- New critical data sets from the LHC on Drell-Yan, top, jets, W/Z+jets
- NNLO predictions available for all of above allowing this data to be included in PDF fits
  - transferring NNLO information to global PDF fits still a bit of an issue, i.e. precision of K-factors (statistical jitter->need to smooth and/or use statistical error), availability of grids
- New NNLO PDFs available (CT18, MSHT20, NNPDF3.1) that make use of this LHC data (NNPDF4.0 not yet publically available)
  - additional technical improvements to the PDF fits
- These PDF sets will be used for the construction of PDF4LHC21

PDF4LHC21

- new PDFs CT18, MSHT2020, NNPDF3.1, containing large amount of LHC data
- some new/different techniques, i.e. fitted charm* for NNPDF3.1



consistency with PDF4LHC15, a bit more of a spread of the gg uncertainty bands than for the 2015 combination; some of gg fusion Higgs uncertainty will be due to spread of central values

*charm is fit as a free PDF rather then being generated through evolution

- exercise: start with a reduced data set large enough to provide constraints, small enough that resulting PDFs should be similar for the different groups
  - add more data sets, ttbar, jets … leading to something close to full data sets
- end result in ~few months: central PDFs and Hessian error sets representing the 3 published PDFs->30-50 error PDFs should be sufficient
- paper will appear on archive (PDF4LHC15 paper has 1200 citations)

Aside: uncertainties

- PDF uncertainties depend first on the experimental uncertainties of the data (the path to 1% precision goes through the data)
- Data from two measurements, or even from within the same measurement, can both be very precise, but the result of adding both to the PDF fit can be an increase in the PDF uncertainty (or more likely) a smaller decrease in uncertainty than expected) if the data are in tension with each other
- The resultant PDF uncertainty relies on the definition of a tolerance, i.e. (in the Hessian fit perspective) what is a significant increase from the global minimum χ^2, i.e. PDF uncertainty can be adjusted by changing the tolerance
- $\Delta\chi^2=1$ is not applicable for ~4000 data points from different experiments
- NB: CT (Tier 2) and MSHT (dynamic tolerance) have introduced criteria to restrict the pull of data sets that disagree with global fit; can lead to non-Gaussian behavior

Datasets: CT18 example

TABLE I. Datasets included in the CT18(Z) NNLO global analyses. Here we directly compare the quality of fit found for CT18 NNLO vs CT18Z NNLO on the basis of χ^2_E, $\chi^2_E/N_{pt,E}$, and S_E, in which $N_{pt,E}$, χ^2_E are the number of points and value of χ^2 for experiment E at the global minimum. S_E is the effective Gaussian parameter [38,42,56] quantifying agreement with each experiment. The ATLAS 7 TeV 35 pb^{-1} W/Z dataset, marked by ‡‡, is replaced by the updated one (4.6 fb^{-1}) in the CT18A and CT18Z fits. The CDHSW data, labeled by †, are not included in the CT18Z fit. The numbers in parentheses are for the CT18Z NNLO fit.

Exp. ID#	Experimental dataset		$N_{pt,E}$	χ^2_E	$\chi^2_E/N_{pt,E}$	S_E
160	HERAI + II 1 fb^{-1}, H1 and ZEUS NC and CC $e^{\pm}p$ reduced cross sec. comb.	[30]	1120	1408 (1378)	1.3 (1.2)	5.7 (5.1)
101	BCDMS F_2^p	[57]	337	374 (384)	1.1 (1.1)	1.4 (1.8)
102	BCDMS F_2^d	[58]	250	280 (287)	1.1 (1.1)	1.3 (1.6)
104	NMC F_2^d/F_2^p	[59]	123	126 (116)	1.0 (0.9)	0.2 (−0.4)
108†	CDHSW F_2^p	[60]	85	85.6 (86.8)	1.0 (1.0)	0.1 (0.2)
109†	CDHSW $x_B F_3^p$	[60]	96	86.5 (85.6)	0.9 (0.9)	−0.7 (−0.7)
110	CCFR F_2^p	[61]	69	78.8 (76.0)	1.1 (1.1)	0.9 (0.6)
111	CCFR $x_B F_3^p$	[62]	86	33.8 (31.4)	0.4 (0.4)	−5.2 (−5.6)
124	NuTeV $\nu\mu\mu$ SIDIS	[63]	38	18.5 (30.3)	0.5 (0.8)	−2.7 (−0.9)
125	NuTeV $\bar{\nu}\mu\mu$ SIDIS	[63]	33	38.5 (56.7)	1.2 (1.7)	0.7 (2.5)
126	CCFR $\nu\mu\mu$ SIDIS	[64]	40	29.9 (35.0)	0.7 (0.9)	−1.1 (−0.5)
127	CCFR $\bar{\nu}\mu\mu$ SIDIS	[64]	38	19.8 (18.7)	0.5 (0.5)	−2.5 (−2.7)
145	H1 σ_r^b	[65]	10	6.8 (7.0)	0.7 (0.7)	−0.6 (−0.6)
147	Combined HERA charm production	[66]	47	58.3 (56.4)	1.2 (1.2)	1.1 (1.0)
169	H1 F_L	[33]	9	17.0 (15.4)	1.9 (1.7)	1.7 (1.4)
201	E605 Drell-Yan process	[67]	119	103.4 (102.4)	0.9 (0.9)	−1.0 (−1.1)
203	E866 Drell-Yan process $\sigma_{pd}/(2\sigma_{pp})$	[68]	15	16.1 (17.9)	1.1 (1.2)	0.3 (0.6)
204	E866 Drell-Yan process $Q^3 d^2\sigma_{pp}/(dQ dx_F)$	[69]	184	244 (240)	1.3 (1.3)	2.9 (2.7)
225	CDF run-1 lepton A_{ch}, $p_{T\ell} > 25$ GeV	[70]	11	9.0 (9.3)	0.8 (0.8)	−0.3 (−0.2)
227	CDF run-2 electron A_{ch}, $p_{T\ell} > 25$ GeV	[71]	11	13.5 (13.4)	1.2 (1.2)	0.6 (0.6)
234	DØ run-2 muon A_{ch}, $p_{T\ell} > 20$ GeV	[72]	9	9.1 (9.0)	1.0 (1.0)	0.2 (0.1)
260	DØ run-2 Z rapidity	[73]	28	16.9 (18.7)	0.6 (0.7)	−1.7 (−1.3)
261	CDF run-2 Z rapidity	[74]	29	48.7 (61.1)	1.7 (2.1)	2.2 (3.3)
266	CMS 7 TeV 4.7 fb^{-1}, muon A_{ch}, $p_{T\ell} > 35$ GeV	[75]	11	7.9 (12.2)	0.7 (1.1)	−0.6 (0.4)
267	CMS 7 TeV 840 pb^{-1}, electron A_{ch}, $p_{T\ell} > 35$ GeV	[76]	11	4.6 (5.5)	0.4 (0.5)	−1.6 (−1.3)
268‡‡	ATLAS 7 TeV 35 pb^{-1} W/Z cross sec., A_{ch}	[77]	41	44.4 (50.6)	1.1 (1.2)	0.4 (1.1)
281	DØ run-2 9.7 fb^{-1} electron A_{ch}, $p_{T\ell} > 25$ GeV	[78]	13	22.8 (20.5)	1.8 (1.6)	1.7 (1.4)
504	CDF run-2 inclusive jet production	[79]	72	122 (117)	1.7 (1.6)	3.5 (3.2)
514	DØ run-2 inclusive jet production	[80]	110	113.8 (115.2)	1.0 (1.0)	0.3 (0.4)

Spartyness, a statistical variable that indicates the degree of tension of a given data set with the resultant PDF

Spartyness should have a normal distribution with a mean of zero and a standard deviation of 1

LHC data

TABLE II. Like Table I, for newly included LHC measurements. The ATLAS 7 TeV W/Z data (4.6 fb^{-1}), labeled by ‡, are included in the CT18A and CT18Z global fits, but not in CT18 and CT18X.

