# PDFs for Higgs Physics at the HL-LHC

Lucian Harland-Lang, University of Oxford

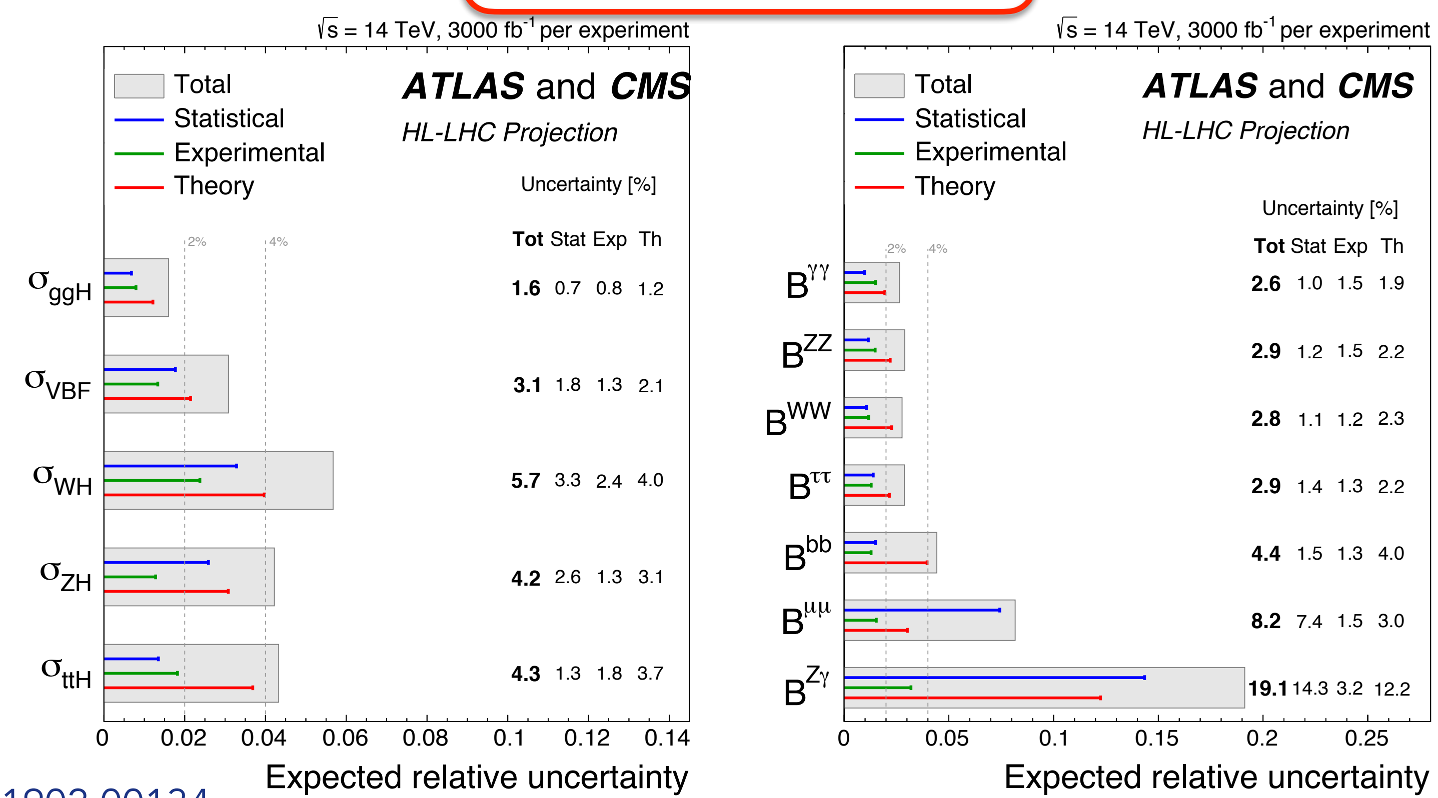
Higgs 2020, 28 October 2020

# Higgs Physics at HL-LHC

- **LHC/HL-LHC**: exciting opportunity to pin down Higgs sector.
- Stress testing SM Higgs hypothesis through high precision data + theory could be route to discovery.

▸ HL-LHC projections ~20 years from now!

(S2) TH uncertainties scaled down by factor 2, EXP scaled according to $\sqrt{\mathscr{L}}$



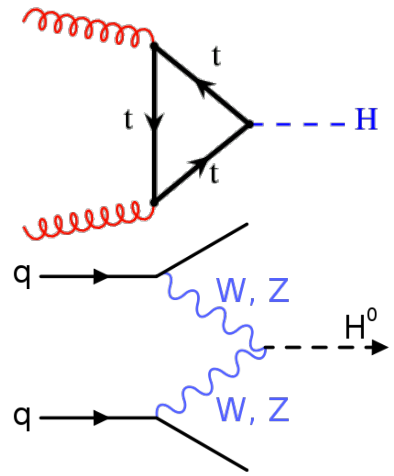

1902.00134

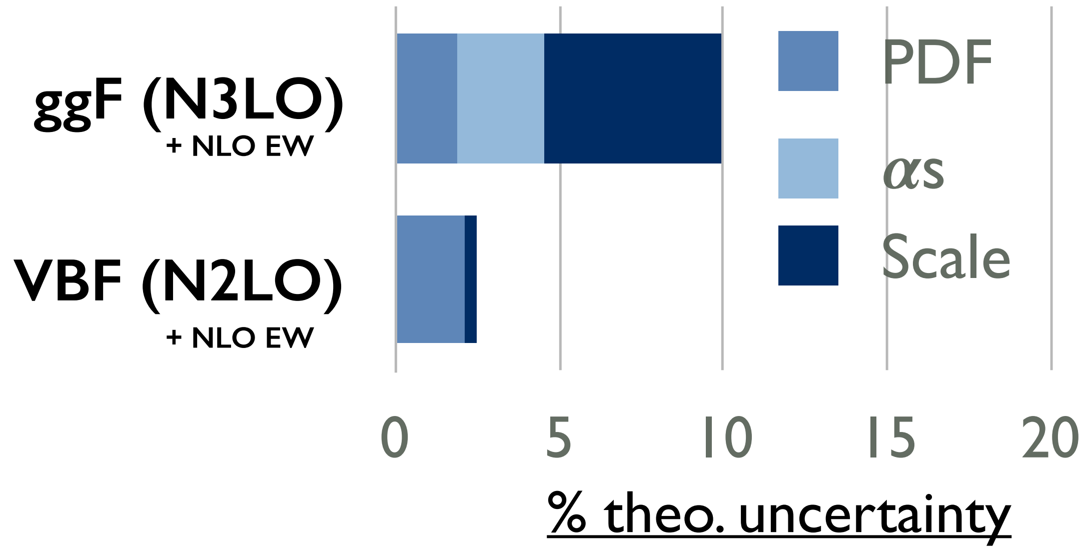
D. de Florian's talk

- Key ingredient in this (and indeed broader LHC precision programme): precise understanding of **PDFs** and their uncertainties/biases.

