Update on the PDF4LHC20 benchmarking exercise

Juan Rojo
VU Amsterdam & Theory group, Nikhef

11/08/2020, PDF4LHC meeting

Motivation

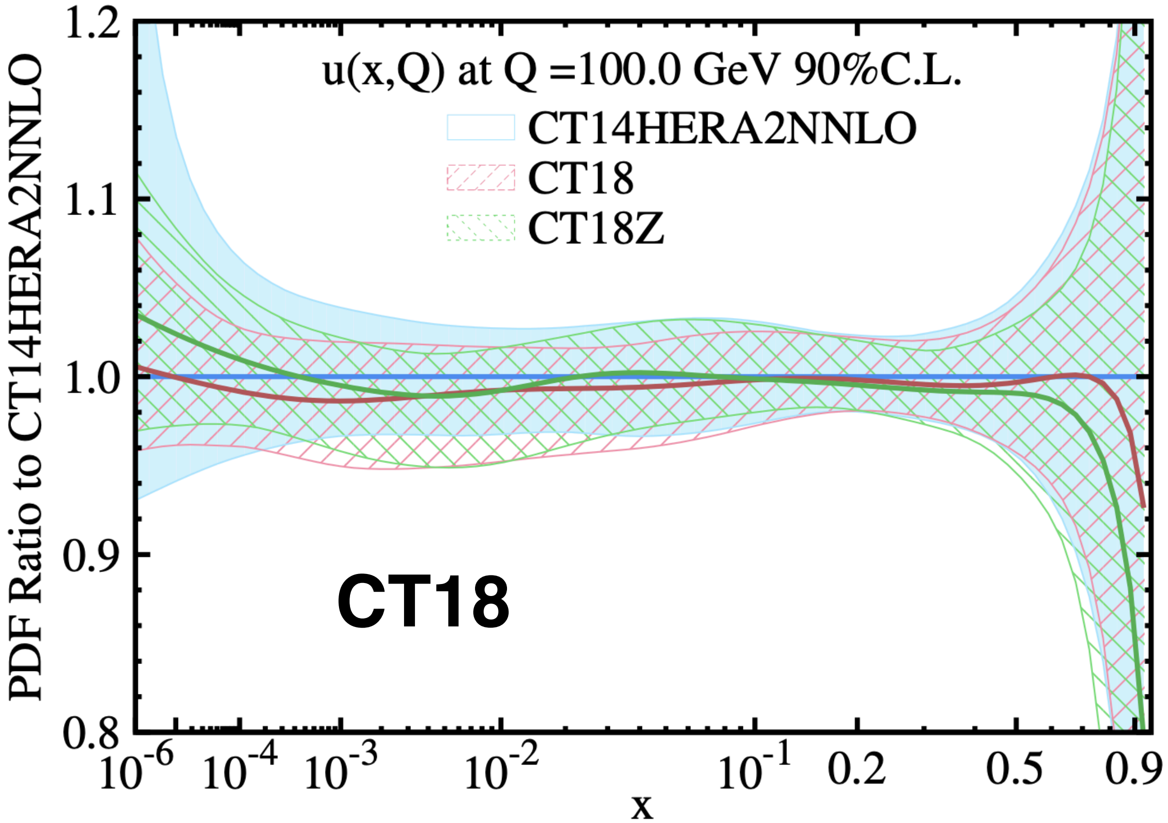
Long tradition of fruitful **PDF benchmarking exercises** that have lead to a fair amount of progress and improved understanding within the community

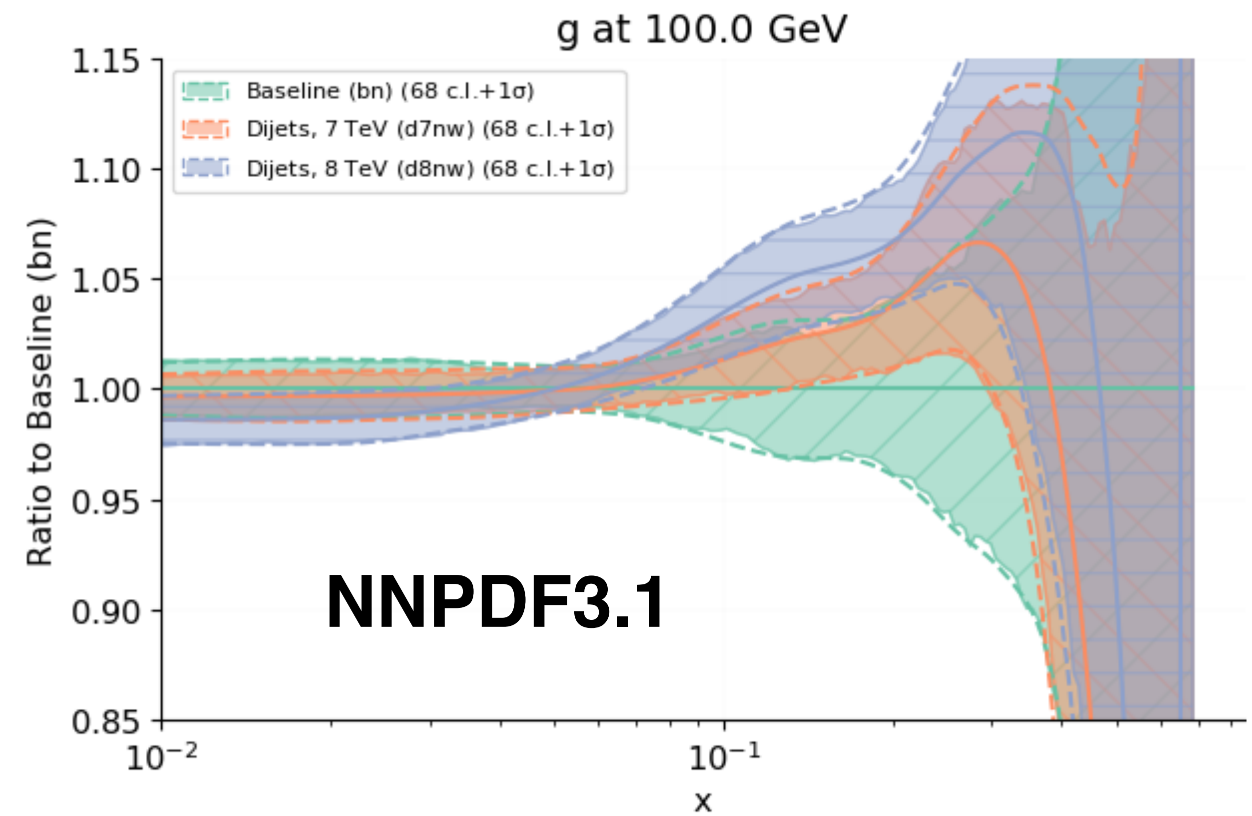
an incomplete list of **past activities** includes:

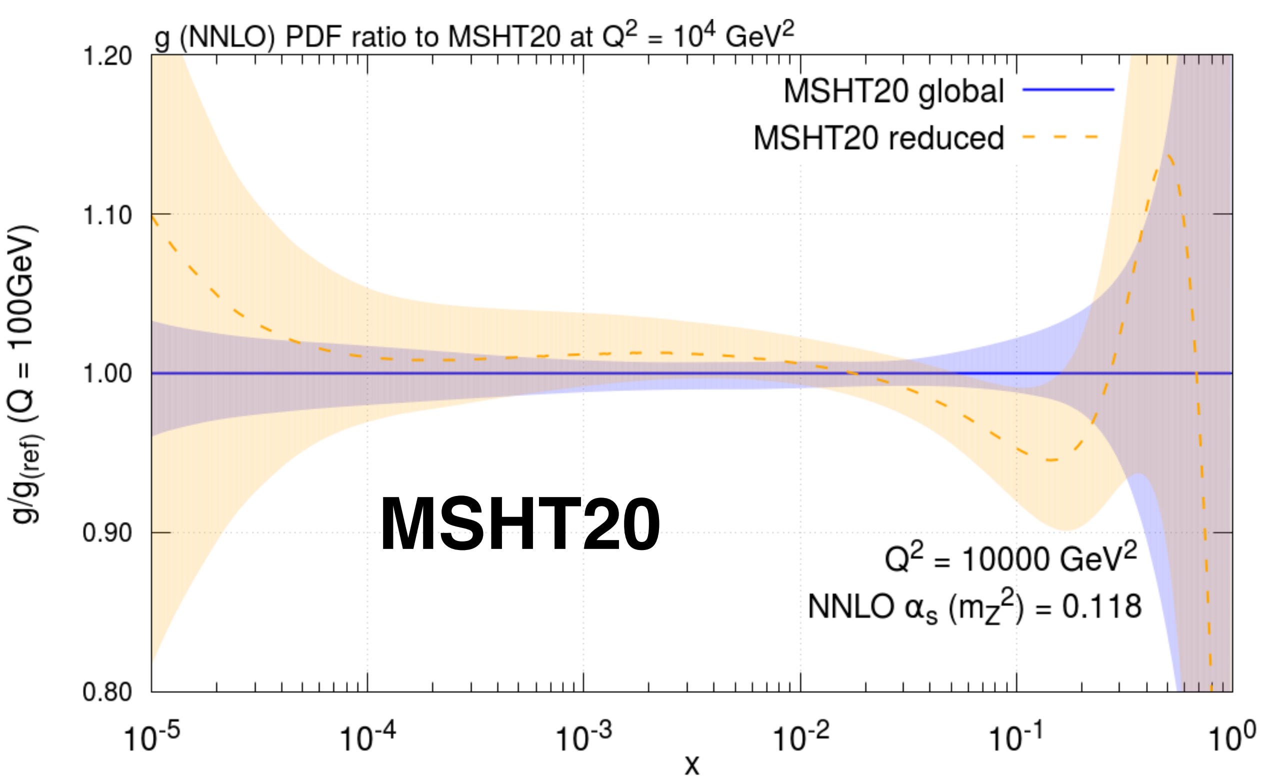
- HERA and the LHC workshop PDF benchmarks (incl. DGLAP evolution comparisons)
- PDF4LHC 2010 benchmark and recommendations for PDF usage at Run I
- Studies in the framework of Higgs Cross-Section Working Group and Yellow Reports
- PDF studies in Snowmass 2013
- 2013 PDF benchmarking study with LHC data
- PDF4LHC 2015 recommendations from Run II
- PDF4LHC 2015 report on the impact of LHC data
- Various PDF activities in the context of the Les Houches workshops
- …….

A great deal of productive activity within the PDF community within the last 10 years!