Exp. ID#	Experimental dataset		$N_{pt,E}$	χ^2_E	$\chi^2_E/N_{pt,E}$	S_E
245	LHCb 7 TeV 1.0 fb^{-1} W/Z forward rapidity cross sec.	[81]	33	53.8 (39.9)	1.6 (1.2)	2.2 (0.9)
246	LHCb 8 TeV 2.0 fb^{-1} $Z \to e^- e^+$ forward rapidity cross sec.	[82]	17	17.7 (18.0)	1.0 (1.1)	0.2 (0.3)
248‡	ATLAS 7 TeV 4.6 fb^{-1}, W/Z combined cross sec.	[39]	34	287.3 (88.7)	8.4 (2.6)	13.7 (4.8)
249	CMS 8 TeV 18.8 fb^{-1} muon charge asymmetry A_{ch}	[83]	11	11.4 (12.1)	1.0 (1.1)	0.2 (0.4)
250	LHCb 8 TeV 2.0 fb^{-1} W/Z cross sec.	[84]	34	73.7 (59.4)	2.1 (1.7)	3.7 (2.6)
253	ATLAS 8 TeV 20.3 fb^{-1}, Z p_T cross sec.	[85]	27	30.2 (28.3)	1.1 (1.0)	0.5 (0.3)
542	CMS 7 TeV 5 fb^{-1}, single incl. jet cross sec., $R = 0.7$ (extended in y)	[86]	158	194.7 (188.6)	1.2 (1.2)	2.0 (1.7)
544	ATLAS 7 TeV 4.5 fb^{-1}, single incl. jet cross sec., $R = 0.6$	[9]	140	202.7 (203.0)	1.4 (1.5)	3.3 (3.4)
545	CMS 8 TeV 19.7 fb^{-1}, single incl. jet cross sec., $R = 0.7$, (extended in y)	[87]	185	210.3 (207.6)	1.1 (1.1)	1.3 (1.2)
573	CMS 8 TeV 19.7 fb^{-1}, $t\bar{t}$ norm. double-diff. top p_T and y cross sec.	[88]	16	18.9 (19.1)	1.2 (1.2)	0.6 (0.6)
580	ATLAS 8 TeV 20.3 fb^{-1}, $t\bar{t}$ p_T^t and $m_{t\bar{t}}$ abs. spectrum	[89]	15	9.4 (10.7)	0.6 (0.7)	−1.1 (−0.8)

CT18: 3681 data points, χ^2=4255 (similar χ^2/dof values for MSHT20, NNPDF3.1)

Benchmarking

- Origins of differences of PDFs
  - due to variations of experimental input, different theory choices, fitting methodologies?
  - so for benchmarking, use common theory settings (i.e. perturbative charm, m_{charm}=1.4 GeV, s=sbar at input scale, $\alpha_s(m_Z)$=0.118, positive-definite PDFs, no deuteron or nuclear corrections…)
- …and start with reduced set that contains targeted data

- NMC deuteron to proton ratio in DIS.
- NuTeV dimuon cross-sections.
- HERA I+II inclusive cross-sections from DIS.
- E866 fixed target Drell-Yan ratio pd/pp data.
- D0 Z rapidity distribution.
- ATLAS W, Z 7 TeV rapidity distribution, only Z peak and central.
- CMS 7 TeV W asymmetry.
- CMS 8 TeV inclusive jet data.
- LHCb 7, 8 TeV W, Z rapidity distributions.
- BCDMS proton and deuteron DIS data.

-later added ATLAS tt data
-now adding additional LHC jet data

Theory/systematic errors

- First step, ensure that theory predictions are the same for each group, i.e. make predictions for data used in reduced fit using PDF4LHC15 predictions
  - some differences observed, understood
- Verify that treatment of systematic errors is the same/similar

χ^2 values for reduced fits

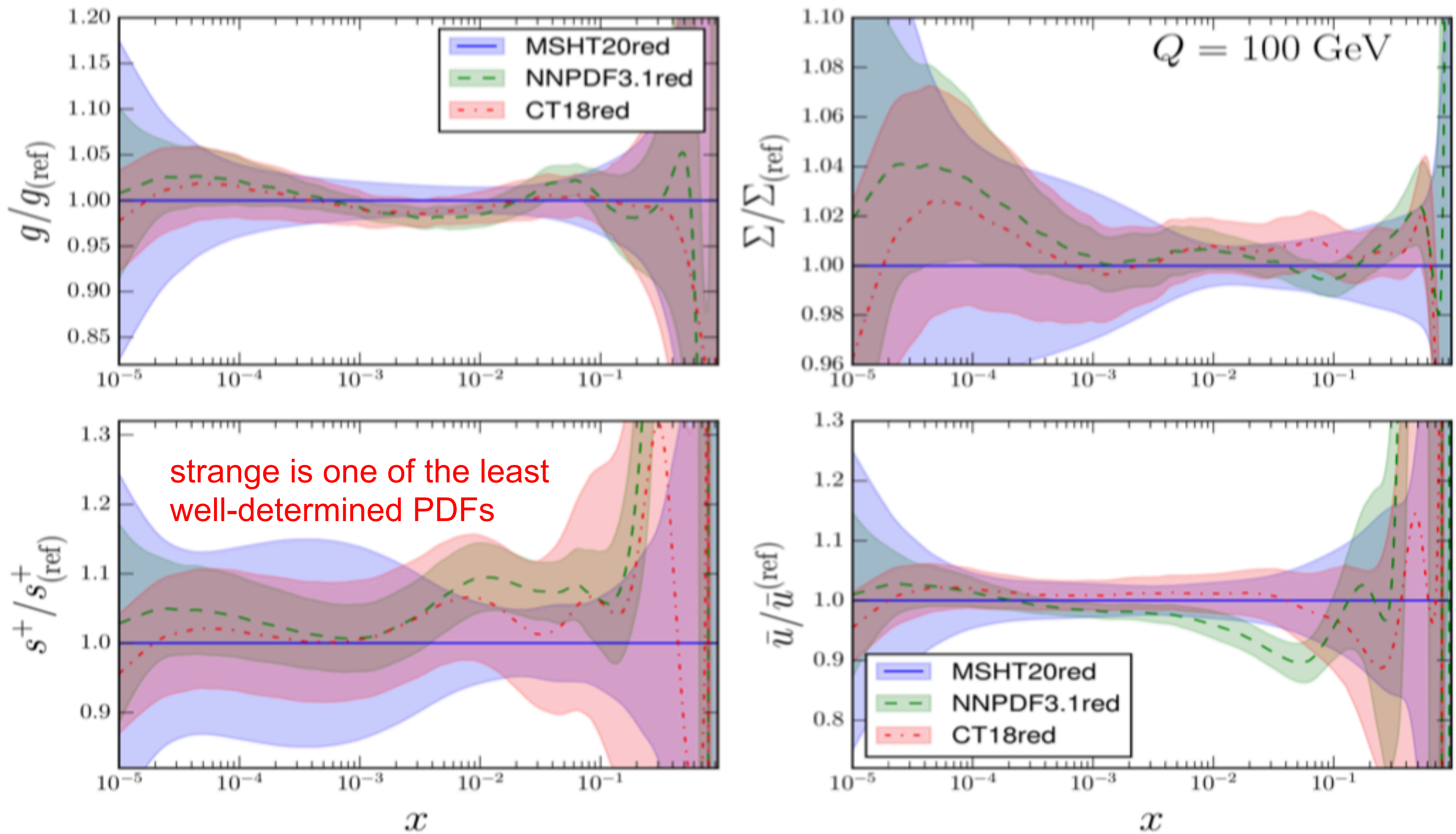
ID	Expt.	N_{pt}	χ^2/N_{pt} (CT)	χ^2/N_{pt} (MSHT)	χ^2/N_{pt} (NNPDF)
101	BCDMS F_2^p	$329/163^{\dagger\dagger}/325^{\dagger}$	1.06	1.00	1.21
102	BCDMS F_2^d	$246/151^{\dagger\dagger}/244^{\dagger}$	1.01	0.88	1.10
104	NMC F_2^d/F_2^p	$118/117^{\dagger}$	0.91	0.93	0.90
124+125	NuTeV $\nu\mu\mu + \bar{\nu}\mu\mu$	$71/73^{\dagger\dagger}$	0.77	0.81	1.22
160	HERAI+II	1120	1.22	1.20	1.22
203	E866 $\sigma_{pd}/(2\sigma_{pp})$	15	1.49	0.79	0.43
245+250	LHCb 7TeV& 8TeV W,Z	29+30	1.16	1.18	1.44
246	LHCb 8TeV $Z \to ee$	17	1.30	1.43	1.57
248	ATLAS 7TeV W,Z(2016)	34	2.13	1.79	2.33
260	D0 Z rapidity	28	0.57	0.58	0.62
267	CMS 7TeV eletron A_{ch}	11	1.43	1.50	0.76
269	ATLAS 7TeV W,Z(2011)	30	1.01	0.93	1.01
545	CMS 8TeV incl. jet	$185/174^{\dagger\dagger}$	1.02	1.39	1.30
Total	N_{pt}	—	2263	1993	2256
Total	χ^2/N_{pt}	—	1.13	1.14	1.20

Table 1: Reduced FITS. χ^2/N_{pt}, **consensus theory**:$s = \bar{s}$, perturbative charm, nuclear-free, ...