# Higgs and PDFs

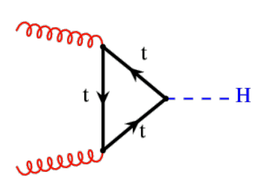
M. Ubiali, Higgs Couplings 2019

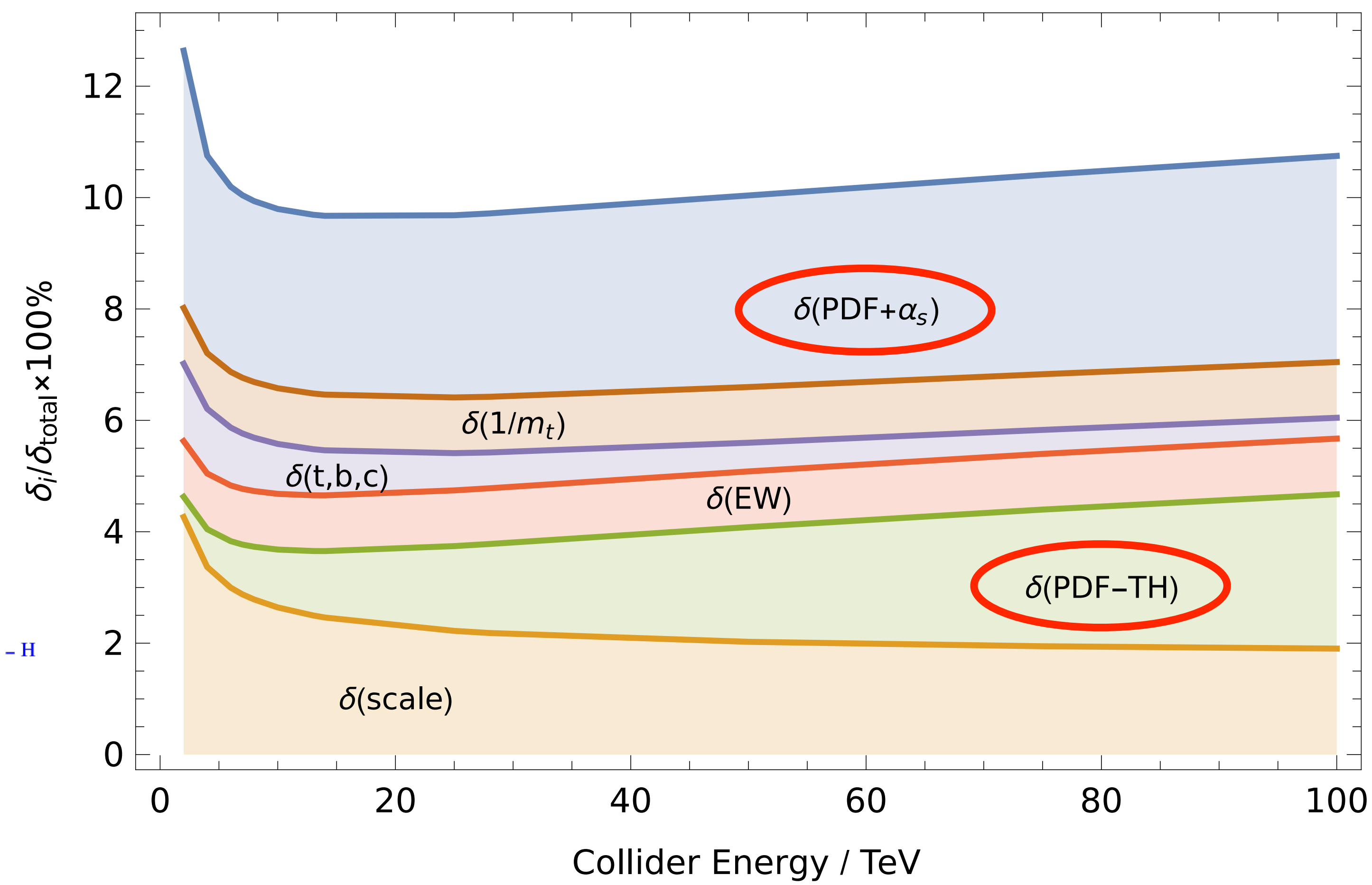



| σ@13 TeV |
|---|
| **48.6 pb** |
| **3.77 pb** |

★ PDF uncertainty important limiting factor in Higgs precision programme.



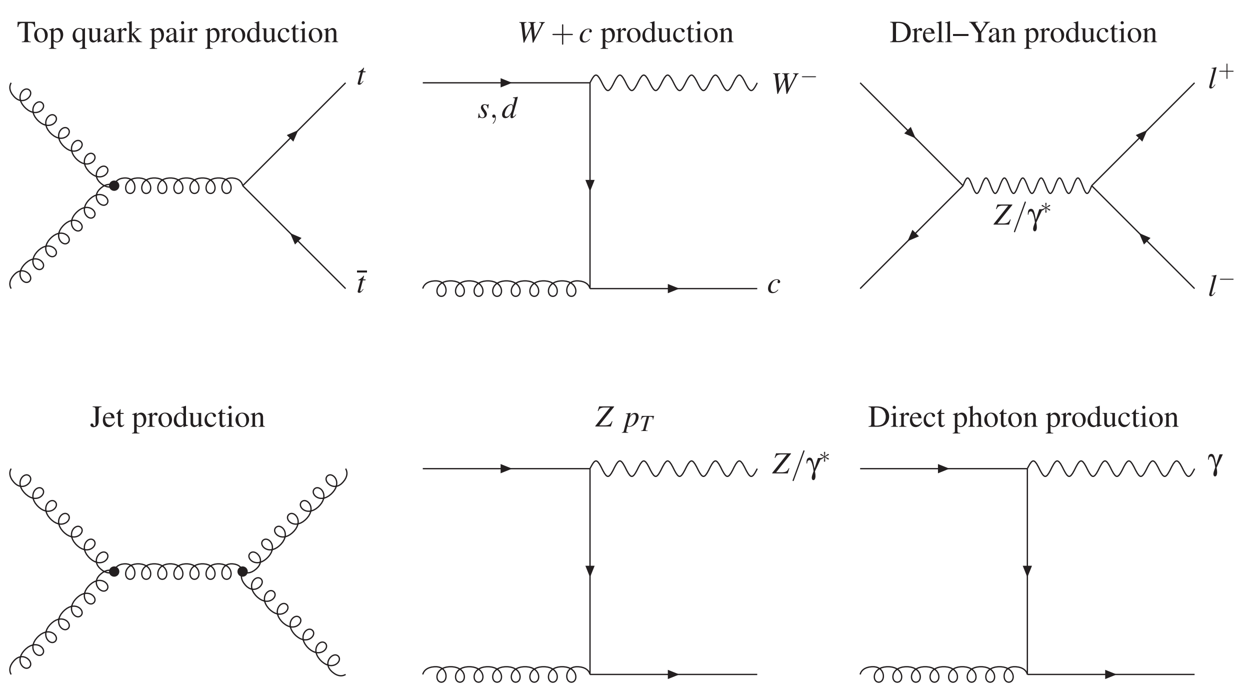




M. Cepeda et al., 1902.00134

# HL-LHC + PDFs: what are we aiming for?

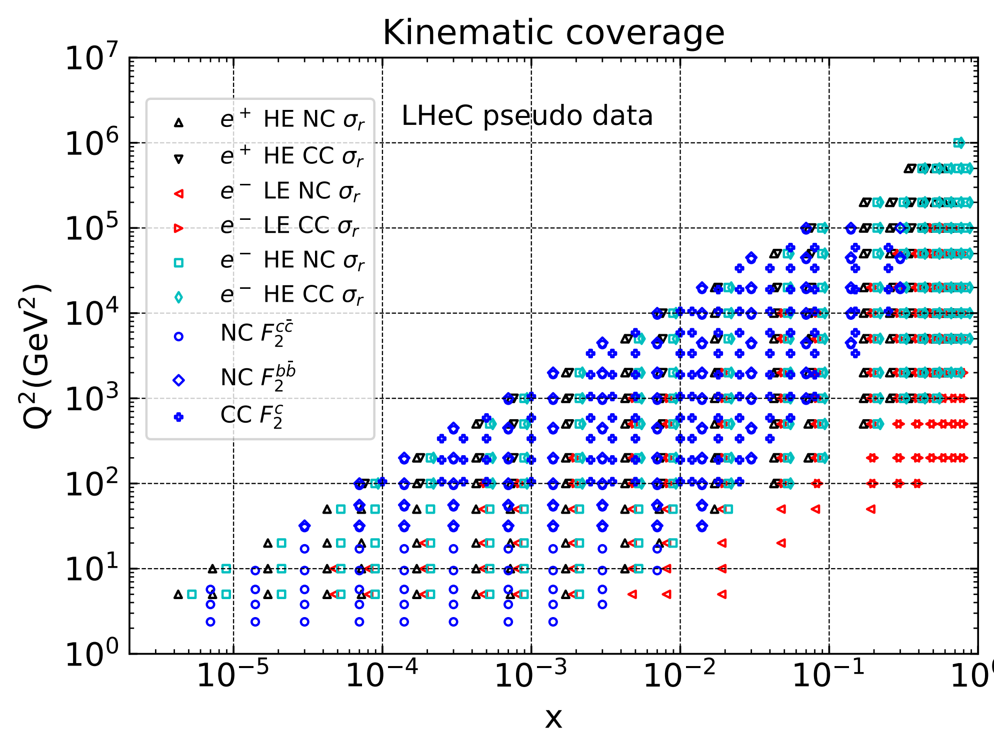
- What can we expect for impact of final **HL-LHC** data on **PDFs**?
- Studies based on straightforward extrapolations of statistical errors and estimates of improvements in systematics.
- Datasets considered non-exhaustive and some allowance for tensions built in.