Current PDF landscape

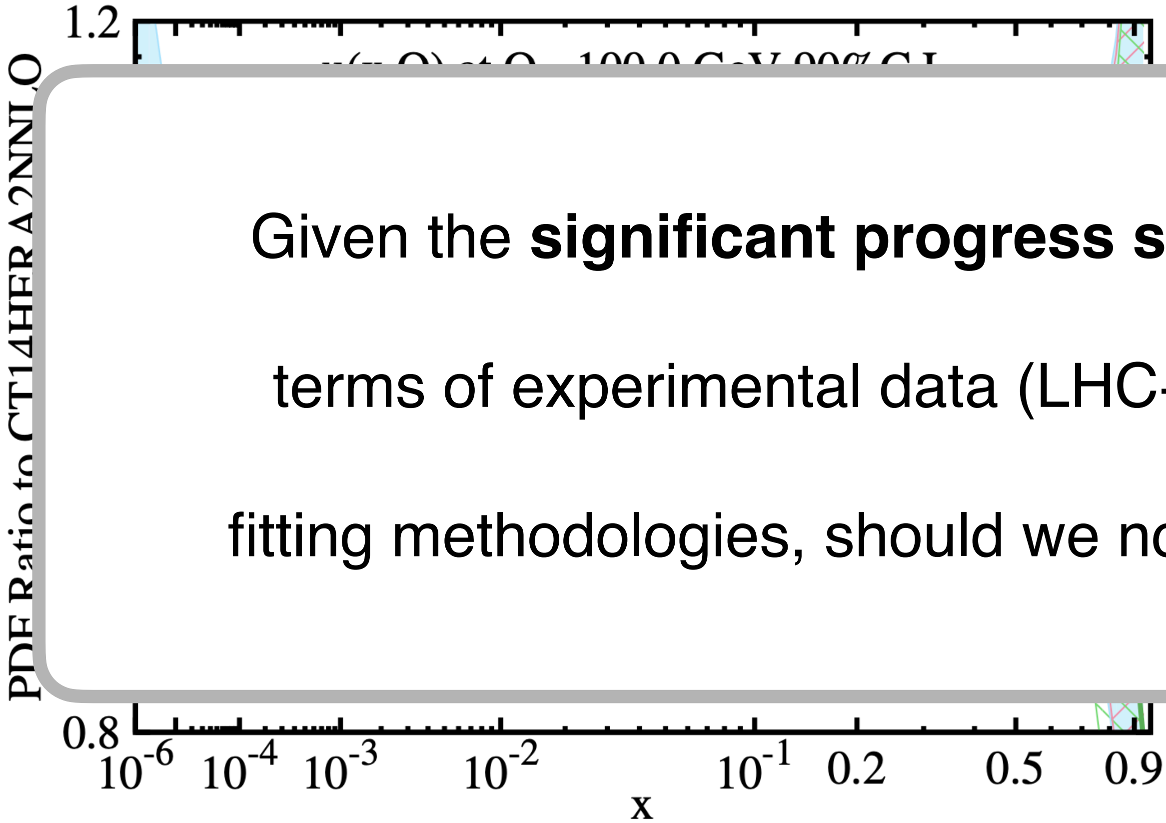


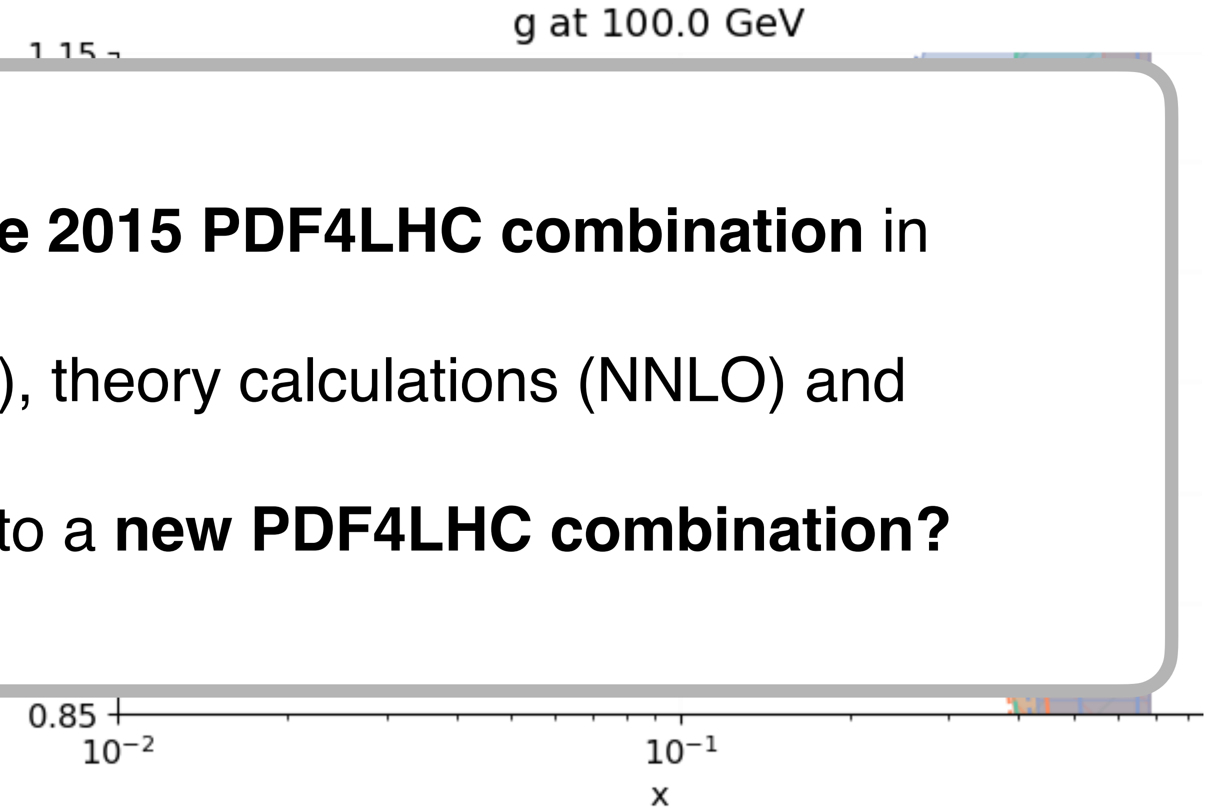




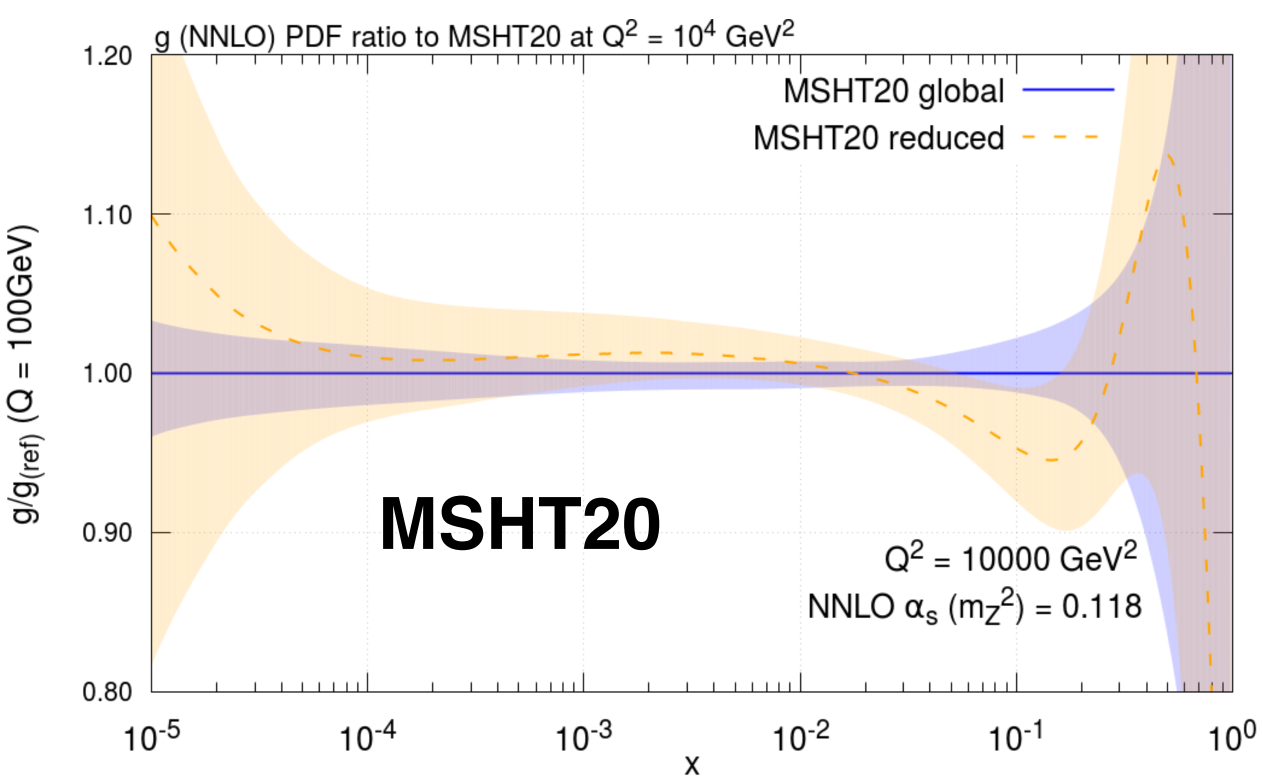
- These three global are based on a **comparable dataset**, with ATLAS W,Z 2011, ATLAS & CMS inclusive jets , LHCb W,Z production, top quark differential ….
- In addition to new data, **methodological improvements** *e.g.* in parametrisation
- Several **spin-offs** and **dedicated studies** concerning top production, jets, strangeness, QED effects and the photon PDF, heavy quarks …

Current PDF landscape





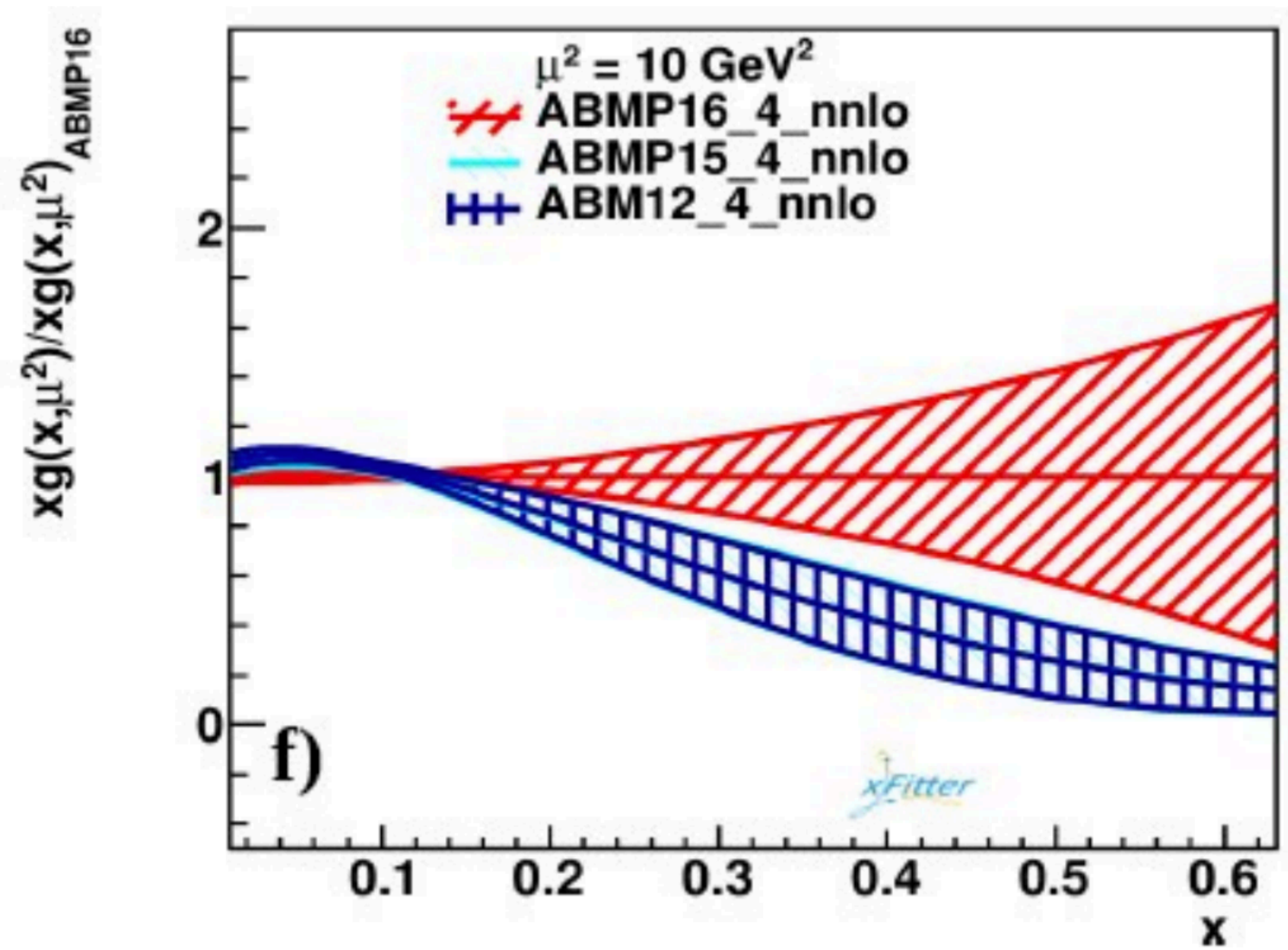
Given the **significant progress since the 2015 PDF4LHC combination** in terms of experimental data (LHC+HERA), theory calculations (NNLO) and fitting methodologies, should we now aim to a **new PDF4LHC combination?**



- These three global are based on a **comparable dataset**, with ATLAS W,Z 2011, ATLAS & CMS inclusive jets , LHCb W,Z production, top quark differential ….
- In addition to new data, **methodological improvements** *e.g.* in parametrisation
- Several **spin-offs** and **dedicated studies** concerning top production, jets, strangeness, QED effects and the photon PDF, heavy quarks …

Current PDF landscape

ABMP16



Significant constraints from LHC data, improved approximate NNLO calculations, impact *e.g.* on large-x gluon

plus many related activities!

- PDF studies within ATLAS and CMS
- Collinear PDFs together with polarised PDFs and FF (JAM collaboration)
- Updates of the HERAPDF analyses
- PDF projections for HL-LHC and LHeC
- Studies of combined PDF & EFT fits
- PDFs including the constraints from lattice QCD calculations
- Novel sensitivity metrics in PDF space
- ……

Goal of the 2020 benchmarking exercise

Understanding as much as possible the origin of the differences between **CT18, MSHT20, and NNPDF3.1**, with the motivation of their eventual combination: **PDF4LHC20**

and what is the methodology adopted to achieve these goals?

- Compare the resulting PDFs obtained in fits based on **identical input dataset and theoretical settings** between the three groups (or as close as possible)
- Compare the values obtained for the **χ^2 for each dataset** (using common definitions)
- Compare the **point-by-point cross-sections** (separately at NLO and NNLO) obtained by the different groups when the theoretical settings (input PDF sets, QCD and QED couplings, heavy quark masses) are the same

in other words: which of the observed differences can be traced back to variations of the experimental input, of the theory settings, and to the methodologies of each respective group?

which of these differences arise from equally valid but different choices between the groups?

Goal of the 2020 benchmarking exercise

Understanding as much as possible the origin of the differences between **CT18, MSHT20, and NNPDF3.1**, with the motivation of their eventual combination: **PDF4LHC20**

What are the main challenges that need to be tackled?

- Impact of **possible tensions** between datasets or groups of processes
- Different implications/interpretations/consequences of these tensions in the three groups
- **Ambiguities** on the choice of differential distributions to fit for some datasets
- PDF uncertainties include experimental and parametrisation uncertainties (via tolerance in the Hessian fits), with differences can still arise from **theory settings** (e.g. scales) and MHOUs. Even at NNLO the latter are expected to be important.