††MSHT †NNPDF

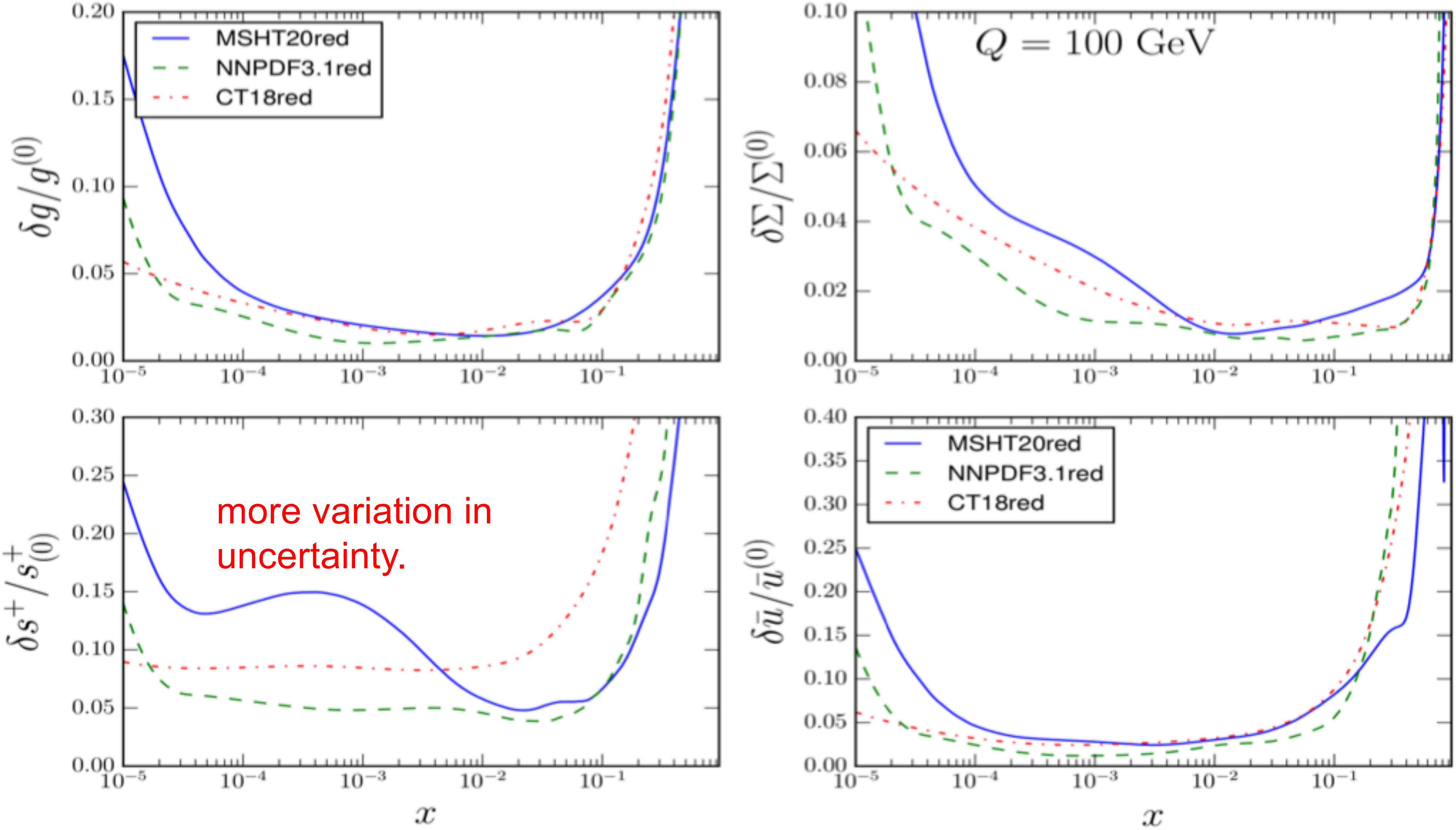
Reduced fits

Reasonable agreement for the most part.

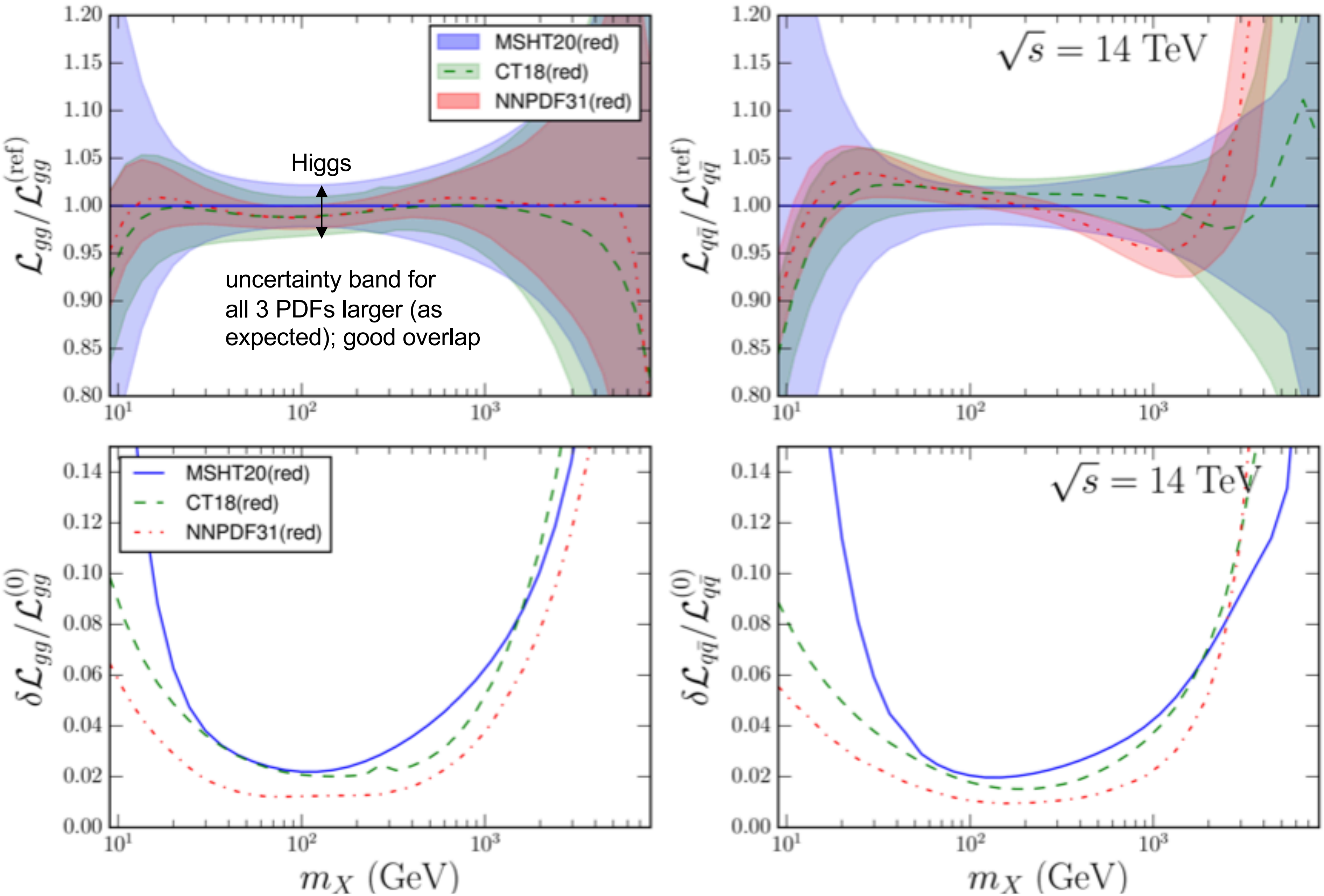


Uncertainties

Uncertainties increased with respect to full global fits, but in agreement with each other.