**R. Abdul Khalek, S. Bailey, J. Gao, LHL, J. Rojo. Eur.Phys.J. C78 (2018) no.11, 962**

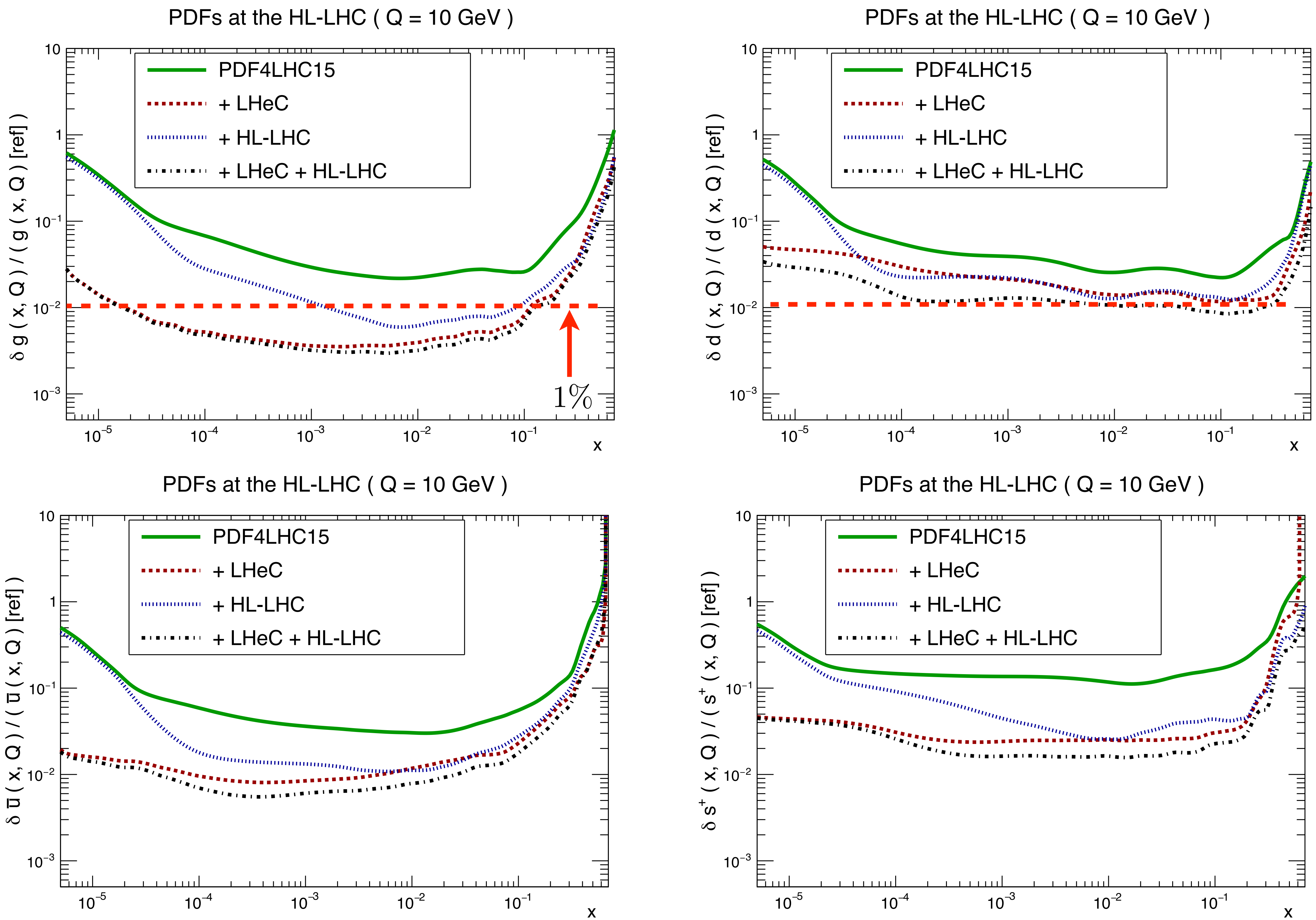



**HL-LHC**

**R. Abdul Khalek, S. Bailey, J. Gao, LHL, J. Rojo. SciPost Phys. 7, 051 (2019)**



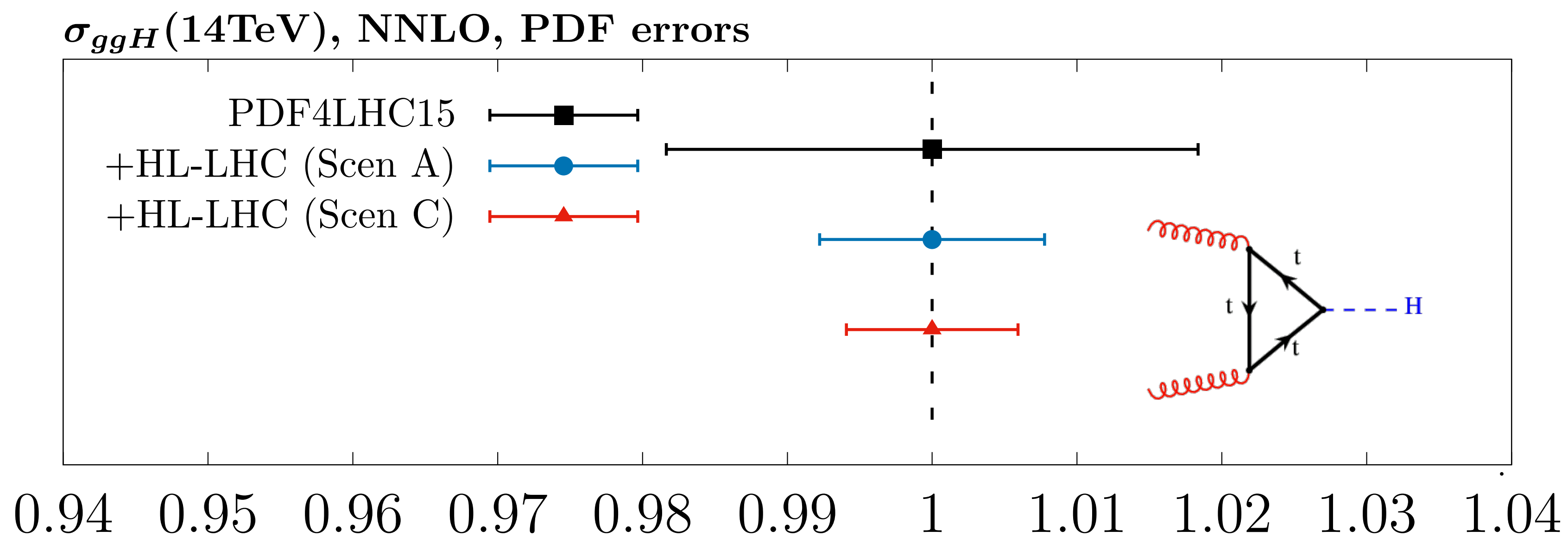

**LHeC**

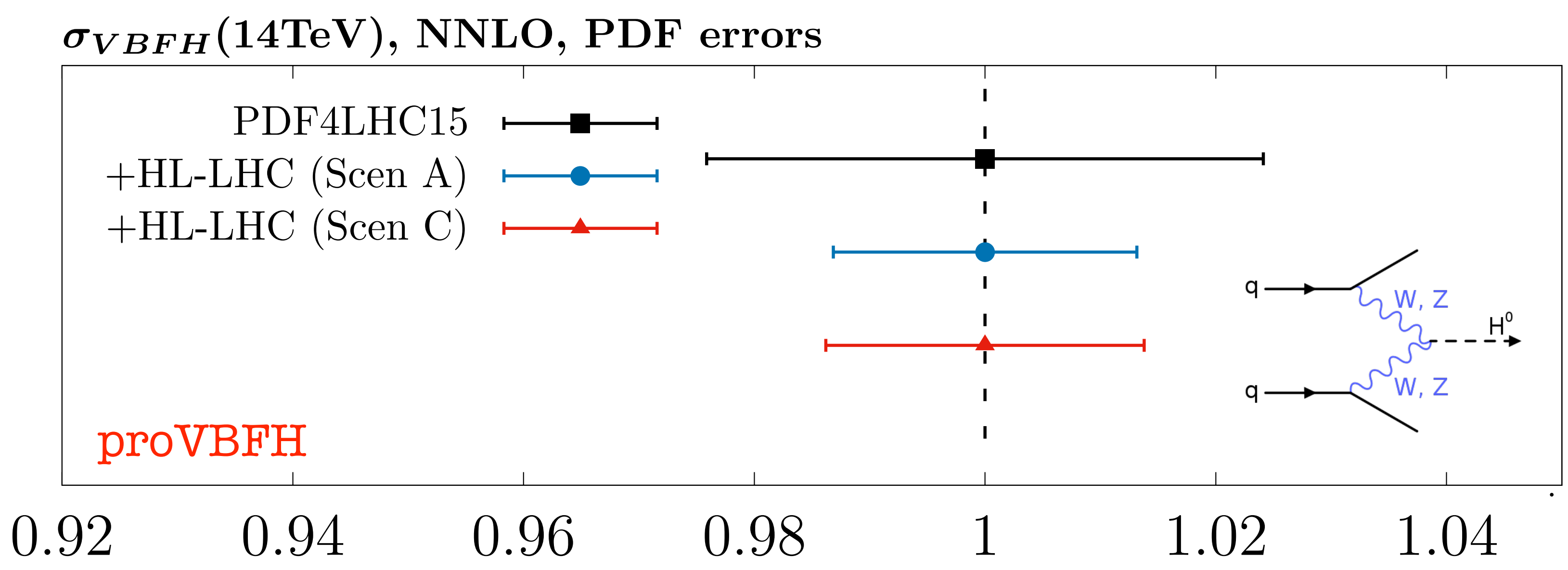



- **Sub percent level** uncertainty in e.g. gluon in some $x$ regions.
- Implication for Higgs physics?

- **Significant improvement** seen in ggH: impact of gluon sensitive LHC data. PDF uncertainty below ~ 1% level.
- **Good improvement** in VBF. Focus of this study on currently stats limited data, further improvement certainly possible here from e.g. precision DY.
- Note baseline is PDFLHC15: does not correspond to latest sets.