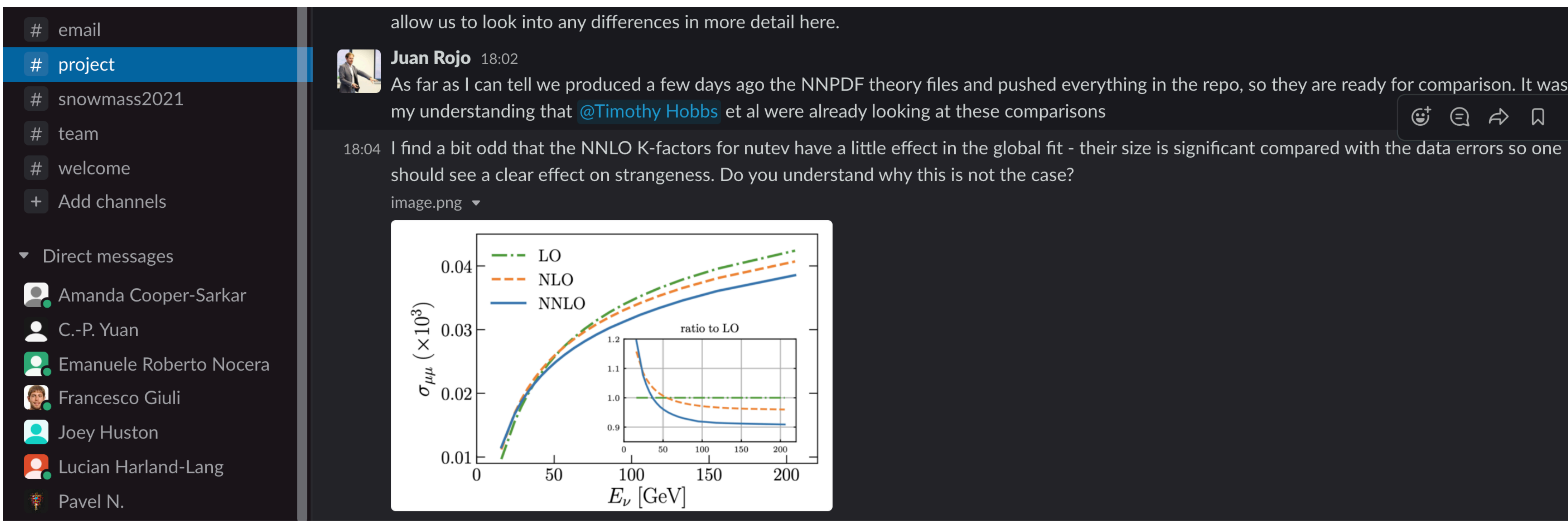
we aim to identify not only the differences, but also to trace back the origin of the **similarities** between the three global fits, to motivate their **eventual combination**

Goal of the 2020 benchmarking exercise

Understanding as much as possible the origin of the differences between **CT18, MSHT20, and NNPDF3.1**, with the motivation of their eventual combination: **PDF4LHC20**

to maximise efficiency, discussion and communication of the results takes place a **dedicated Slack workspace** (everyone interested is encouraged to join!)

pdf4lhc20benchmarking.slack.com

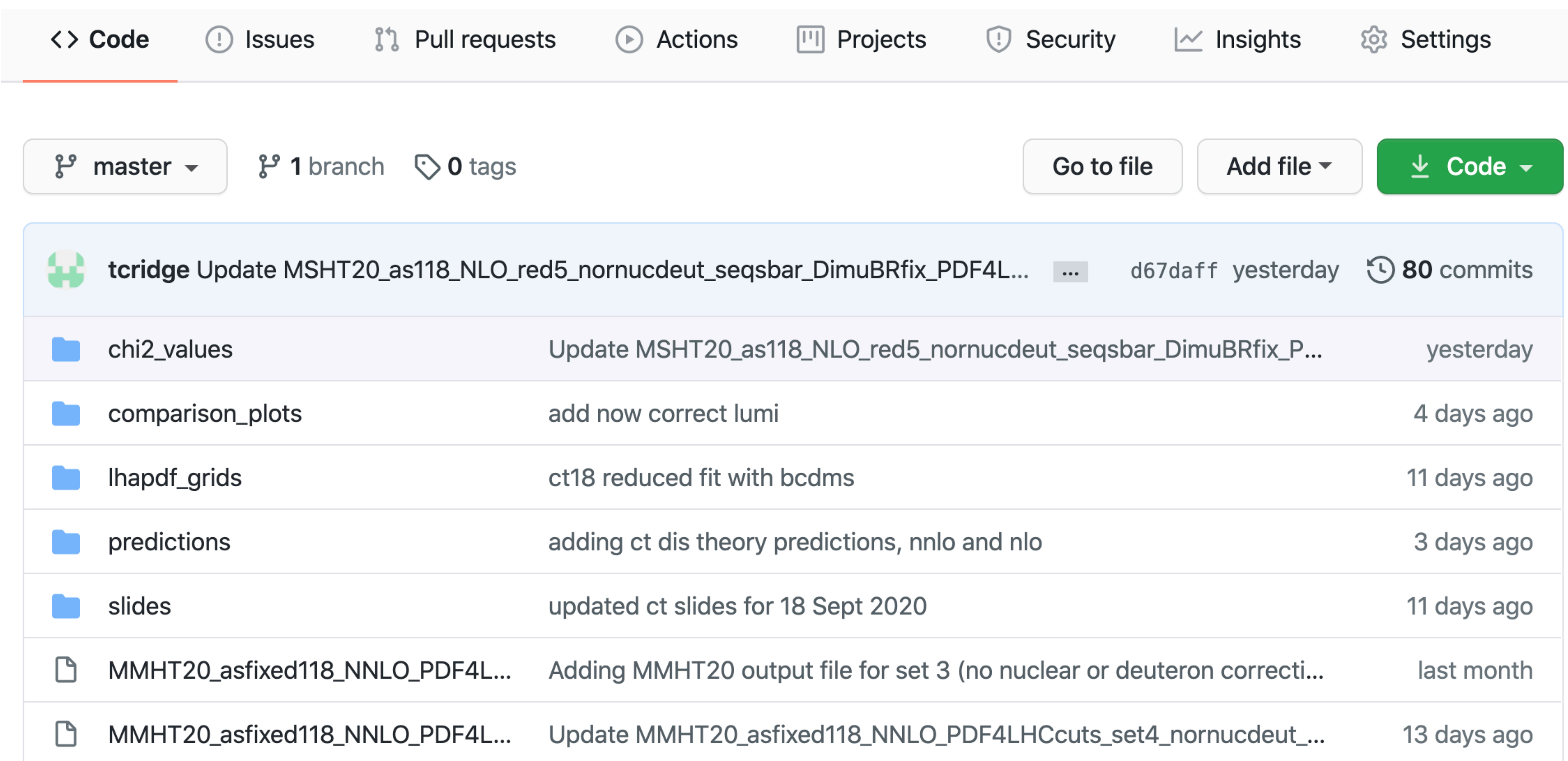


Goal of the 2020 benchmarking exercise

Understanding as much as possible the origin of the differences between **CT18, MSHT20, and NNPDF3.1**, with the motivation of their eventual combination: **PDF4LHC20**

further, we **collect the results** of this benchmarking exercise (LHAPDF grids, χ^2 tables, cross-section tables, comparison plots of PDFs and luminosities) in a **GitHub repo**

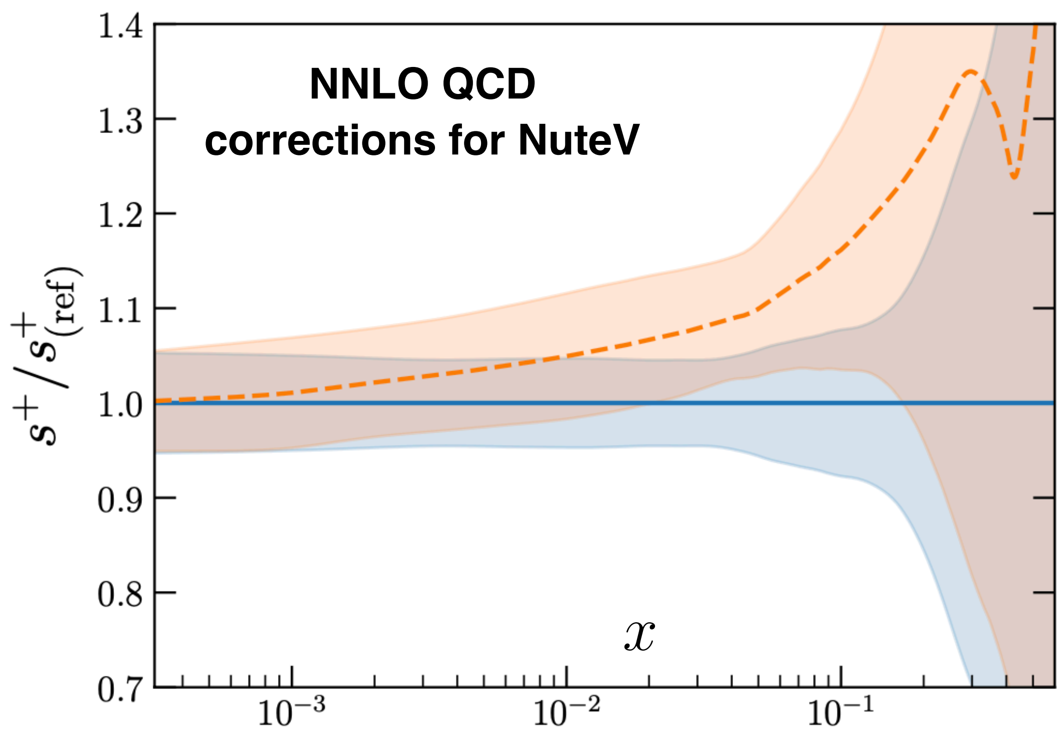
https://github.com/juanrojochacon/pdf4lhc20

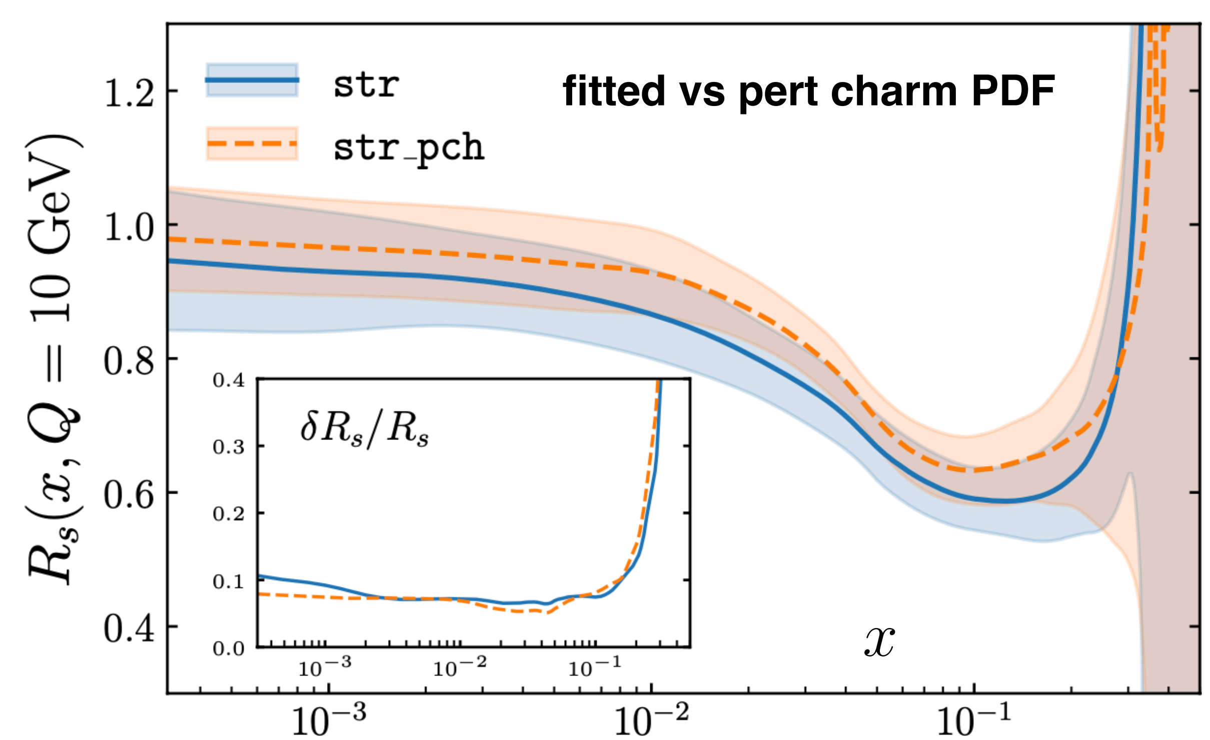


PDF fits based on a common dataset

- For the sake of the benchmarking, adopted the following **common theoretical settings:**
  - ☑ No deuteron or nuclear corrections
  - ☑ Vanishing strangeness asymmetry
  - ☑ Positive-definite PDFs in the flavour basis
  - ☑ Perturbative charm
  - ☑ NNLO massive corrections for the NuTeV dimuon data
  - ☑ Fixed branching fraction to muons for the NuTeV dimuon data
  - ☑ Common values of the strong coupling and heavy quark masses

note that this are not necessarily well-motivated choices: we adopt them solely for the benchmarking





PDF fits based on a common dataset

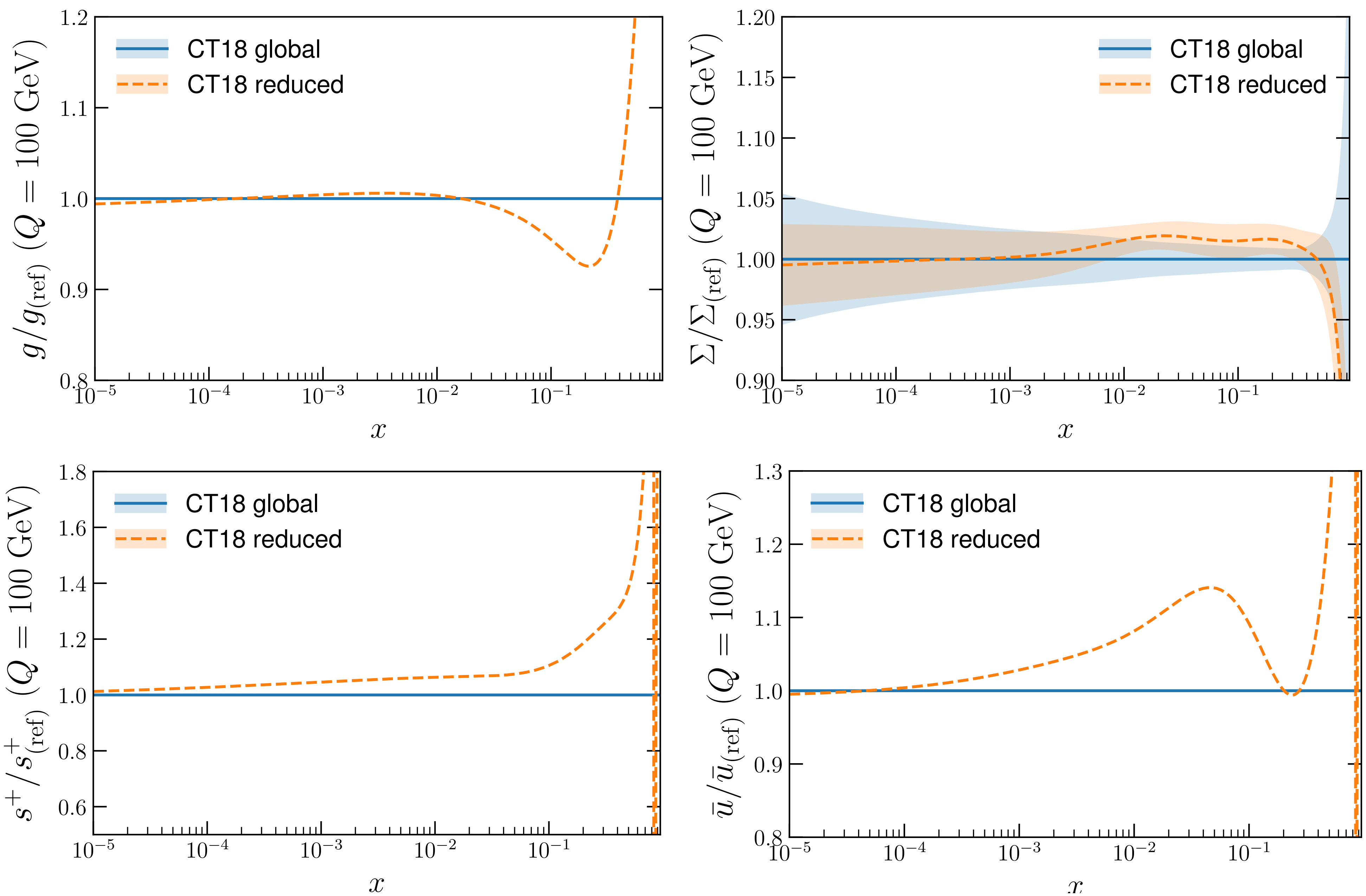
- For the sake of the benchmarking, adopted the following **common dataset:**
  - **NMC** deuteron-to-proton ratio
  - **BCDMS** proton and deuteron structure functions
  - **NuTeV** dimuon cross-sections
  - **HERA I+II** combination of inclusive structure functions
  - Drell-Yan **E866** deuteron-to-proton ratio
  - **D0** Z rapidity distribution
  - **ATLAS W,Z** inclusive 2010+2011 (only central rapidity region)
  - **CMS W** electron asymmetry
  - **CMS inclusive jets** at 8 TeV
  - **LHCb 7,8 TeV W,Z** rapidity distributions

see the documentation available on Slack concerning the specific details of each dataset