PDF luminosities



High x gluon

- Of great interest for both SM physics and searches
- 3 main datasets sensitive to high x gluon: jet data, top data, Z p_T data
- Tensions between data sets, tensions within data sets
  - correlated systematics important
- Consider ATLAS 8 TeV top datasets: $m_{tt}, y_t, y_{tt}, p_T^{top}$
  - MSHT, CT, ATLAS cannot get good fit to all correlated distributions together, or to y_t, y_{tt} separately, in either reduced fit or full global fit
  - NNPDF able to fit rapidity distributions if systematics for each observable are de-correlated; for correlated case find same large chisquares as the other PDFs
- Theory predictions check out, i.e. common theory used by all groups
- NB: top data sets have a low number of data points; NNPDF cannot divide into training and validation, so all data in training
  - small data sets are effectively double-weighted (e.g. E866, CMS W charge asymmetry)

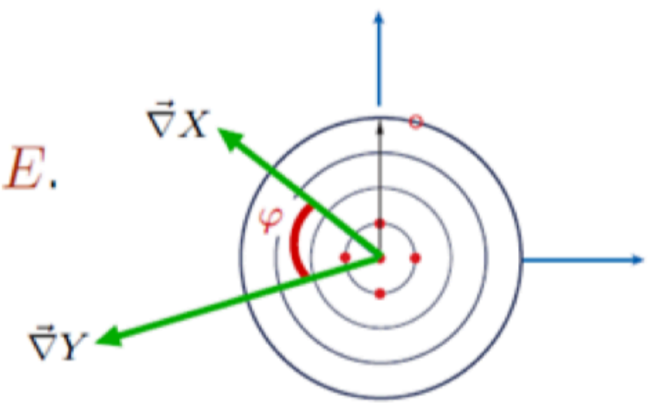
Useful tool

Tim Hobbs, Pavel Nadolsky

L_2 sensitivity, definition

$S_{f,L_2}(E)$ for experiment E is the estimated $\Delta\chi^2_E$ for this experiment when a PDF $f_a(x_i, Q_i)$ increases by the +68% c.l. Hessian PDF uncertainty

Take $X = f_a(x_i, Q_i)$ or $\sigma(f)$; $Y = \chi^2_E$ for experiment E.

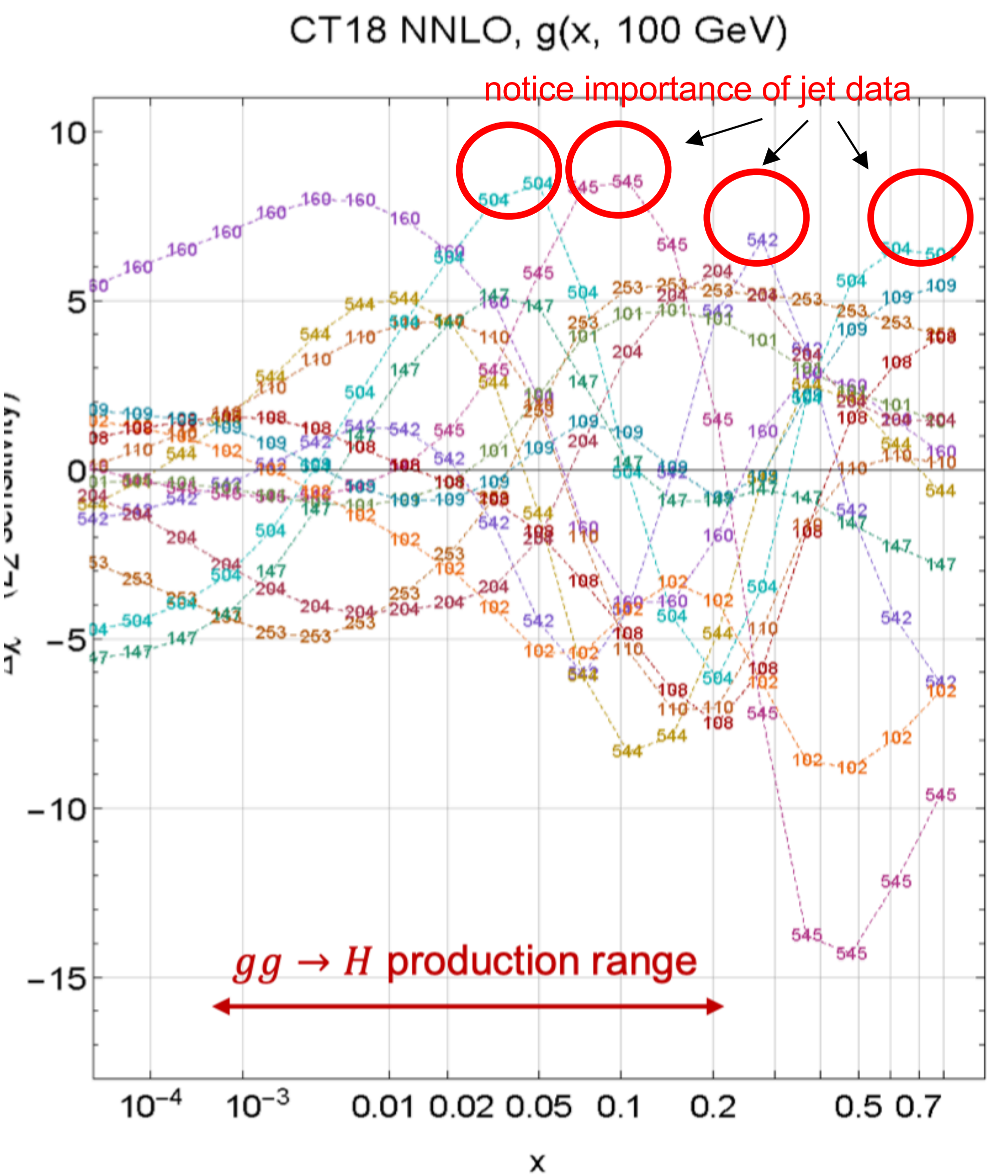


$$S_{f,L_2} \equiv \Delta Y(\vec{z}_{m,X}) = \vec{\nabla}Y \cdot \vec{z}_{m,X} = \vec{\nabla}Y \cdot \frac{\vec{\nabla}X}{|\vec{\nabla}X|} = \Delta Y \cos\varphi.$$

A fast version of the Lagrange Multiplier scan of χ^2_E along the direction of $f_a(x_i, Q_i)$!