**Different scenarios = different projected systematics.**



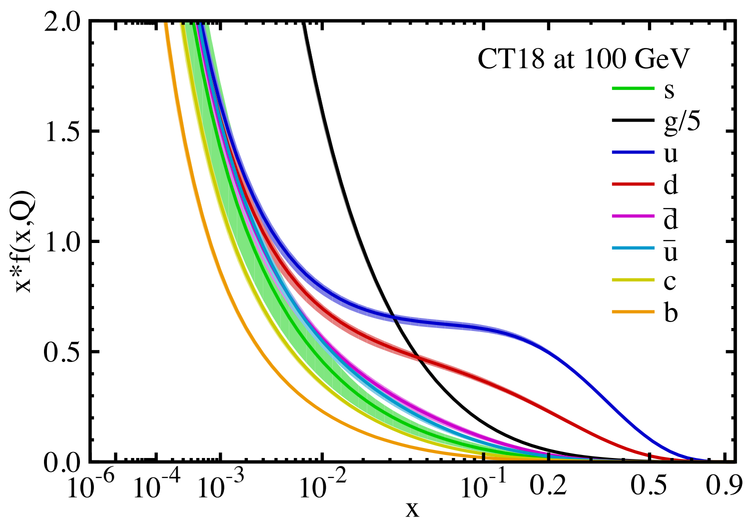
# PDF Fits Today

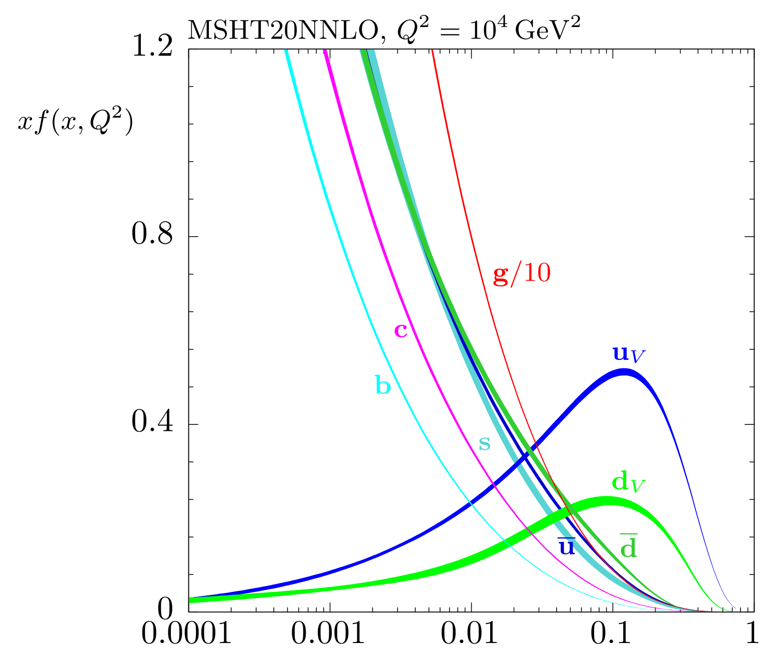
# Latest Releases

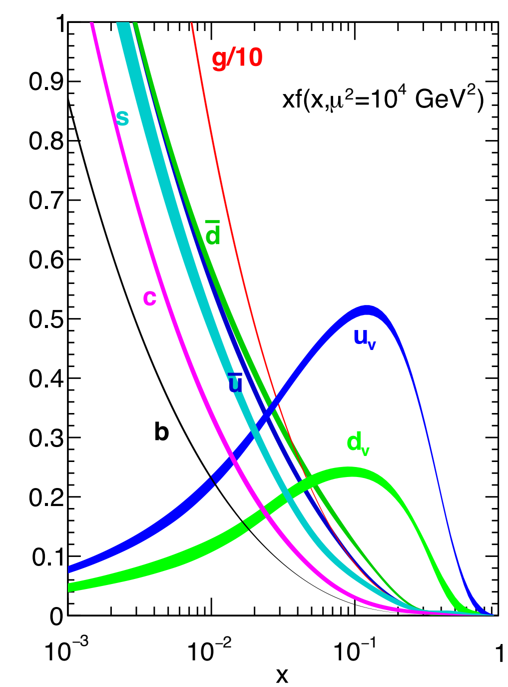
- **Projection studies**: (sub) % level PDF uncertainty achievable by end of HL-LHC. Can we get there? Where are we are now?
- ‘Post-Run I’ sets now exist from three major global fitters:

**CT18** **MSHT20*** **NNPDF3.1**

- In all cases, focus on including significant amount of **new data**, higher **precision theory** and on **methodological improvements**.




- LHC data also playing key role in **ABM** fits, while ATLAS/CMS continue to provide their own PDF analyses.