- Note that many other datasets **frequently used in PDF fits** are not considered here since *e.g.* they have not implemented by the three groups or are treated differently
  - CHORUS neutrino DIS, HERA F2charm, ATLAS one-jet, top quark pair production, …

PDF fits based on a common dataset

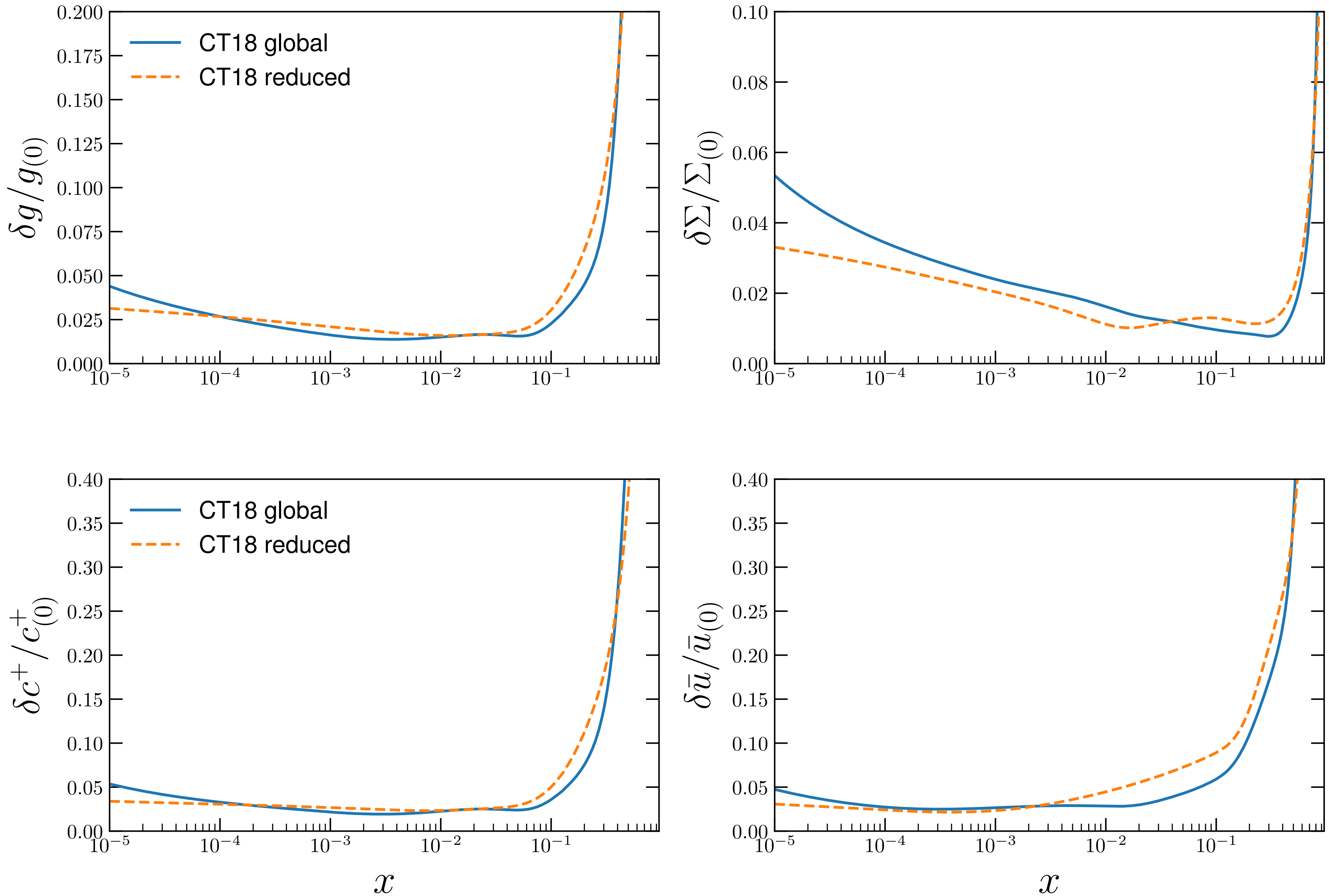
- One can compare the **global and the reduced fits** separately for each of the three groups
- Again, in most cases **differences are expected and understood**
- For the NNPDF reduced sets, the **baseline** is the fit produced in the recent strangeness study



- For CT18 in general good overall compatibility
- The reduced dataset enhances ubar and singlet
- Impact also on gluon
- Only very moderate increase in PDF uncertainties

PDF fits based on a common dataset

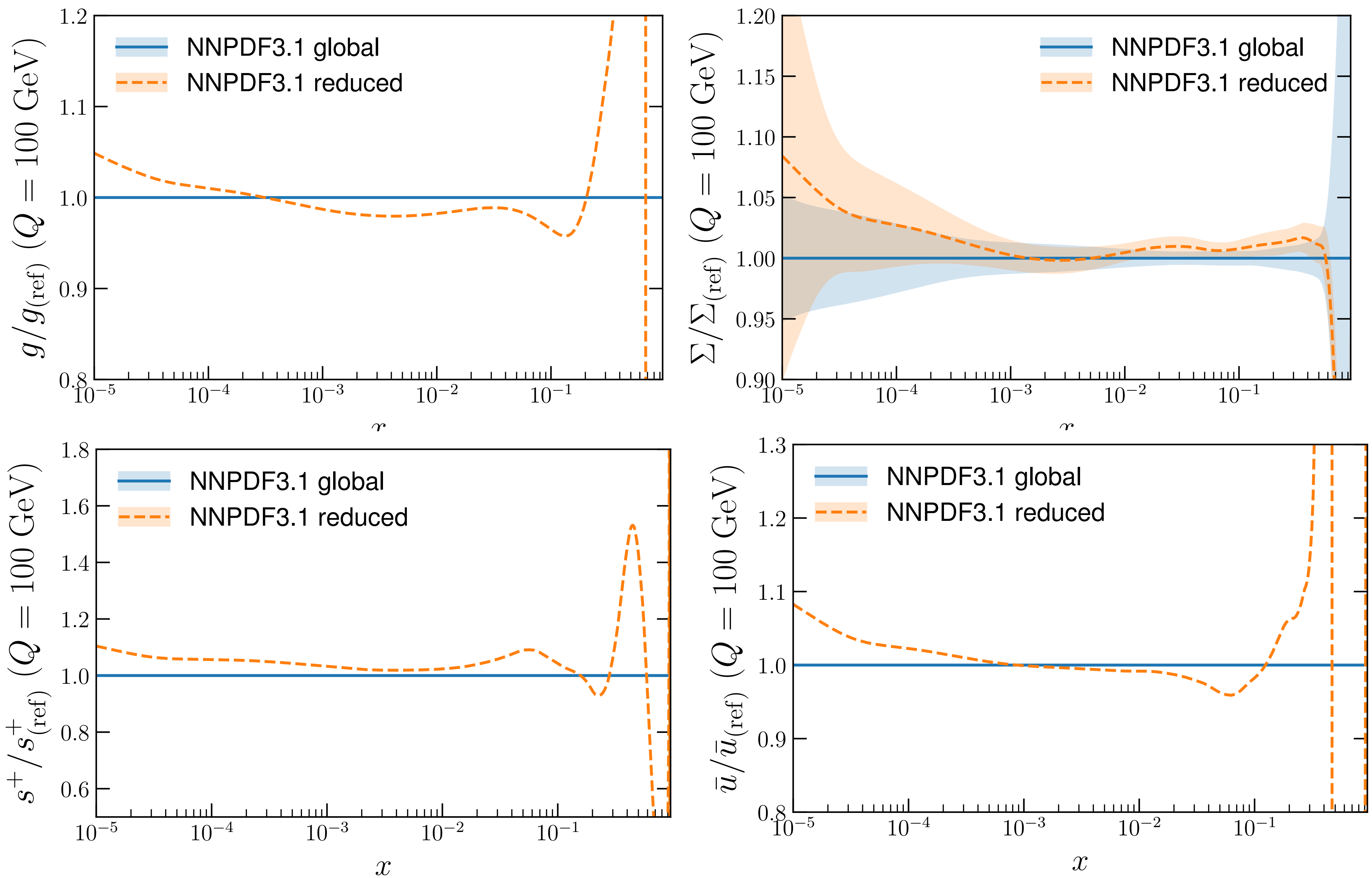
- One can compare the **global and the reduced fits** separately for each of the three groups
- Again, in most cases **differences are expected and understood**
- For the NNPDF reduced sets, the **baseline** is the fit produced in the recent strangeness study



- For CT18 in general good overall compatibility
- The reduced dataset enhances ubar and singlet
- Impact also on gluon
- Only very moderate increase in PDF uncertainties

PDF fits based on a common dataset

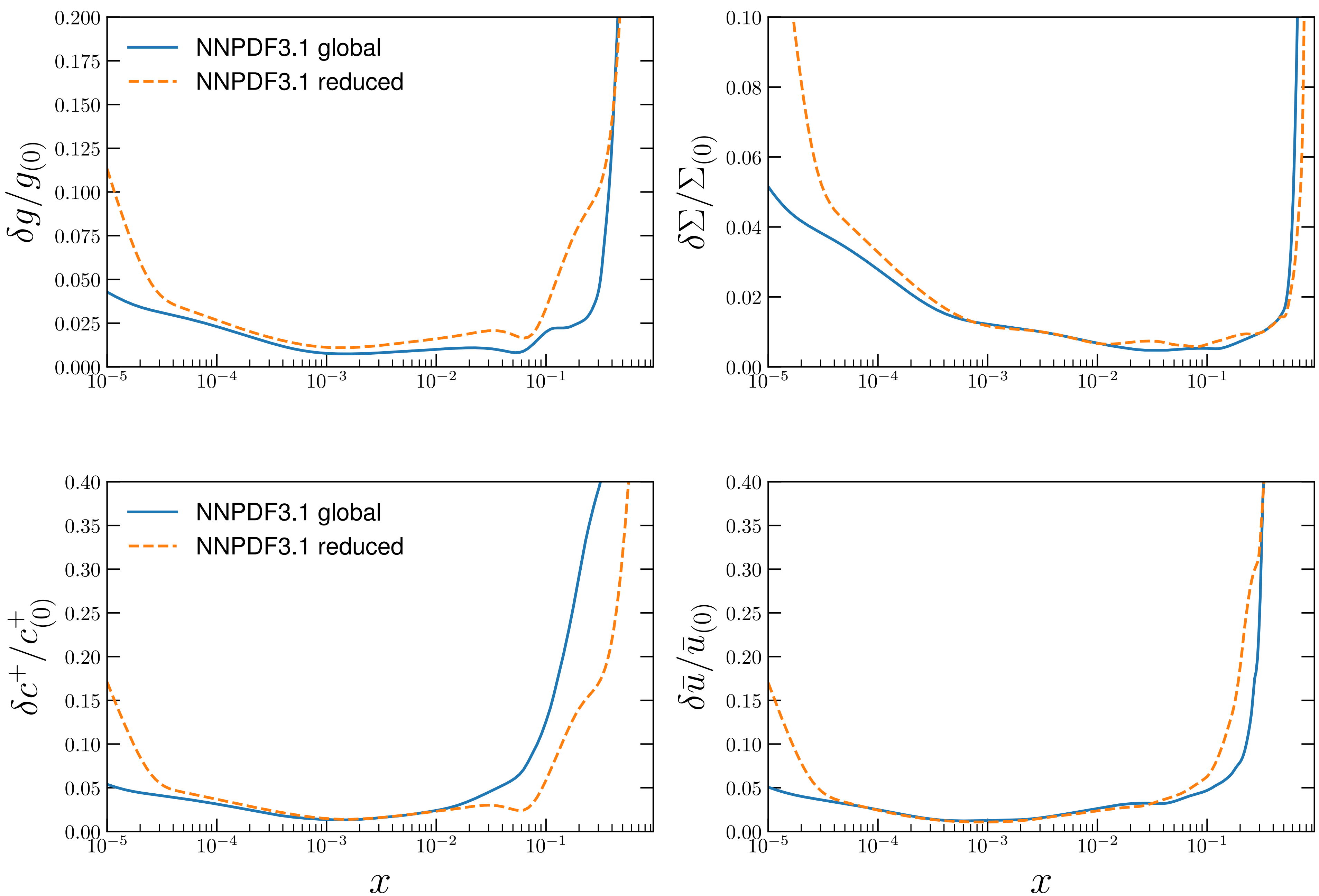
- One can compare the **global and the reduced fits** separately for each of the three groups
- Again, in most cases **differences are expected and understood**
- For the NNPDF reduced sets, the **baseline** is the fit produced in the recent strangeness study