Estimated χ^2 pulls from experiments

(L_2 sensitivity, T. Hobbs et al., arXiv:1904.00222, v. 2)



Plots of L_2 sensitivities to various PDFs:
https://ct.hepforge.org/PDFs/ct18/figures/L2Sensitivity/

CT18 NNLO, gluon at Q=100 GeV

15 core-minutes

Most sensitive experiments

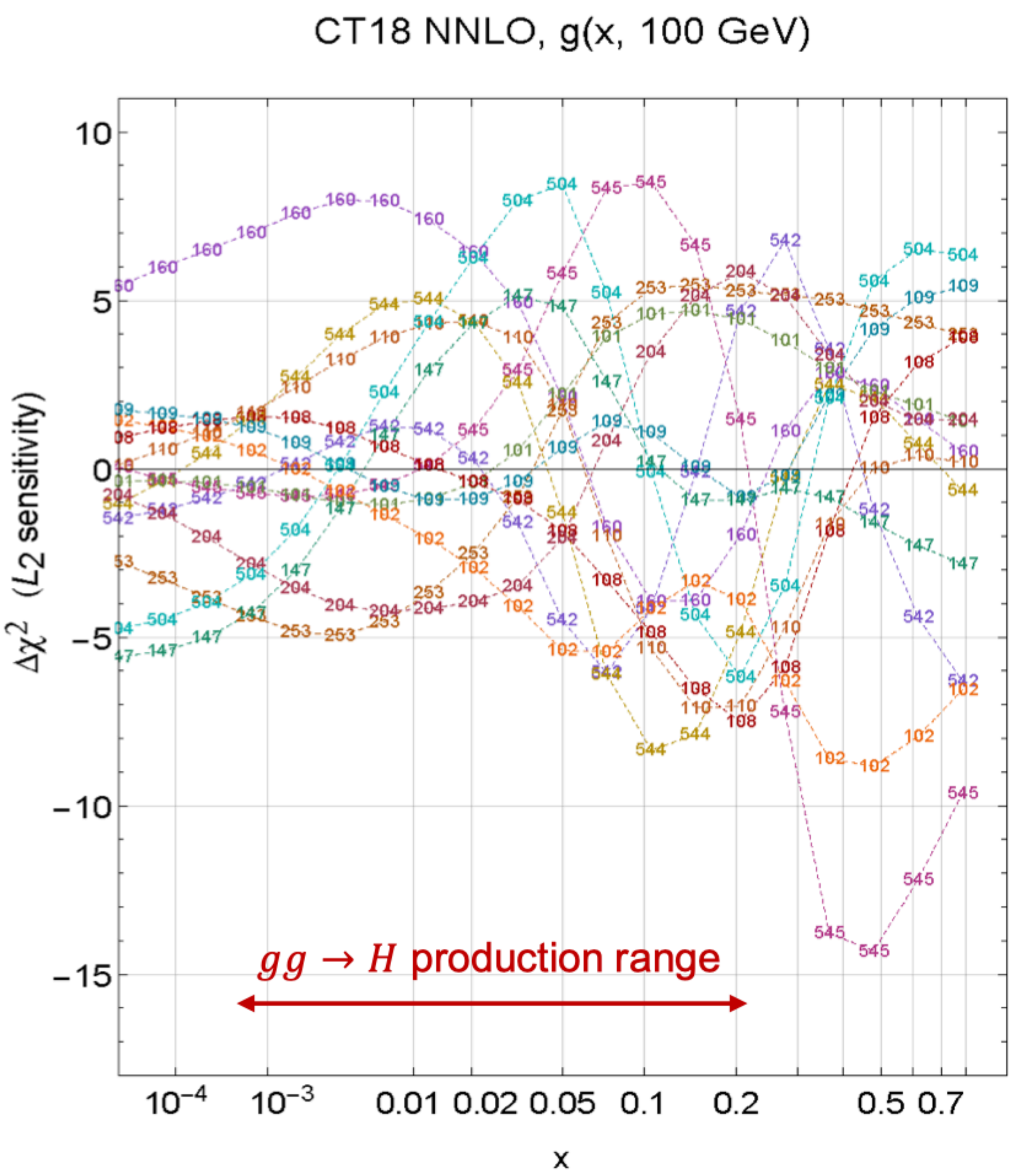
- 253 ATL8ZpTbT
- 542 CMS7jtR7y6T
- 544 ATL7jtR6uT
- 545 CMS8jtR7T
- 160 HERAIpII
- 101 BcdF2pCor
- 102 BcdF2dCor
- 108 cdhswf2
- 109 cdhswf3
- 110 ccfrf2.mi
- 147 Hn1X0c
- 204 e866ppxf
- 504 cdf2jtCor2

Experiments with large $\Delta\chi^2 > 0$ $[\Delta\chi^2 < 0]$ pull $g(x, Q)$ in the negative [positive] direction at the shown x

ideally, the sum at any x value should be zero; non-zero indicates non-Gaussianity

Estimated χ^2 pulls from experiments

(L_2 sensitivity, T. Hobbs et al., arXiv:1904.00222, v. 2)



Plots of L_2 sensitivities to various PDFs:
https://ct.hepforge.org/PDFs/ct18/figures/L2Sensitivity/