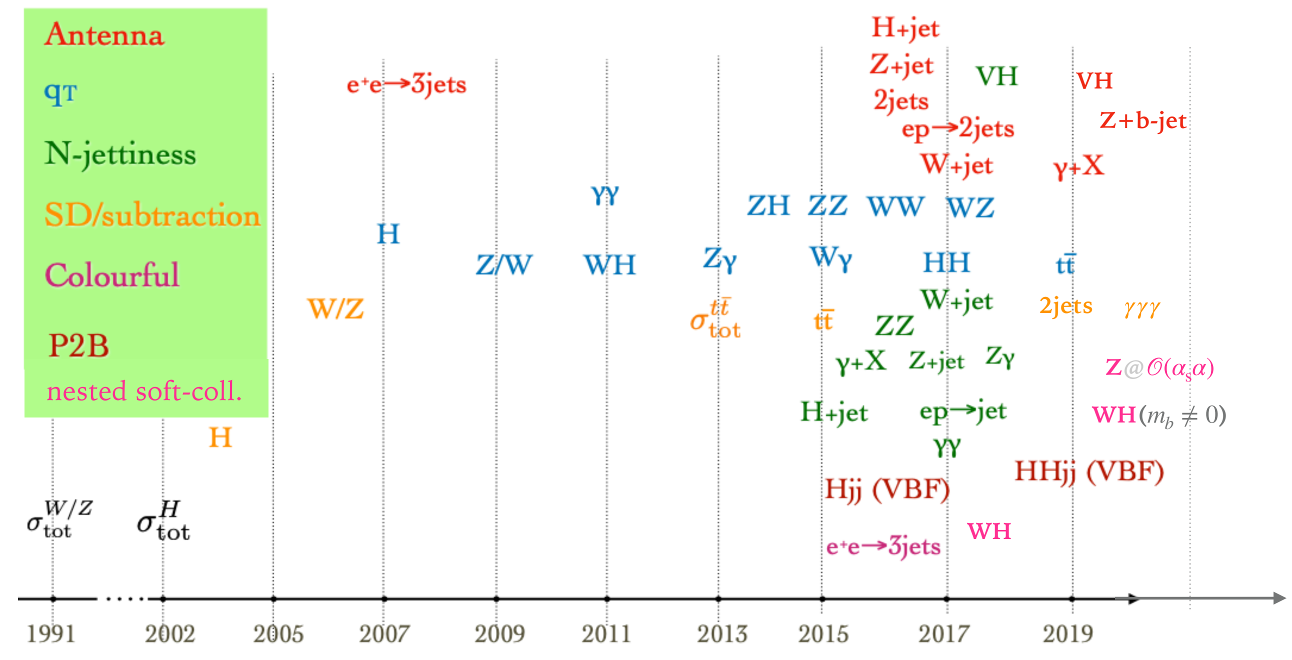


*Paper in prep., PDFs available on request.

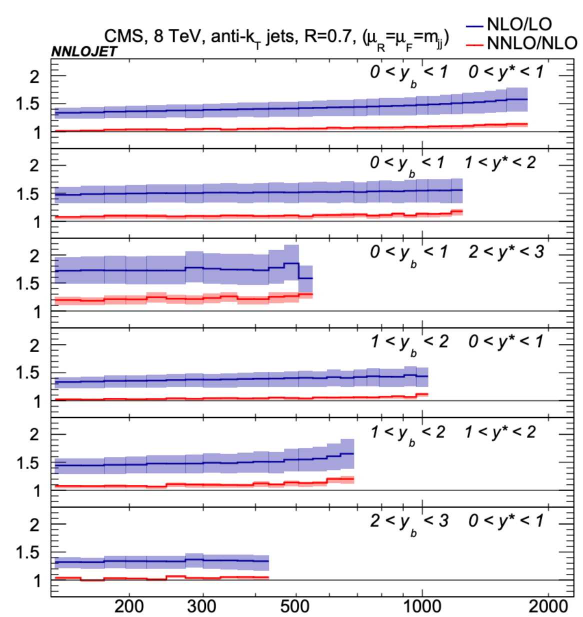
# Precision Theory

A. Huss, QCD@LHC20

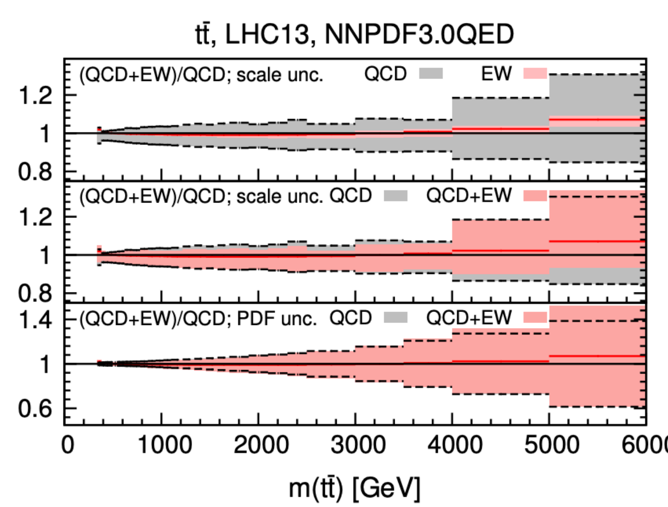
- Vast majority of processes included in fits have full **NNLO** QCD theory (+ **NLO** EW where relevant) available and included.



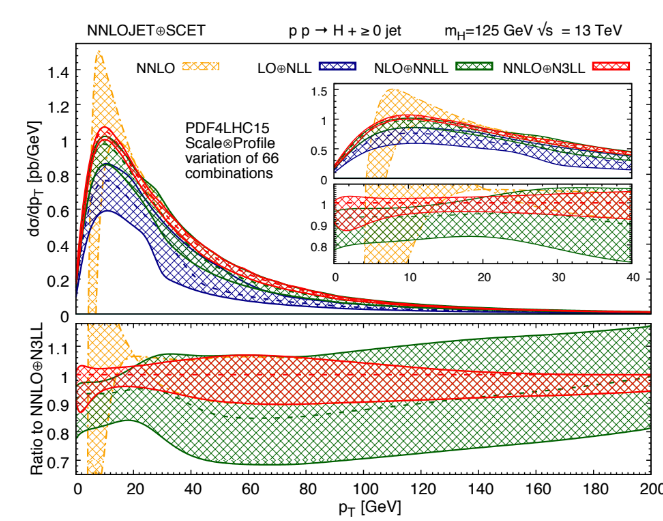

[based on slide by M. Grazzini; QCD@LHC 2019]