- For NNPDF also good consistency **within PDF uncertainties**
- Differences in **large-*x* gluon** from removal of top data
- Reduced dataset increases errors specially for the gluon

PDF fits based on a common dataset

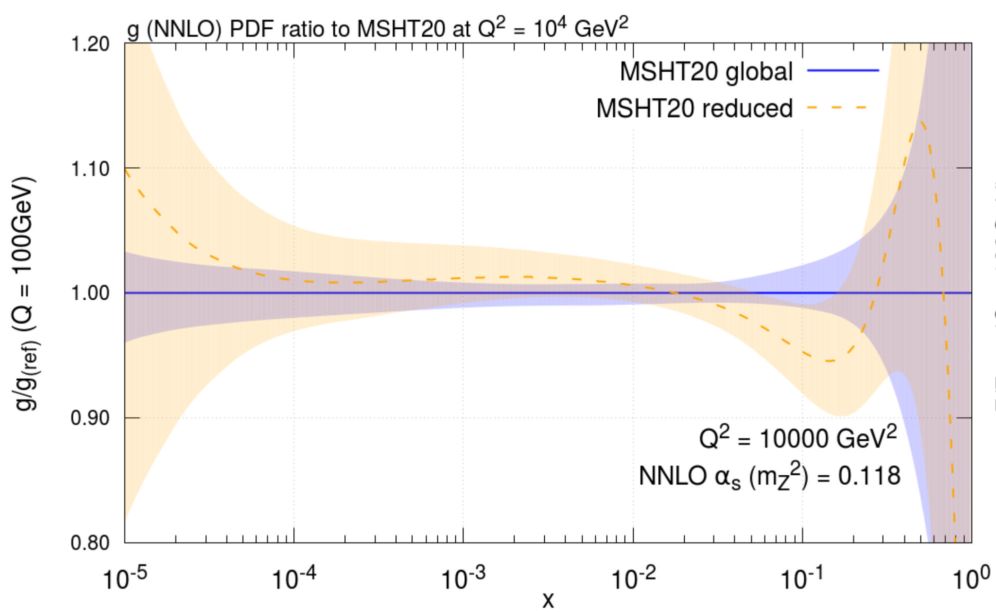
- One can compare the **global and the reduced fits** separately for each of the three groups
- Again, in most cases **differences are expected and understood**
- For the NNPDF reduced sets, the **baseline** is the fit produced in the recent strangeness study

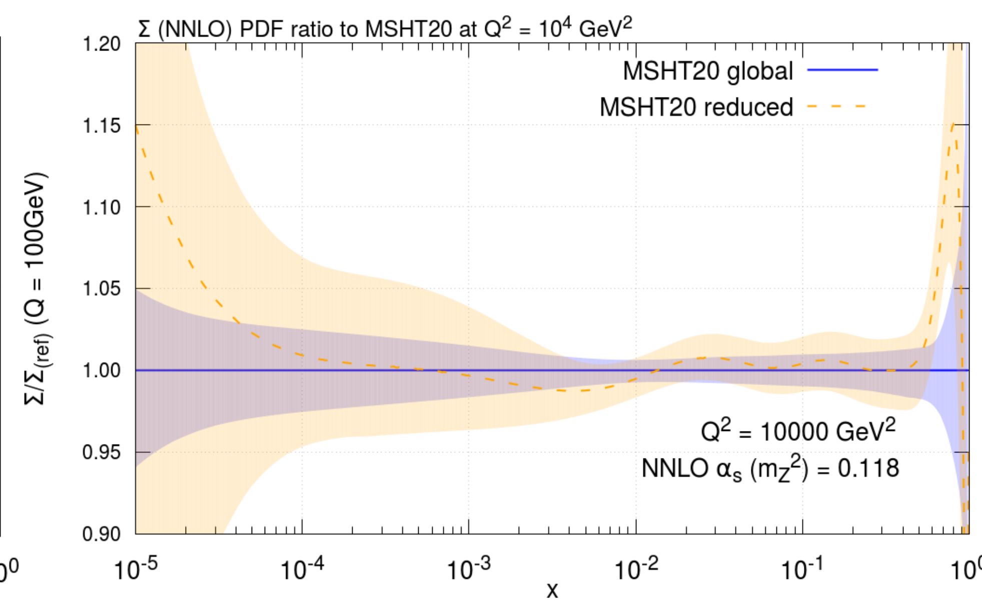


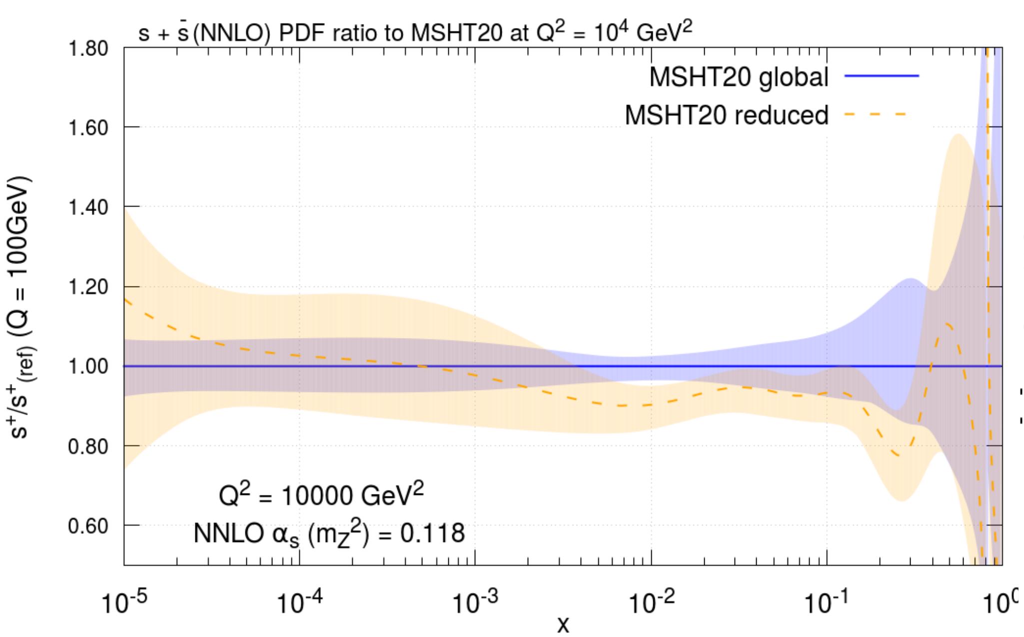
- For NNPDF also good consistency **within PDF uncertainties**
- Differences in **large-*x* gluon** from removal of top data
- Reduced dataset increases errors specially for the gluon

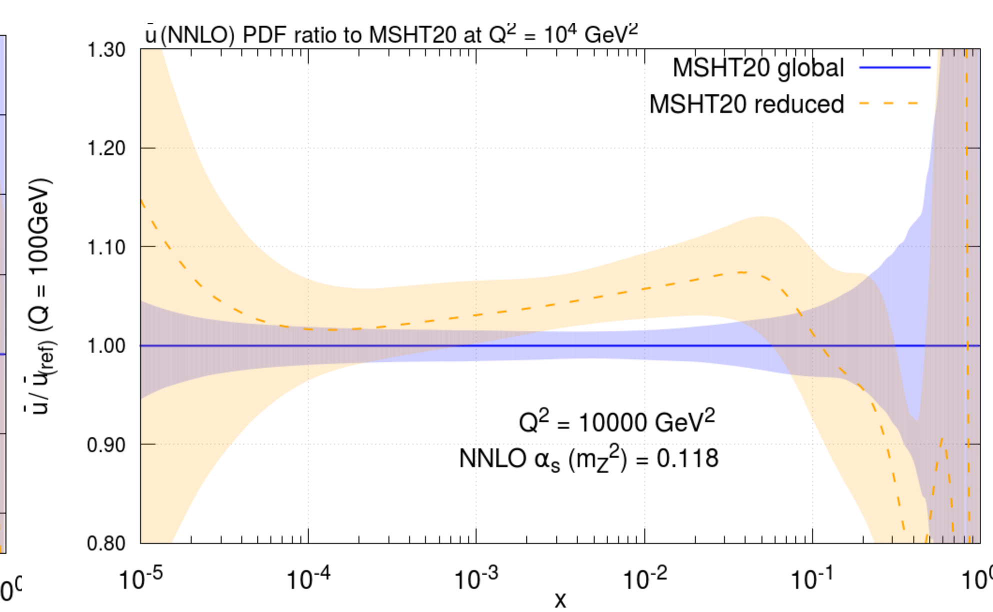
PDF fits based on a common dataset

- One can compare the **global and the reduced fits** separately for each of the three groups
- Again, in most cases **differences are expected and understood**
- For the NNPDF reduced sets, the **baseline** is the fit produced in the recent strangeness study





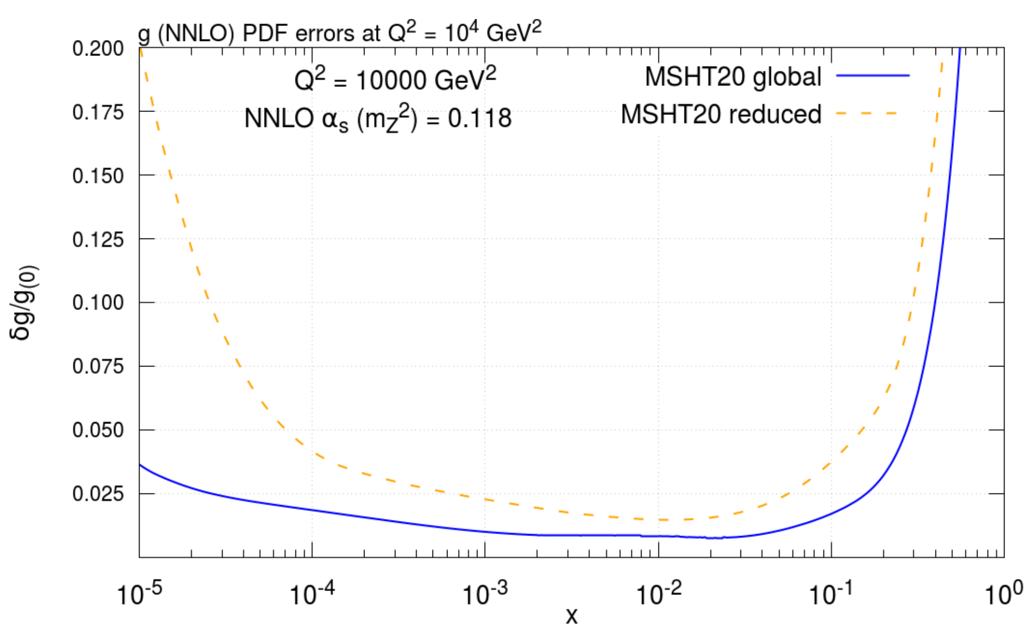


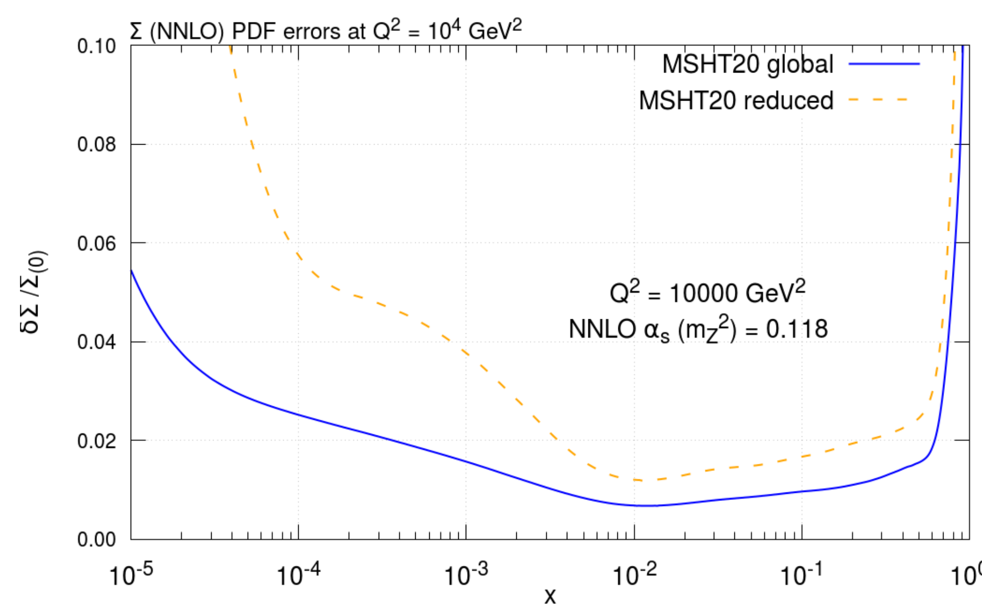


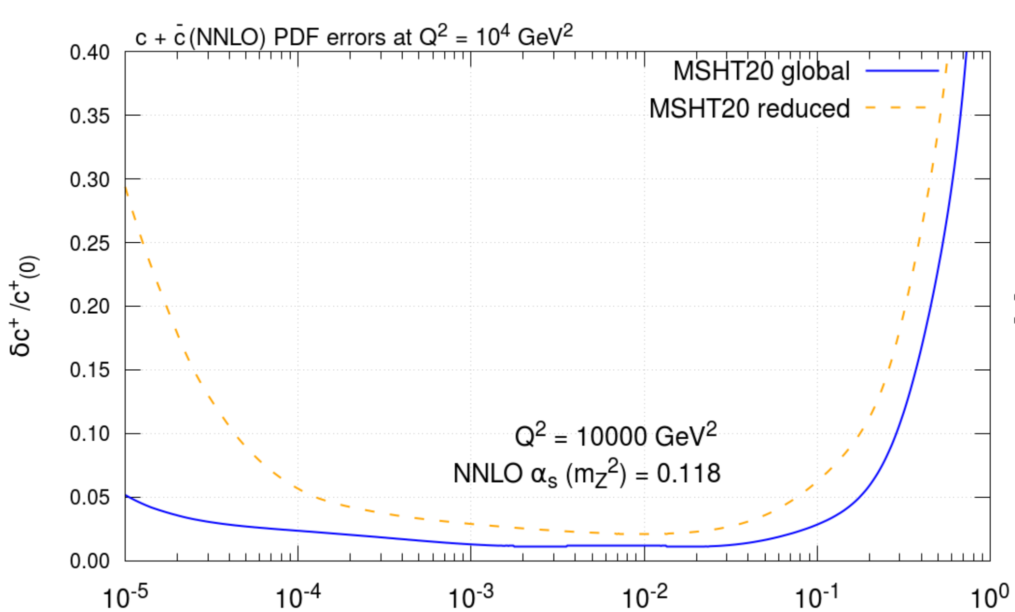
- For MSHT also good consistency (marginal for some flavours)
- Harder large-x valence quarks and medium-x up sea quark
- Reduce dataset brings in changes in **flavour decomposition**
- Increase in PDF errors in reduced dataset fit a bit more marked than for CT18/NNPDF3.1