CT18 NNLO, gluon at Q=100 GeV

15 core-minutes

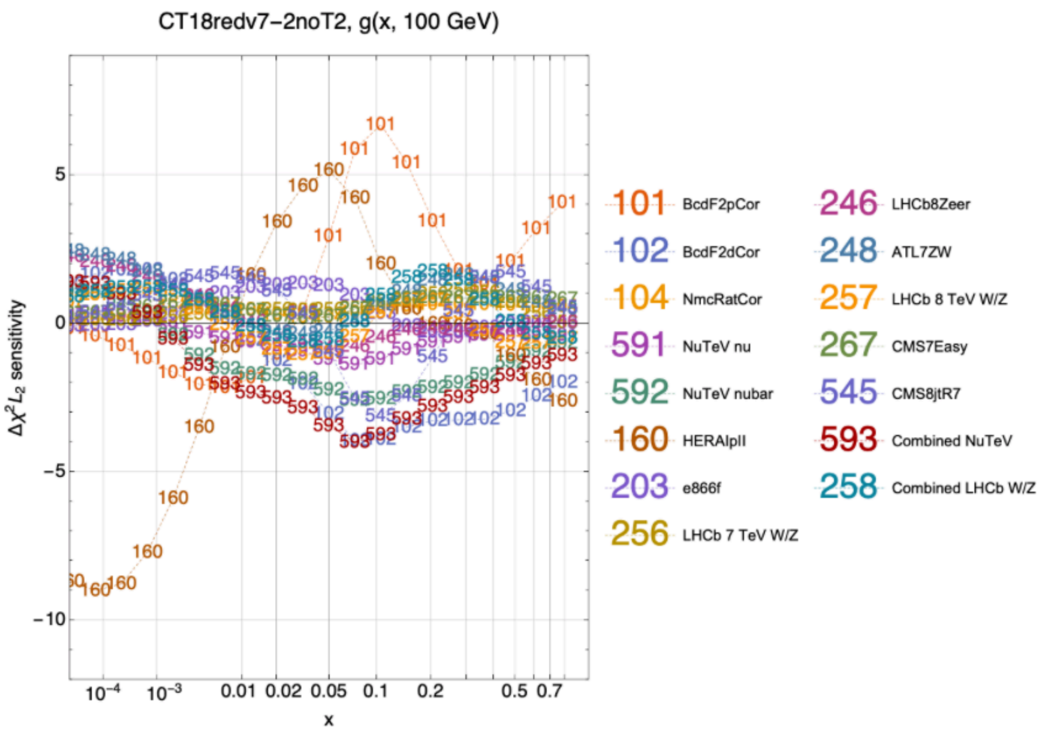
Most sensitive experiments

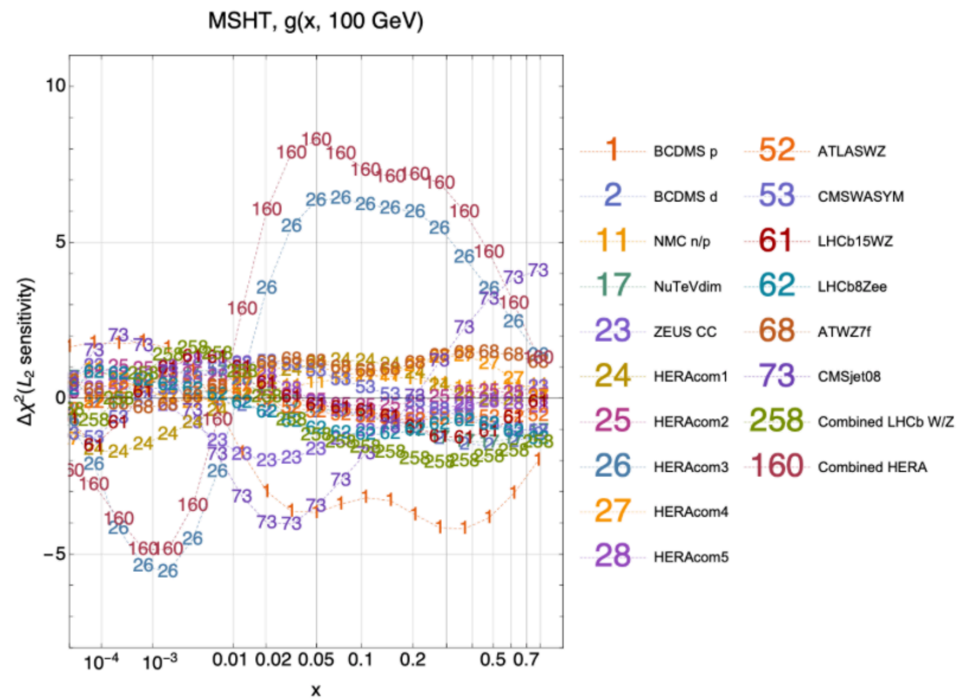
- 253 ATL8ZpTbT
- 542 CMS7jtR7y6T
- 544 ATL7jtR6uT
- 545 CMS8jtR7T
- 160 HERAIpII
- 101 BcdF2pCor
- 102 BcdF2dCor
- 108 cdhswf2
- 109 cdhswf3
- 110 ccfrf2.mi
- 147 Hn1X0c
- 204 e866ppxf
- 504 cdf2jtCor2

Note opposite pulls (tensions) in some x ranges between HERA I+II DIS (ID=160); CDF (504), ATLAS 7 (544), CMS 7 (542), CMS 8 jet (545) production; E866pp DY (204); ATLAS 8 Z pT (253) production; BCDMS and CDHSW DIS

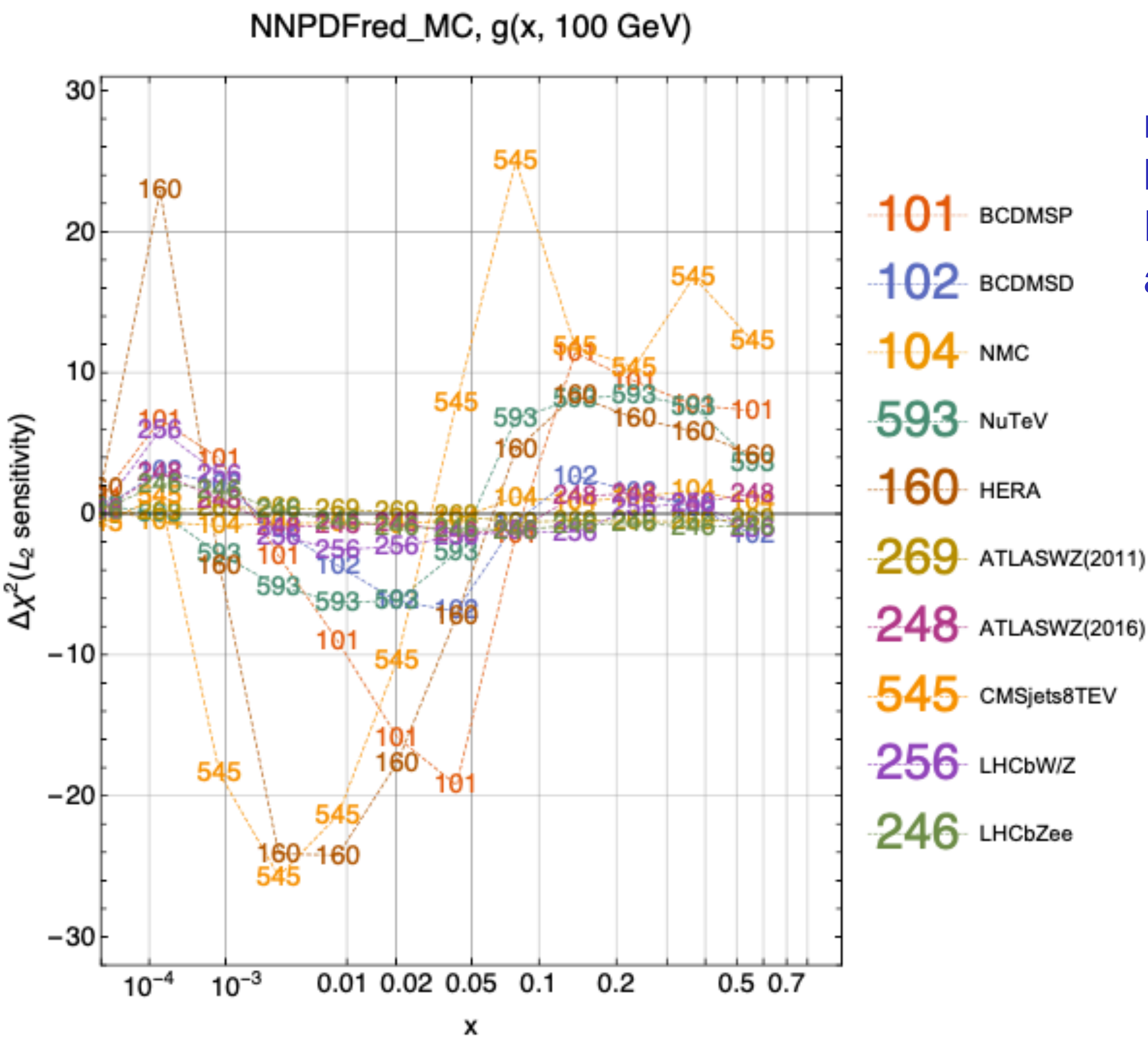
L2 sensitivity for CT18red and MSHTred

use tolerance^2=10, no Tier 2 penalty for CT18red
… and remember that the gluon distributions for all reduced PDFs are similar





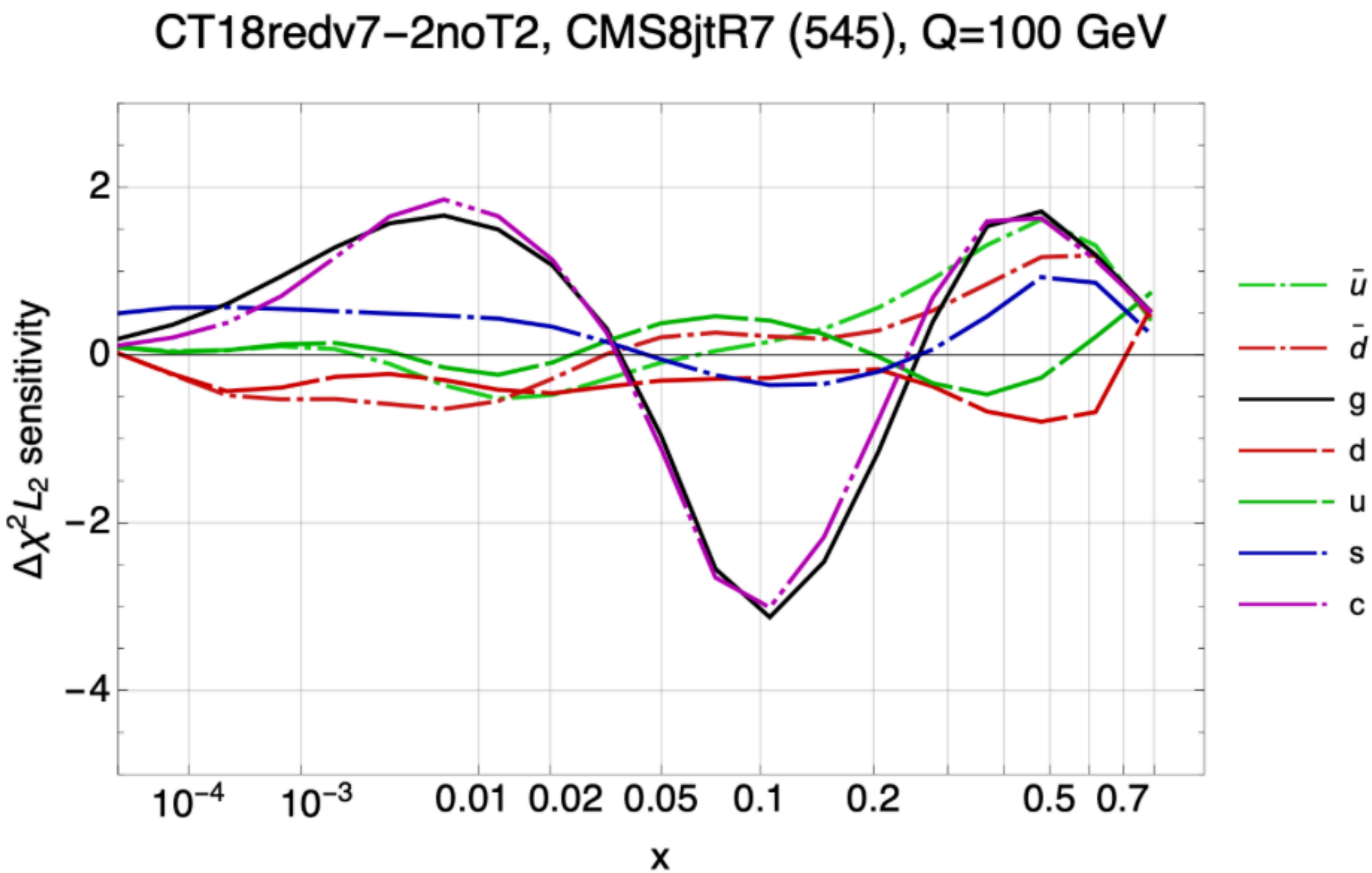
L2 sensitivity for NNPDF3.1red for the gluon distribution