Inclusive jets/dijets



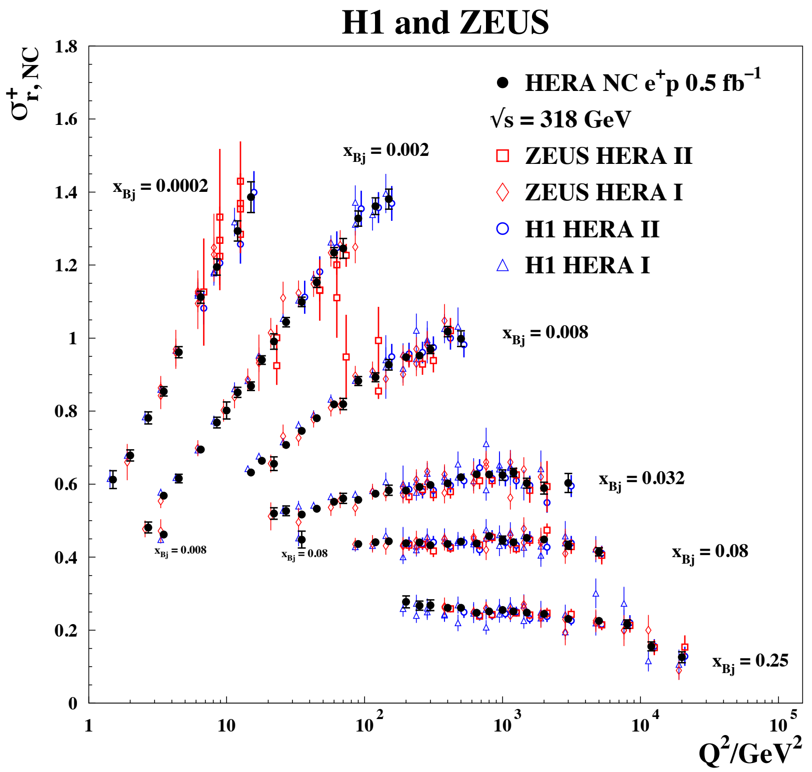

Top quarks - single/ double differential

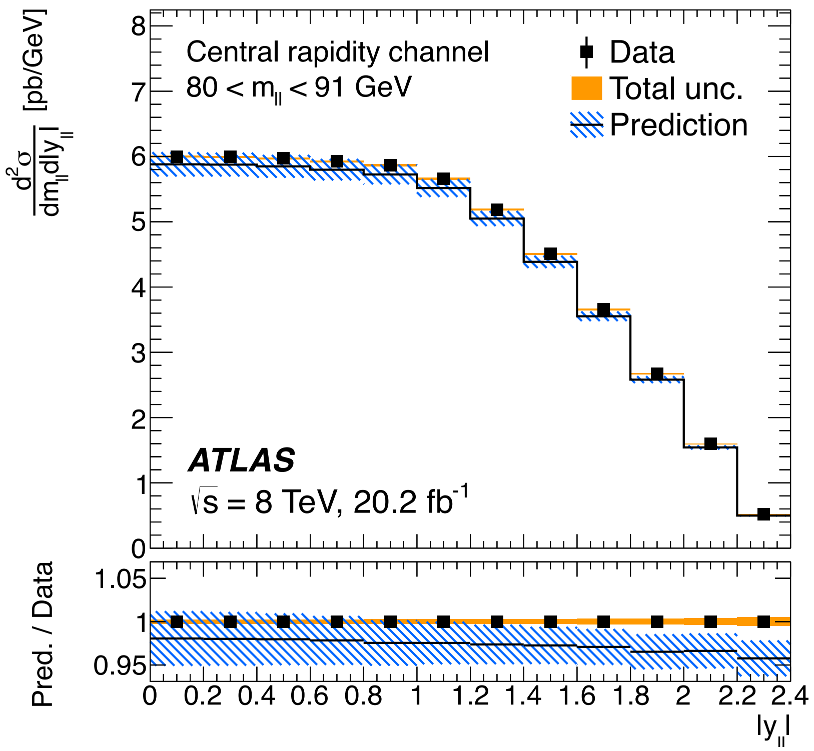


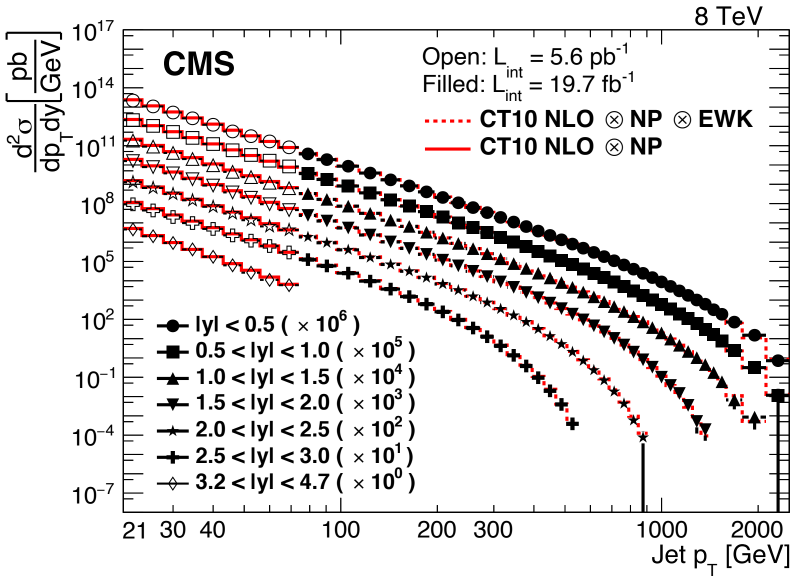

W, Z transverse momentum distributions

# New Data

- Can divide into 3 broad (non-exhaustive) categories:




★ Final HERA H1 + ZEUS combination data on inclusive and heavy flavour **DIS**.

★ High precision multi-differential **DY** data. Flavour decomposition.

★ Inclusive jet, $Z\,p_\perp$, differential $t\bar{t}$. High/intermediate $x$ partons.

- Impact of these datasets has been analysed in detail in many individual studies. No time to discuss here, but see **S. Amoroso's talk** for nice overview.