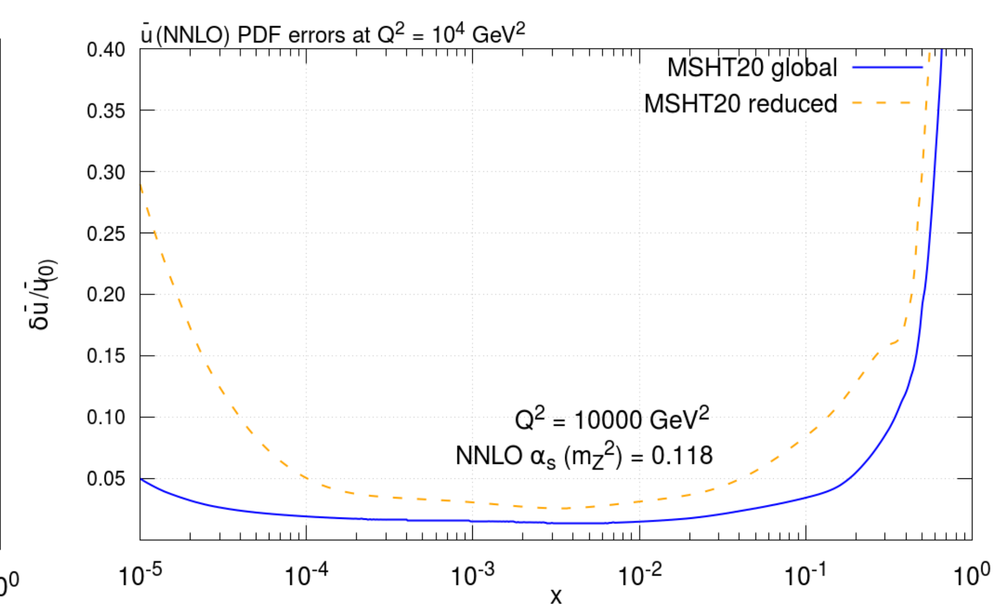
PDF fits based on a common dataset

- One can compare the **global and the reduced fits** separately for each of the three groups
- Again, in most cases **differences are expected and understood**
- For the NNPDF reduced sets, the **baseline** is the fit produced in the recent strangeness study





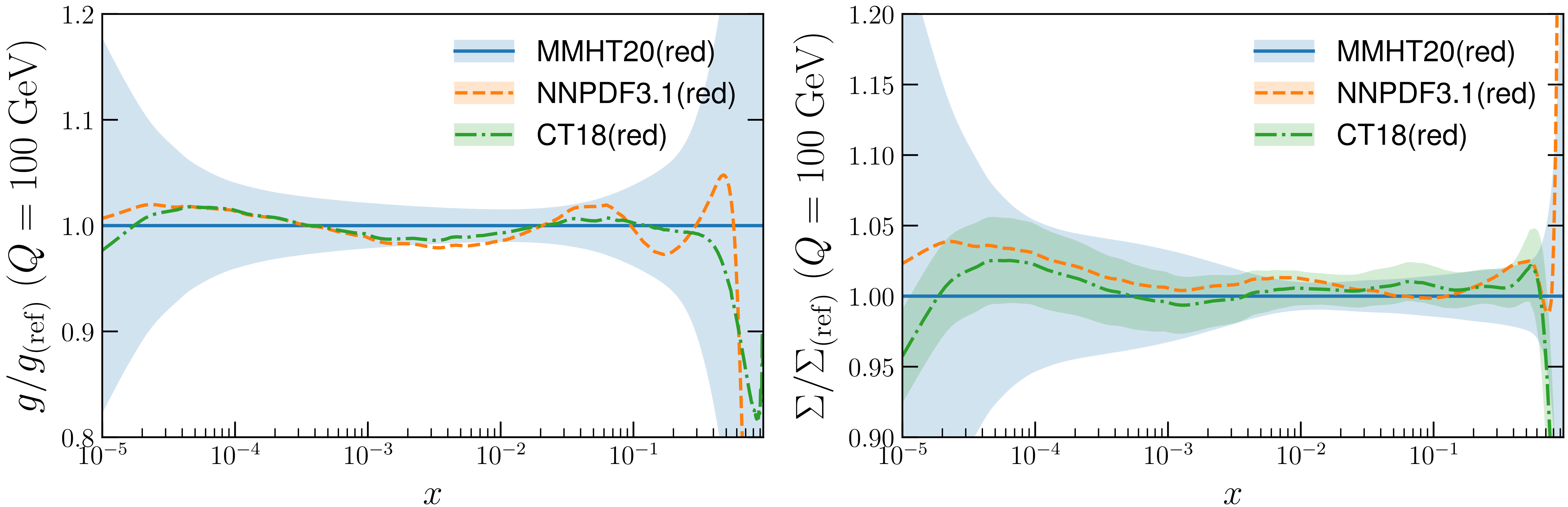


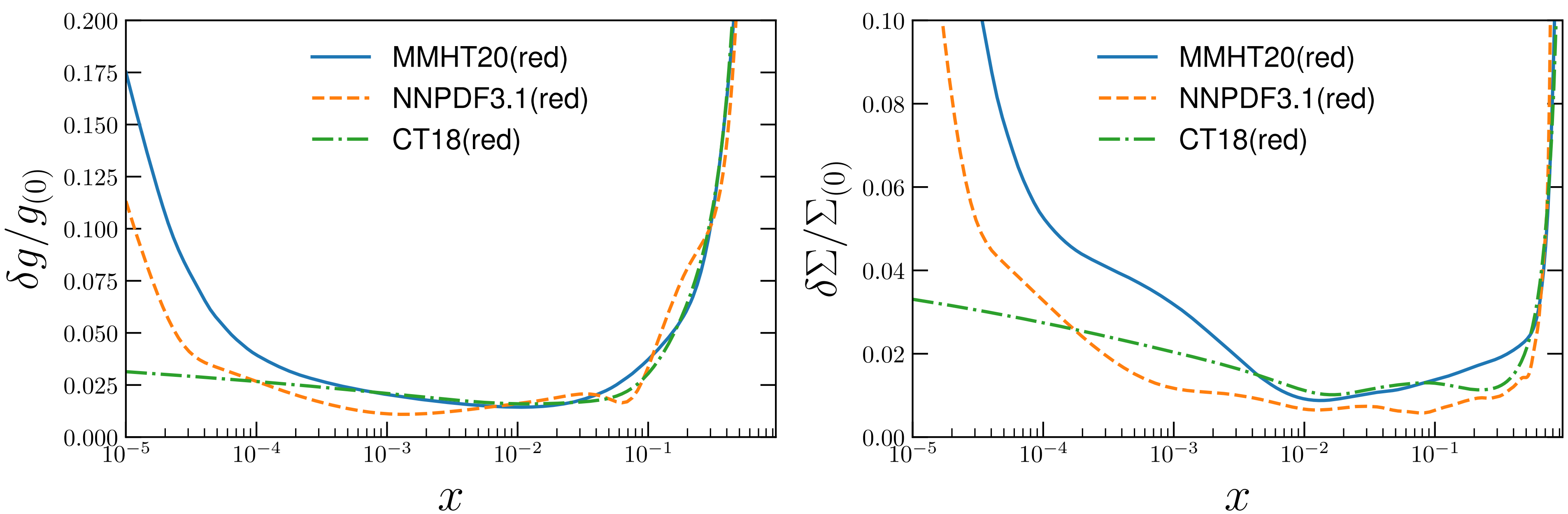


- For MSHT also good consistency (marginal for some flavours)
- Harder large-x valence quarks and medium-x up sea quark
- Reduce dataset brings in changes in **flavour decomposition**
- Increase in PDF errors in reduced dataset fit a bit more marked than for CT18/NNPDF3.1

Comparison of benchmark fits

The goal of adopting a common dataset and the same theory settings was to **reduce the potential sources of differences at the PDF level** between the three groups. Is this the case?