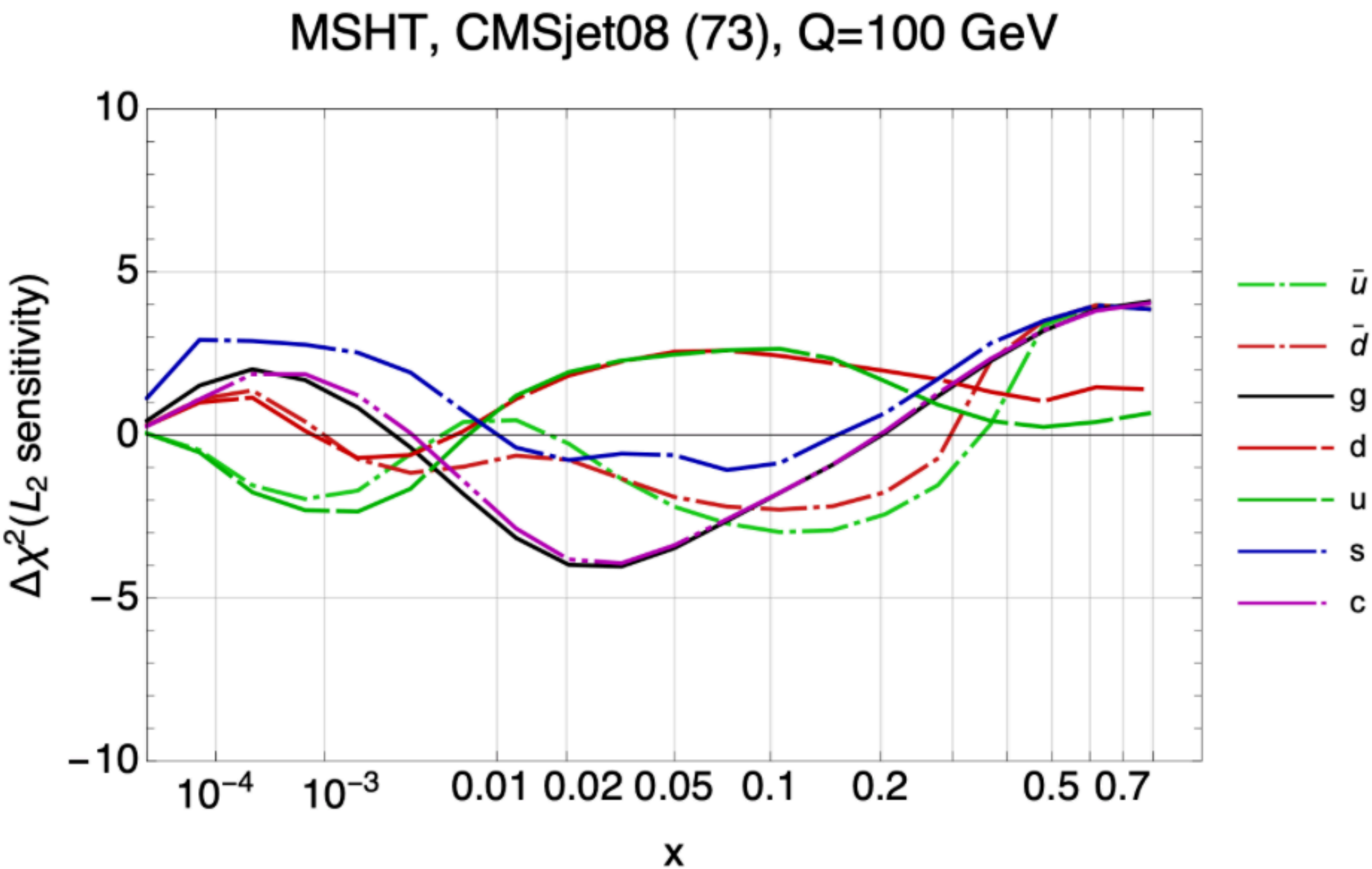
…this is using the MC replicas; they have also been converted into a Hessian error set; results are equivalent

Can also look at sensitivities for individual experiments

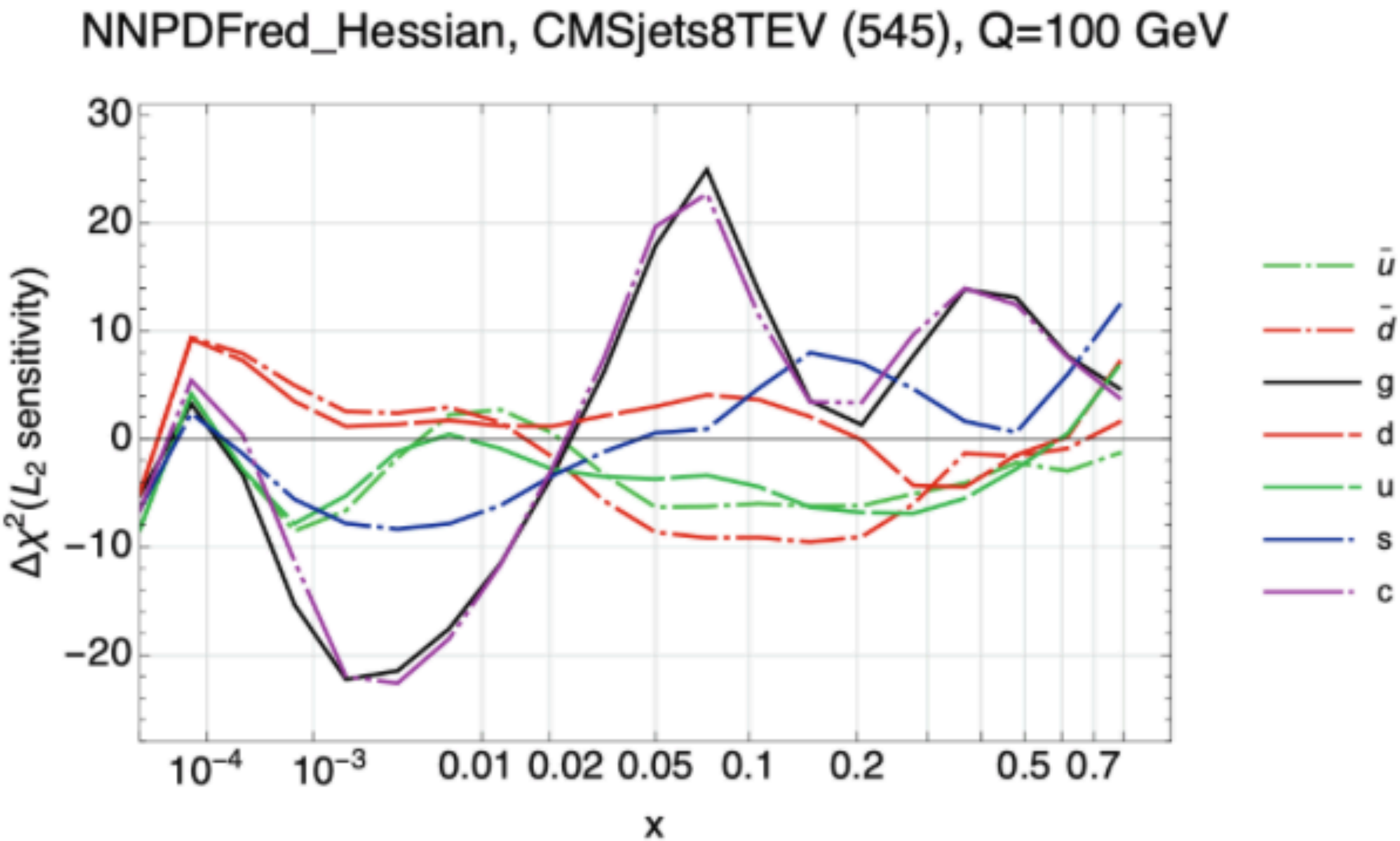
…expect L2 sensitivities for each PDF flavor to average to about zero (Gaussian behavior); asymmetric errors from dynamic (Tier 2) tolerance can change that