See backup for a couple of recent examples

# New Collider Data

**MSHT20: new collider data**

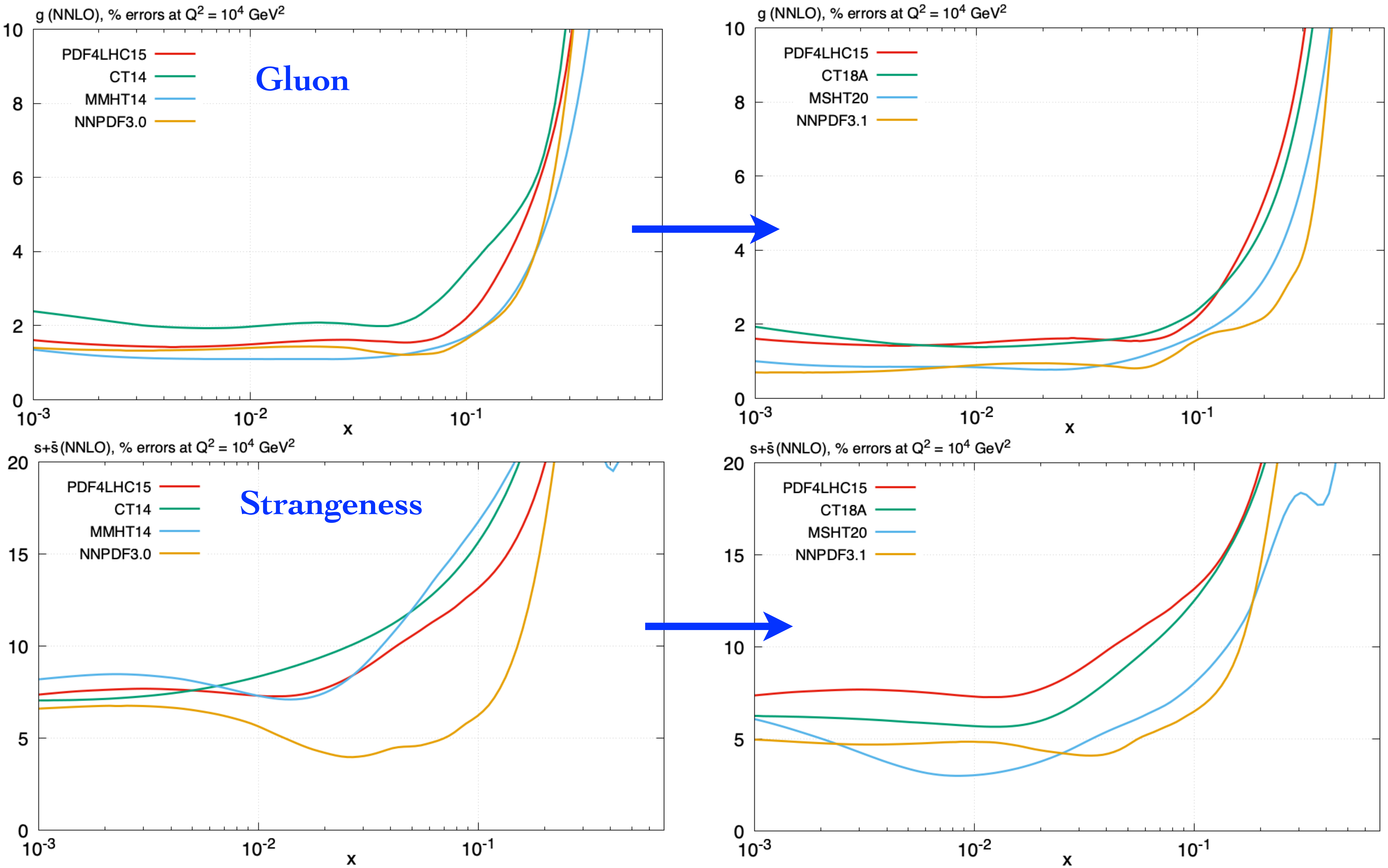
**MSHT20 fit (MMHT14 prediction)**

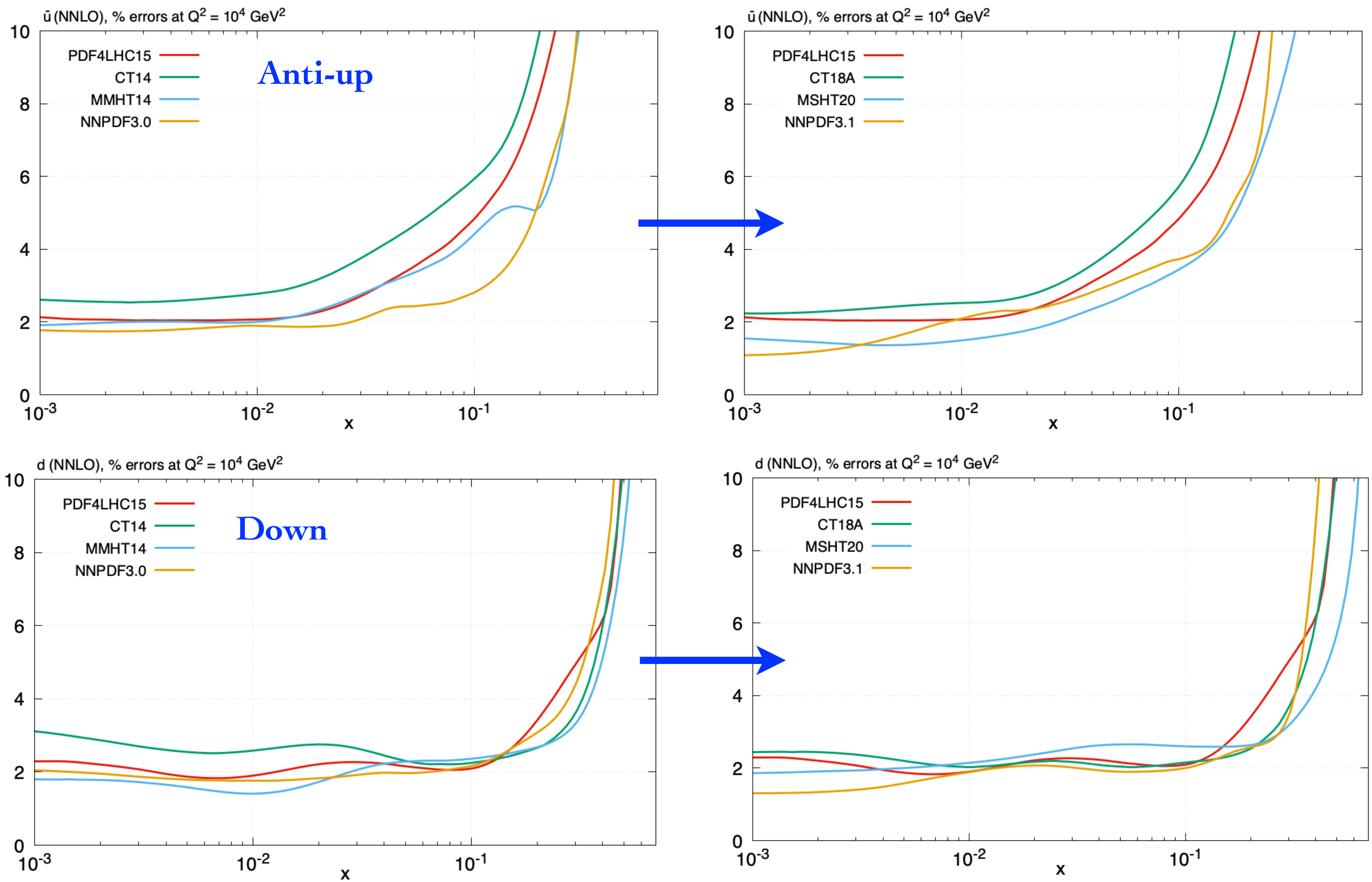
| | Points | NLO $\chi^2/N_{pts}$ | NNLO $\chi^2/N_{pts}$ |
|---|---|---|---|
| DØ $W$ asymmetry | 14 | 0.94 (2.53) | 0.86 (14.7) |
| $\sigma_{t\bar{t}}$ [69]- [70] | 18 | 1.27 (1.31) | 0.81 (0.83) |
| LHCb 7+8 TeV $W+Z$ [71, 72] | 67 | 1.71 (2.35) | 1.48 (1.55) |
| LHCb 8 TeV $Z \to ee$ [73] | 17 | 2.29 (2.89) | 1.54 (1.78) |
| CMS 8 TeV $W$ [74] | 22 | 1.05 (1.79) | 0.58 (1.30) |
| CMS 7 TeV $W+c$ [75] | 10 | 0.82 (0.85) | 0.86 (0.84) |
| ATLAS 7 TeV jets $R=0.6$ [18] | 140 | 1.62 (1.59) | 1.59 (1.68) |
| ATLAS 7 TeV $W+Z$ [20] | 61 | 5.00 (7.62) | 1.91 (5.58) |
| CMS 7 TeV jets $R=0.7$ [76] | 158 | 1.27 (1.32) | 1.11 (1.17) |
| ATLAS 8 TeV $Zp_T$ [54] | 104 | 2.26 (2.31) | 1.81 (1.59) |
| CMS 8 TeV jets $R=0.7$ [77] | 174 | 1.64 (1.73) | 1.50 (1.59) |
| ATLAS 8 TeV $t\bar{t} \to l+j$ sd [78] | 25 | 1.56 (1.50) | 1.02 (1.14) |
| ATLAS 8 TeV $t\bar{t} \to l^+l^-$ sd [79] | 5 | 0.94 (0.82) | 0.68 (1.10) |
| ATLAS 8 TeV high-mass DY [52] | 48 | 1.79 (1.99) | 1.18 (1.26) |
| ATLAS 8 TeV $W^+W^-$+ jets [80] | 25 | 1.36 (1.36) | 0.72 (0.69) |
| CMS 8 TeV $(d\sigma_{\bar{t}t}/dp_{T,t}dy_t)/\sigma_{\bar{t}t}$ [81] | 15 | 2.19 (2.20) | 1.50 (1.47) |
| ATLAS 8 TeV $W^+W^-$ [82] | 22 | 3.85 (13.9) | 2.61 (5.25) |
| CMS 2.76 TeV jets [83] | 81 | 1.53 (1.59) | 1.27 (1.39) |
| CMS 8 TeV $\sigma_{\bar{t}t}/dy_t$ [84] | 9 | 1.43 (1.02) | 1.47 (2.14) |
| ATLAS 8 TeV double differential $Z$ [53] | 59 | 2.67 (3.26) | 1.45 (5.16) |

- Impact of data on fit clear via **MMHT14/MSHT20** difference.
- With the addition of newer higher precision LHC data, the necessity of NNLO becoming increasingly clear.

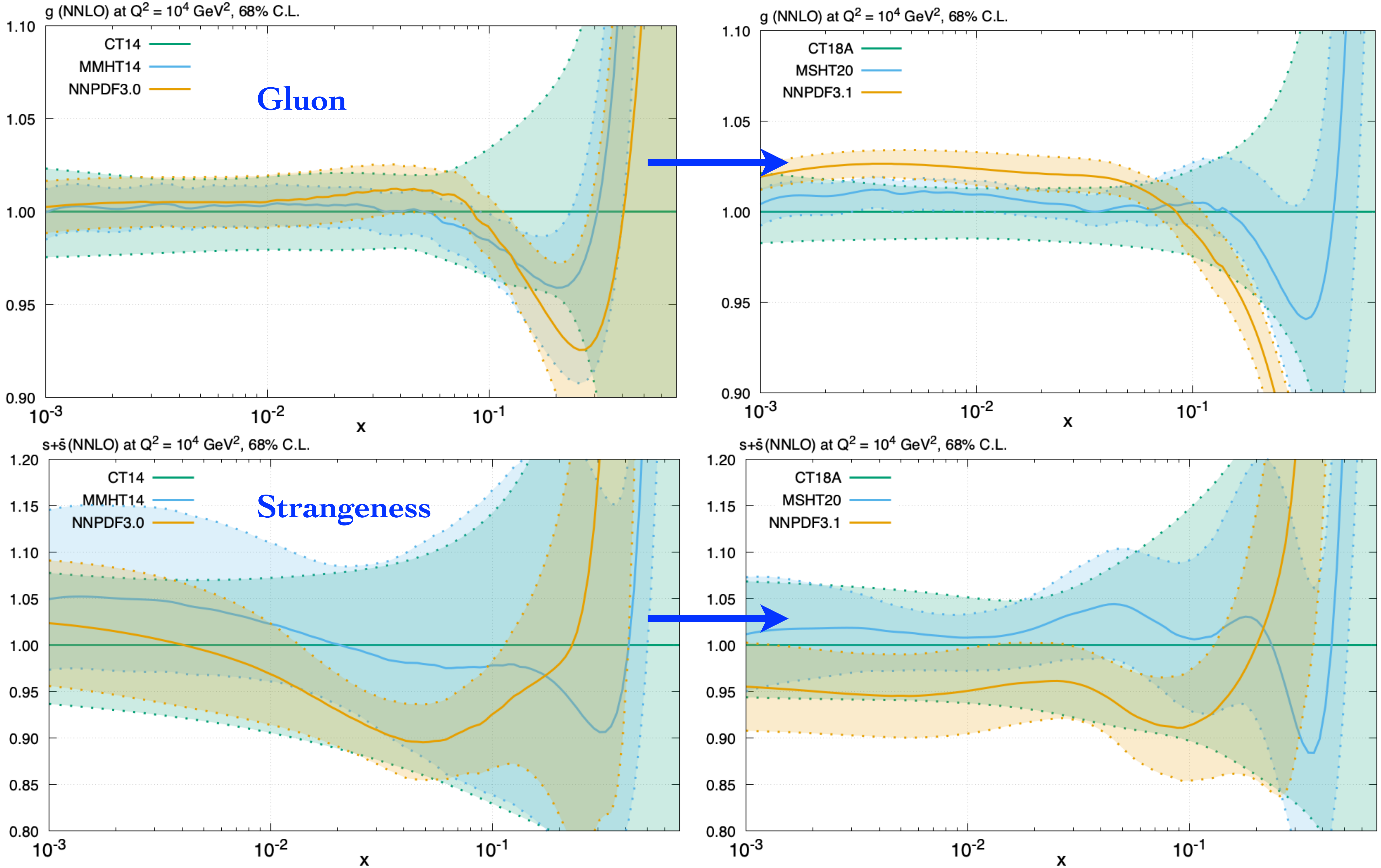
# What do the PDFs look like?