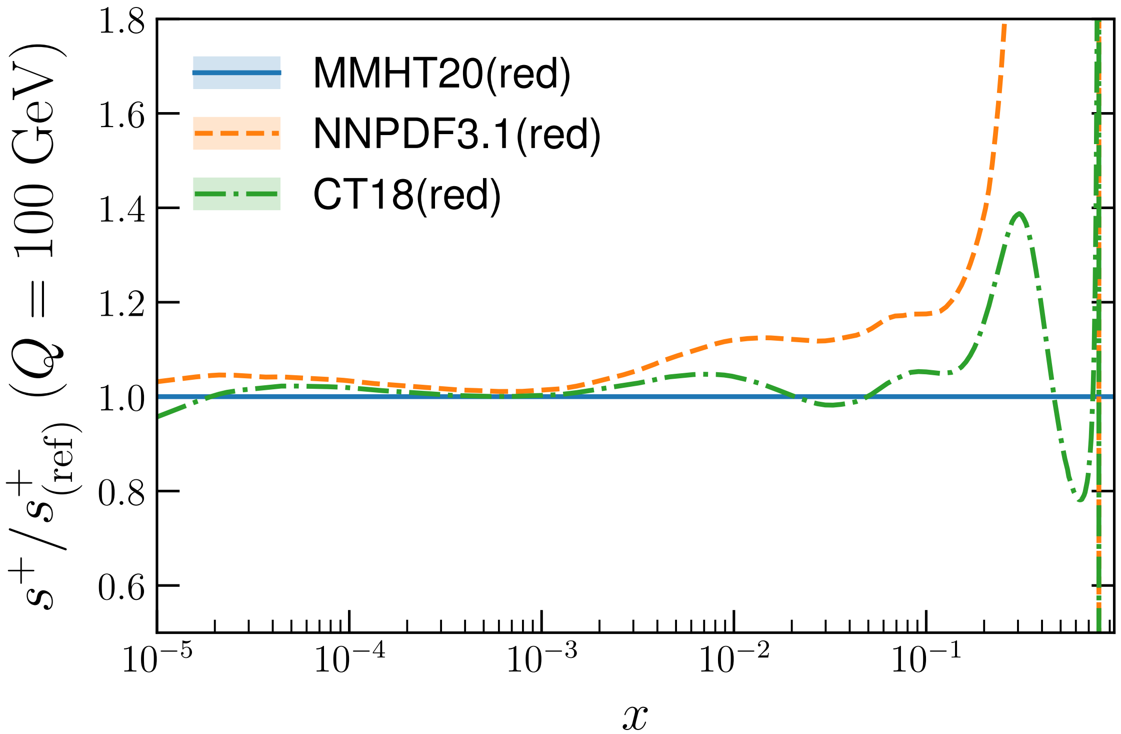


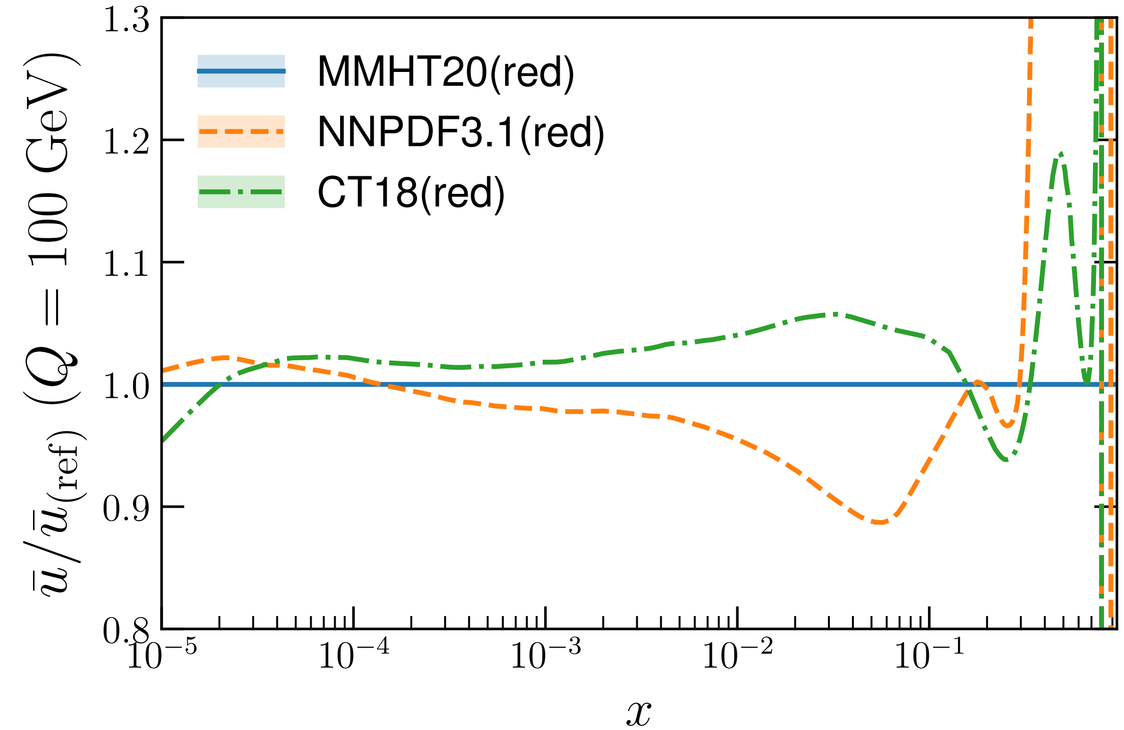


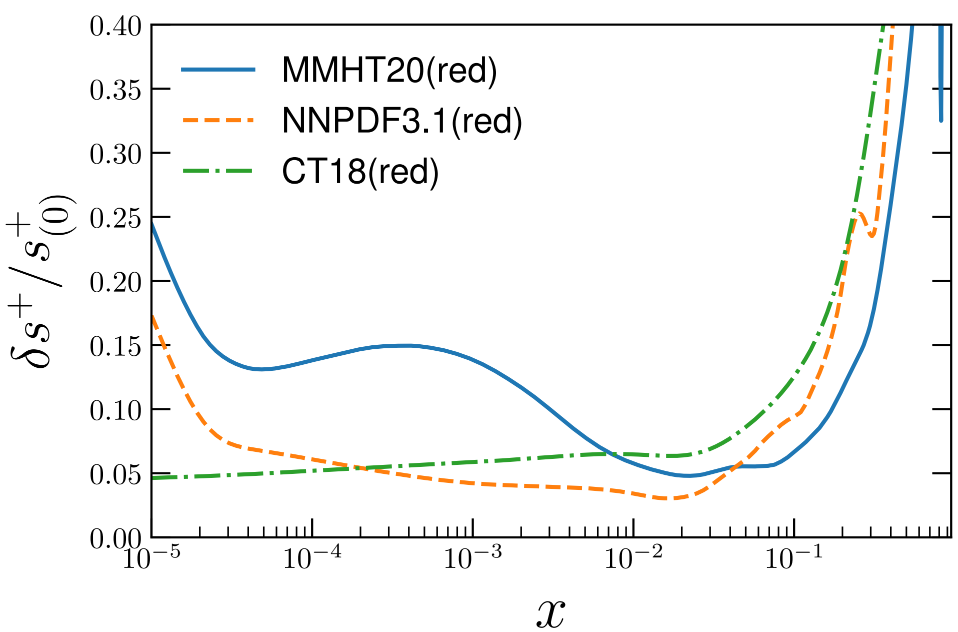
- Good agreement for **gluon and singlet** in full range of *x*
- Similar PDF uncertainties (except extrapolation region)
- Same applies for charm PDF

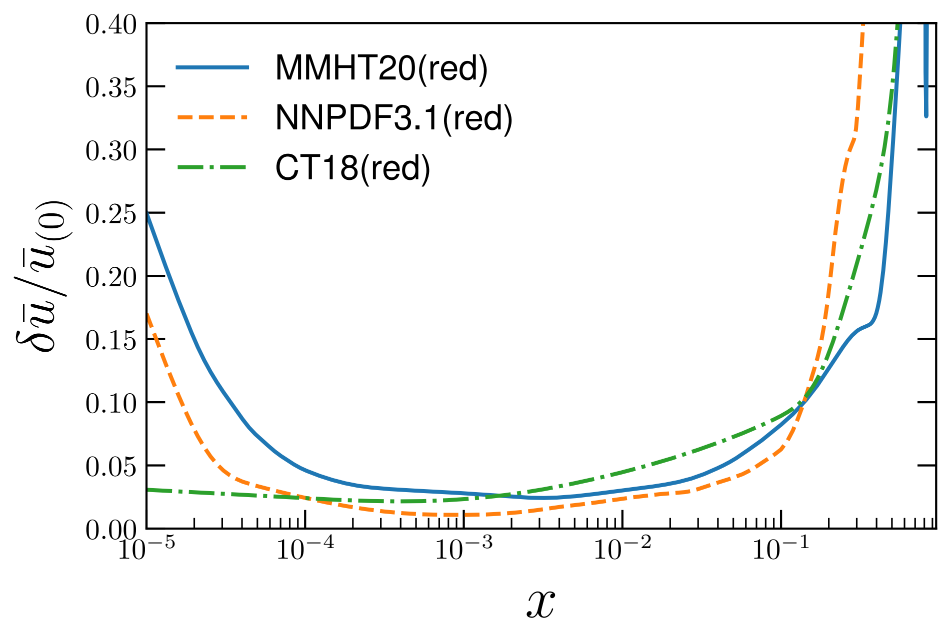
Comparison of benchmark fits

The goal of adopting a common dataset and the same theory settings was to **reduce the potential sources of differences at the PDF level** between the three groups. Is this the case?





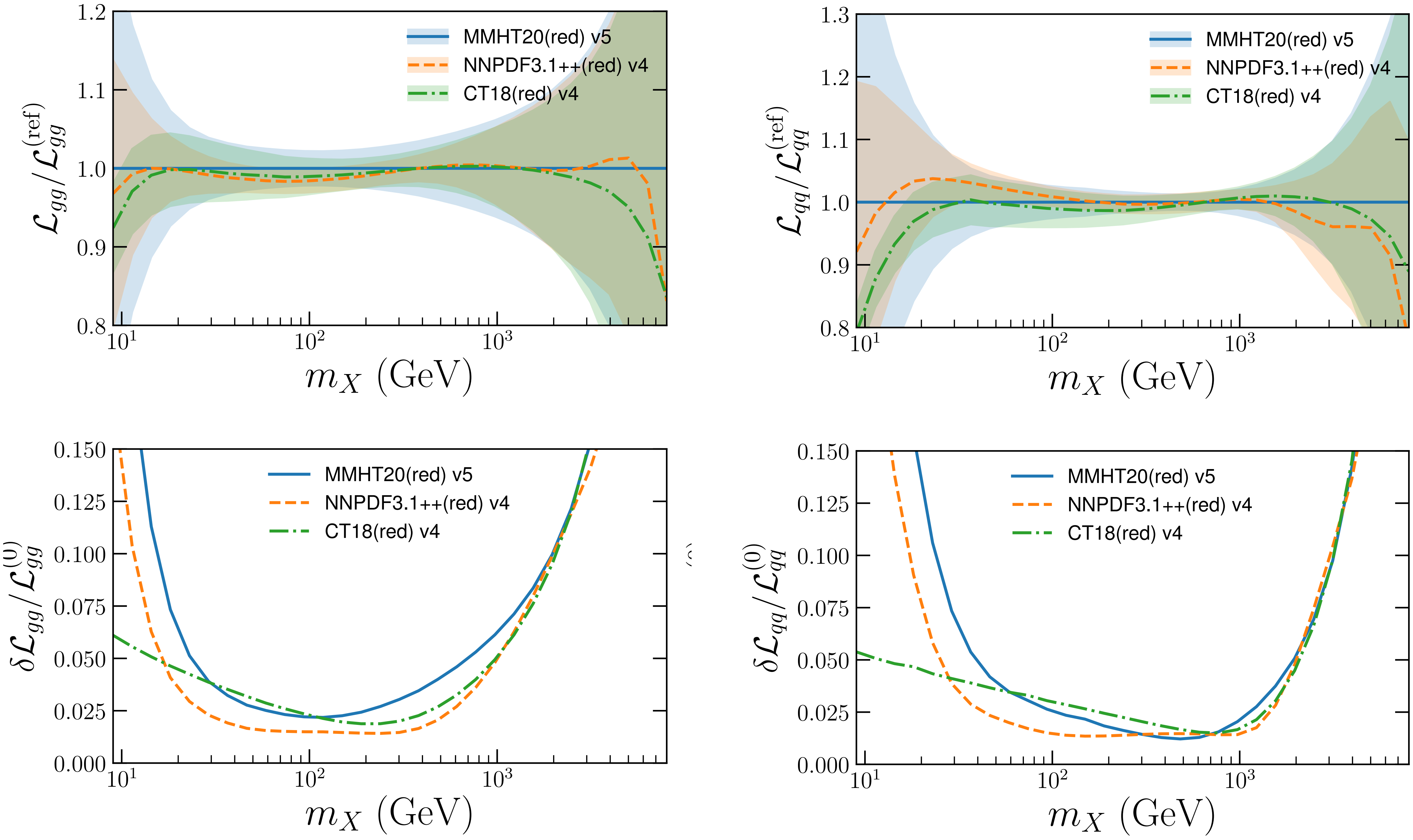




- Good agreement **for sea quarks for x < 0.01**
- Differences in **quark flavour separation for x > 0.01,** specially for NNPDF3.1

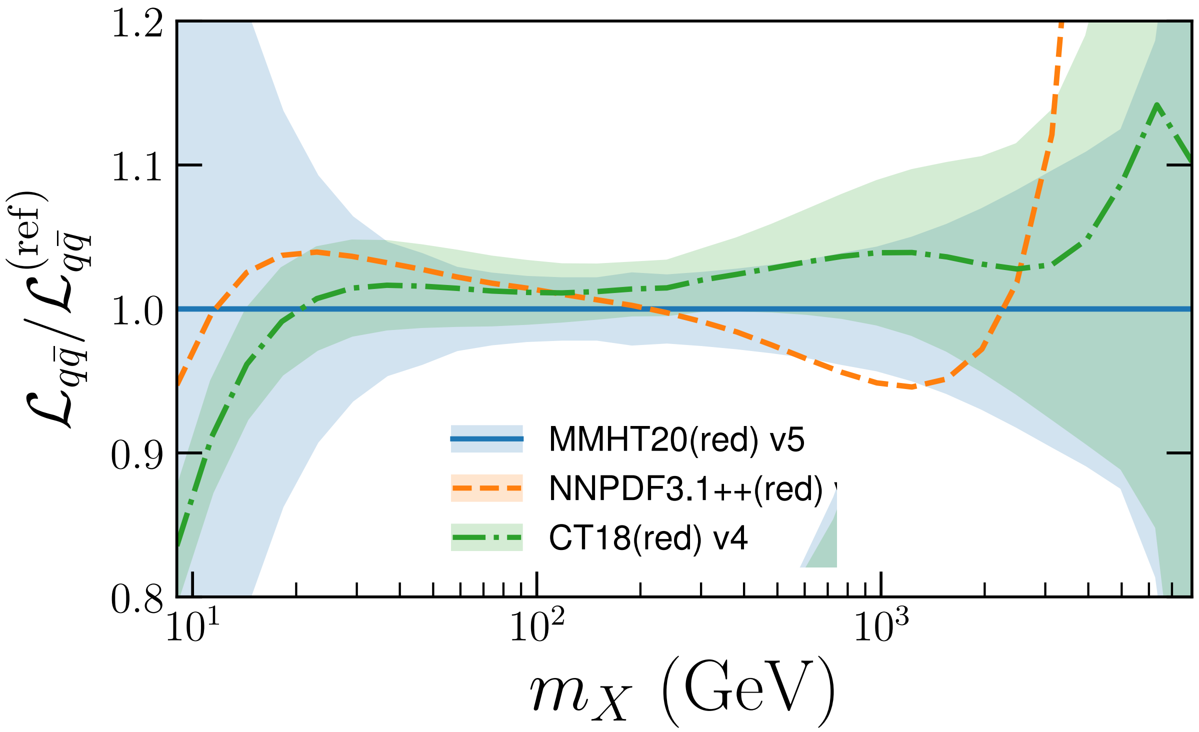
Comparison of benchmark fits

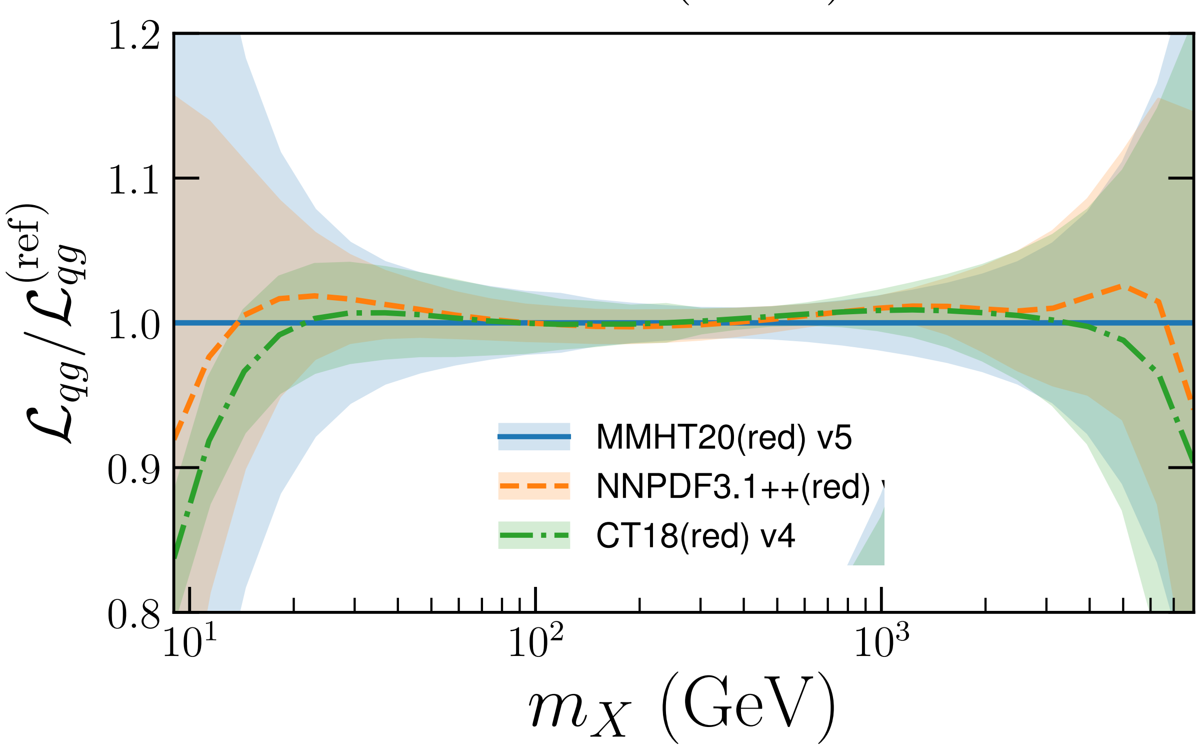
The goal of adopting a common dataset and the same theory settings was to **reduce the potential sources of differences at the PDF level** between the three groups. Is this the case?