Can also look at sensitivities for individual experiments



Can also look at sensitivities for individual experiments

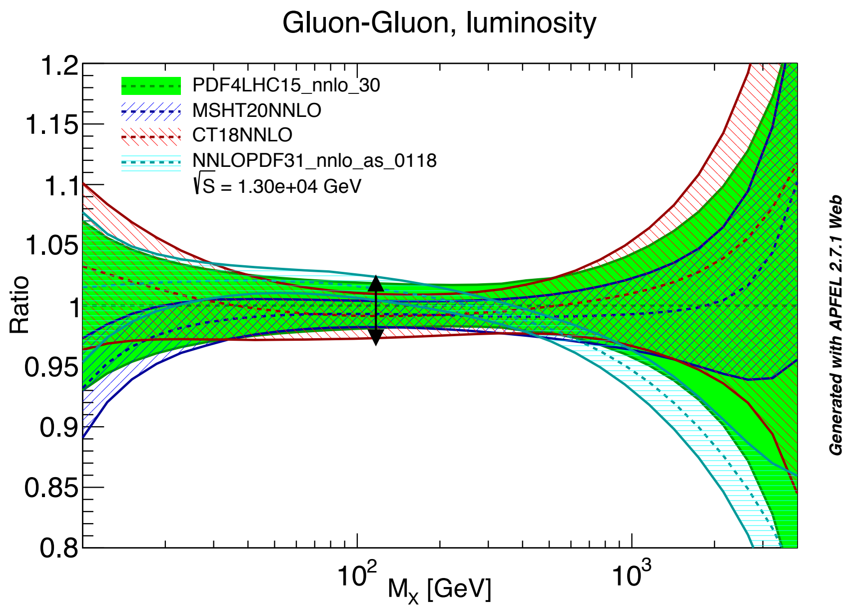


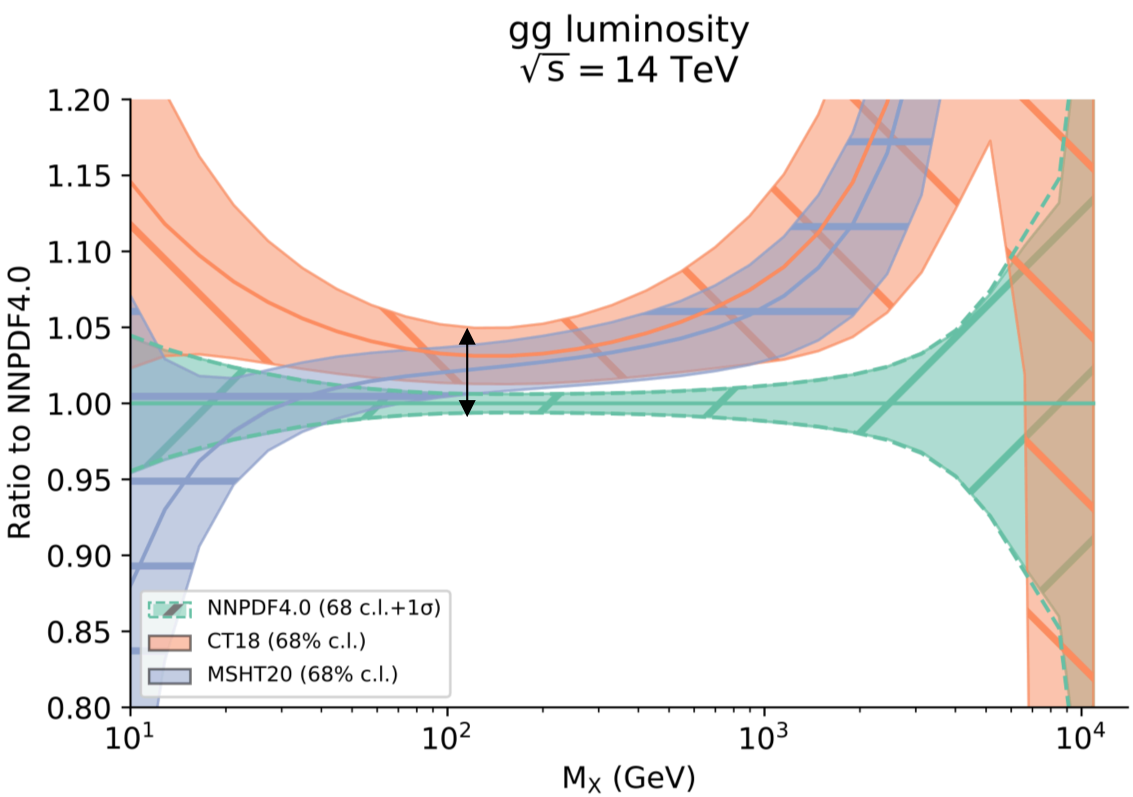
ok, what's the plan

- Add additional jet data sets into reduced set (in progress), consider impact on ATLAS tt and on gluon distribution
  - ATLAS and CMS 7 TeV data
- Use L2 sensitivity to understand impact of each data set on PDF fits (in progress)
  - differences between impact of Tier 2/dynamic tolerances on full fit compared to reduced fit
- Expand to complete CT18, MSHT20, NNPDF3.1 global fits
- Provide PDF4LHC21 PDF sets, collect even more citations

PDF4LHC21 and NNPDF4.0

from Maria's talk on Monday, Tommaso's talk this session: the situation for gg looks different for NNPDF4.0 than for 3.1; spread of central PDFs would stillcontribute to gg PDF uncertainty (but plan is to use NNPDF3.1 in PDF4LHC21)





Some points for discussion

- Any lessons from experience with PDF4LHC15 (from a user perspective) that we should take into account?
- There are some variations that we could consider for additional PDF4LHC21 results
  - perturbative vs fitted charm for NNPDF
  - small x resummation effects->affects low x gluon
    - very important for 100 TeV collider (PDF4100TeV?)
  - any need for such additional sets?
- A new question: inclusive jets vs dijets
  - my personal opinion: should be no significant difference between inclusive jets and dijets; with correlations can use both (studies underway in ATLAS)


You know, it's very strange
I have been in the Remove Trump business so long, now that it's over, I don't know what to do with the rest of my life
Have you ever considered PDF benchmarking?
[redacted]

PDF4LHC15 exercise

- 300 Monte Carlo replicas generated from error sets of each of the 3 global PDF sets; information can be summarized in far fewer error PDFs