Comparison of benchmark fits

The goal of adopting a common dataset and the same theory settings was to **reduce the potential sources of differences at the PDF level** between the three groups. Is this the case?





- **Spot-on agreement** for the gluon-gluon, gluon-quark, and quark-quark luminosity!
- Larger differences for the **quark-antiquark lumi** for m > 600 GeV, but still consistent within errors
- Also differences for strange-induced lumis

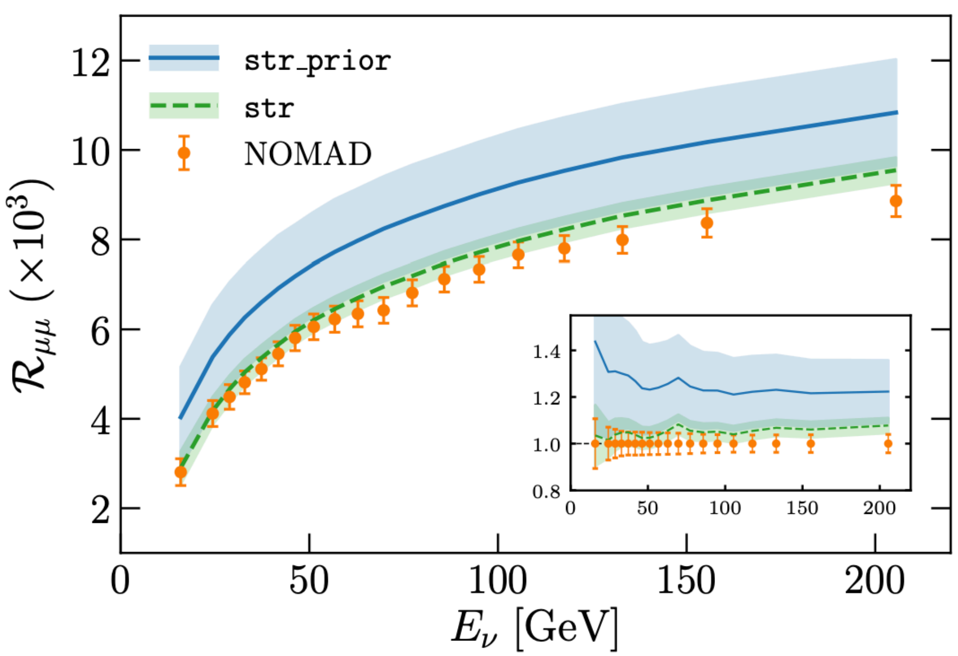
some preliminary conclusions

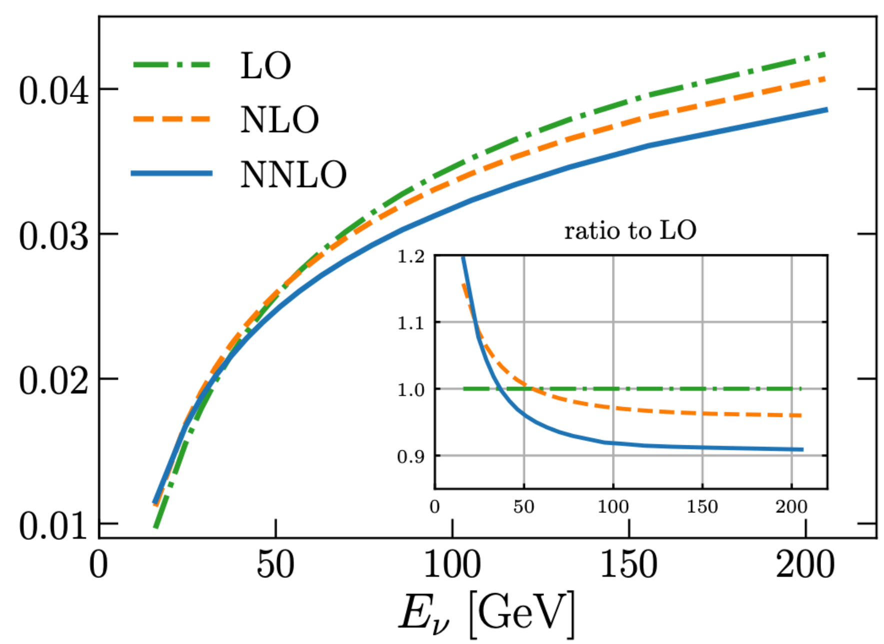
- So far the **benchmark is mostly succesful**: if we use same data and theory, achieve similar PDFs and lumis
- The current point of attention is **quark flavour separation for x>0.01**, but here differences could come eg from parametrisation choices
- Still need to understand better **strangeness**

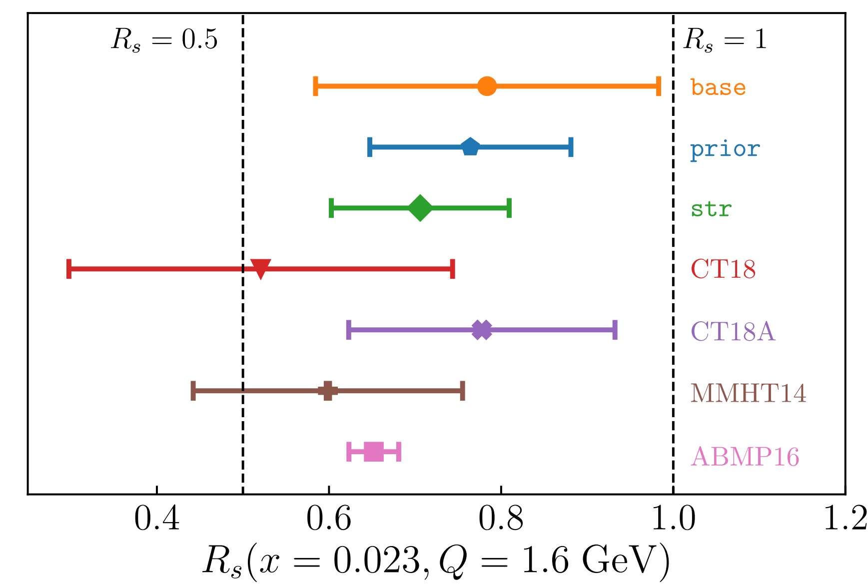
The strangest proton?

PDF4LHC20 benchmark: mostly succesful except for **strangeness** and **quark flavour separation** for x>0.01. Why this can be?

some possibility useful lessons from dedicated study arXiv:2009:00014 (Faura, Iranipour, Nocera, JR, Ubiali)







- Neutrino DIS data (NuTeV & NOMAD) is internally consistent, NNLO QCD corrections important
- **Neutrino DIS and LHC agree on strangeness**
- Agreement for strangeness in global fits seems better than in reduced fits: missing relevant constraints in the latter

Process	Dataset	$n_{\rm dat}$	$\chi^2_{\rm bs}$	$\chi^2_{\rm pr}$	$\chi^2_{\rm str}$
νDIS ($\mu\mu$)		76/76/95	0.76	0.71	0.53
	NuTeV [9]	76/76/76	0.76	0.71	0.53
	NOMAD [10]	—/—/19	[9.3]	[8.8]	0.55
W, Z (incl.)		391/418/418	1.45	1.40	1.40
	ATLAS [12]	34/61/61	1.96	1.65	1.67
W+c		—/37/37	[0.73]	0.68	0.60
	CMS [17, 18]	—/15/15	[1.04]	0.98	0.96
	ATLAS [16]	—/22/22	[0.52]	0.48	0.42
W+jets	ATLAS [15]	—/32/32	[1.58]	1.18	1.18
Total		3981/4077/4096	1.18	1.17	1.17

Validation at the χ^2 level

even for common dataset and theory settings, there **remain differences** between the three groups, *e.g.* chi2 definition, NNLO *K*-factors, heavy quark schemes ….

some of these are well understood and expected, such as heavy quark schemes

Another of the benchmarking tests we are carrying out is comparing the values of the χ^2 which each of the three codes produces when same theory settings and input PDFs are adopted

we should of course compare apples with apples with a common chi2 definition!

$$\chi^2 = \sum_{i,j}^{N_{\rm pt}} (T_i - D_i)({\rm cov}^{-1})_{ij}(T_j - D_j),$$

Experimental definition

$$({\rm cov})_{ij} = \delta_{ij} s_i^2 + \left(\sum_{\alpha=1}^{N_c} \sigma_{i,\alpha}^{(c)} \sigma_{j,\alpha}^{(c)} + \sum_{\alpha=1}^{N_{\mathcal{L}}} \sigma_{i,\alpha}^{(\mathcal{L})} \sigma_{j,\alpha}^{(\mathcal{L})} \right) D_i D_j,$$

T0 definition

$$({\rm cov})_{ij} = \delta_{ij} s_i^2 + \sum_{\alpha=1}^{N_c} \sigma_{i,\alpha}^{(c)} \sigma_{j,\alpha}^{(c)} D_i D_j + \sum_{\alpha=1}^{N_{\mathcal{L}}} \sigma_{i,\alpha}^{(\mathcal{L})} \sigma_{j,\alpha}^{(\mathcal{L})} T_i^{(0)} T_j^{(0)}$$

Extended T0 definition

$$({\rm cov})_{ij} = \delta_{ij} s_i^2 + \left(\sum_{\alpha=1}^{N_c} \sigma_{i,\alpha}^{(c)} \sigma_{j,\alpha}^{(c)} + \sum_{\alpha=1}^{N_{\mathcal{L}}} \sigma_{i,\alpha}^{(\mathcal{L})} \sigma_{j,\alpha}^{(\mathcal{L})} \right) T_i^{(0)} T_j^{(0)}$$

Validation at the χ^2 level

comparison using the **resulting PDFs from the benchmark fits** from each group

ID	Expt.	N_{pt}	χ^2 (CT)	χ^2 (MSHT)	χ^2 (NNPDF)
101	BCDMS F_2^p	$329/163^{\dagger\dagger}/325^{\dagger}$	348.96	163.86	393.06
102	BCDMS F_2^d	$246/151^{\dagger\dagger}/244^{\dagger}$	260.50	133.68	265.35
104	NMC F_2^d/F_2^p	$118/117^{\dagger}$	111.54	109.87	103.38
124+125	NuTeV $\nu\mu\mu + \bar{\nu}\mu\mu$	38+33	62.03	48.46	92.53
160	HERAI+II	1120	1378.87	1344.87	1358.04
203	E866 $\sigma_{pd}/(2\sigma_{pp})$	15	21.35	12.52	5.49
245+250	LHCb 7TeV& 8TeV W,Z	29+30	64.65	70.17	82.57
246	LHCb 8TeV $Z \to ee$	17	23.42	24.91	26.24
248	ATLAS 7TeV W,Z(2016)	34	65.65	66.28	75.54
260	D0 Z rapidity	28	17.51	16.28	17.39
267	CMS 7TeV eletron A_{ch}	11	6.93	17.63	8.22
269	ATLAS 7TeV W,Z(2011)	30	31.35	28.04	29.07
545	CMS 8TeV incl. jet	$185/174^{\dagger\dagger}$	185.40	241.93	232.55
Total	N_{pt}	—	2263	1991	2256
Total	χ^2	—	2584	2278.51	2689.42
			$\chi^2/n = 1.14$	$\chi^2/n = 1.15$	$\chi^2/n = 1.19$

PRELIMINARY

reasonably similar chi2 values at the global level, but note that definitions are still different between each group

Validation at the χ^2 level

comparison using **PDF4LHC15 as input PDF** for the calculations of each group

ID	Expt.	N_{pt}	χ^2 (CT)	χ^2 (MSHT)	χ^2 (NNPDF)
101	BCDMS F_2^p	$329/163^{\dagger\dagger}/325^{\dagger}$	442.07	195.90	479.67
102	BCDMS F_2^d	$246/151^{\dagger\dagger}/244^{\dagger}$	239.36	191.57	298.47
104	NMC F_2^d/F_2^p	$118/117^{\dagger}$	109.11	110.31	108.22
124+125	NuTeV $\nu\mu\mu + \bar{\nu}\mu\mu$	38+33	60.76	35.94	34.04
160	HERAI+II	1120	1421.59	1394.38	1782.61
203	E866 $\sigma_{pd}/(2\sigma_{pp})$	15	6.68	8.15	7.70
245+250	LHCb 7TeV& 8TeV W,Z	29+30	103.04	79.21	154.97
246	LHCb 8TeV $Z \to ee$	17	22.93	28.10	45.63
248	ATLAS 7TeV W,Z(2016)	34	228.16	253.80	241.70
260	D0 Z rapidity	28	17.10	16.12	16.78
267	CMS 7TeV eletron A_{ch}	11	33.18	5.54	7.99
269	ATLAS 7TeV W,Z(2011)	30	35.92	36.91	40.73
545	CMS 8TeV incl. jet	$185/174^{\dagger\dagger}$	282.23	329.00	326.81
Total	N_{pt}	—	2263	1991	2256
Total	χ^2	—	3002.13	2684.92	3545.34

χ^2/n = 1.32 **χ^2/n = 1.34** **χ^2/n = 1.57**

PRELIMINARY

again note that definitions are still different between each group

some marked differences between groups that require further attention: work in progress!

Summary and next steps

- A significant amount of work in the last few months has resulted in several interesting results concerning the **PDF4LHC20 benchmarking exercise**

- Fits based on common datasets and theory settings lead to **very similar** gluon, singlet, charm PDFs and sea quarks for $x < 0.01$. Some differences remain for strangeness and quark flavour separation for $x > 0.01$. **PDF luminosities** in good agreement, except q=qbar

- Work in progress aims to pin down the origin of the residual differences by means of comparing **point-by-point cross-sections** and **chi2s obtained with common definitions**

- The results of this benchmarking exercise will provide useful information concerning whether or not we can (or should) proceed with a **PDF4LHC20 PDF combination**

Many thanks for all the participants of this PDF4LHC20 benchmarking exercise, and specially to Thomas Cridge, Tim Hobbs, and Emanuele Nocera for providing many of the results presented in this talk!