- For most recent public fits, in general clear reduction in **individual errors**. Driven by greatly increased datasets, in particular from **LHC**.

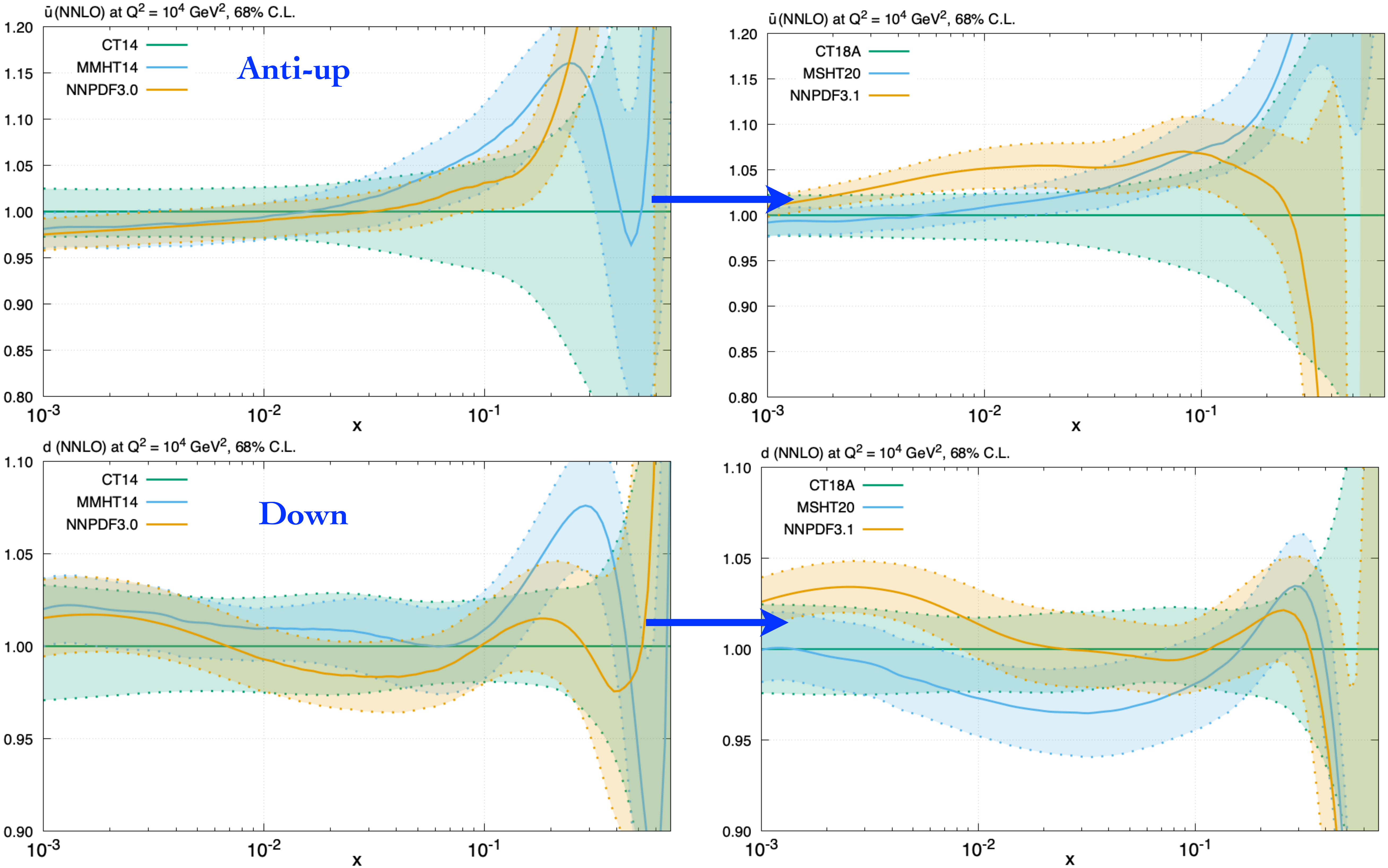





- However picture less clear when looking at central values….

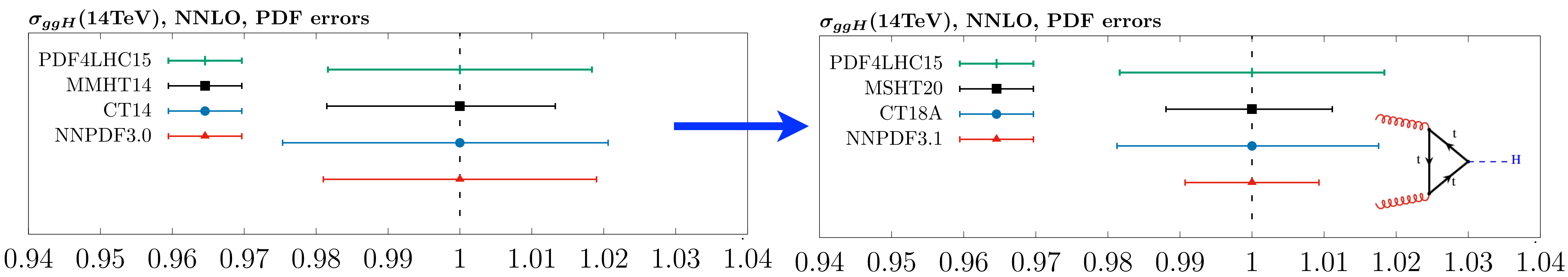


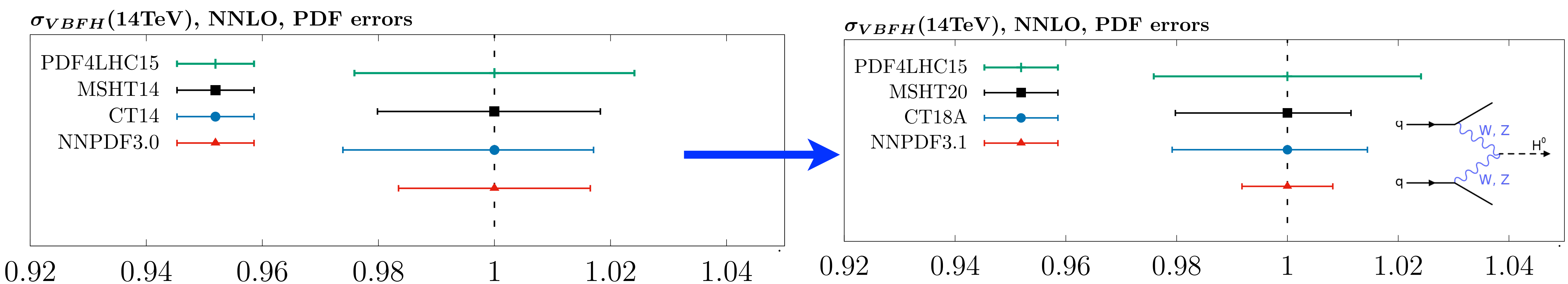

- Encouraging picture for strangeness, but spread in gluon not necessarily reduced/has even increased!



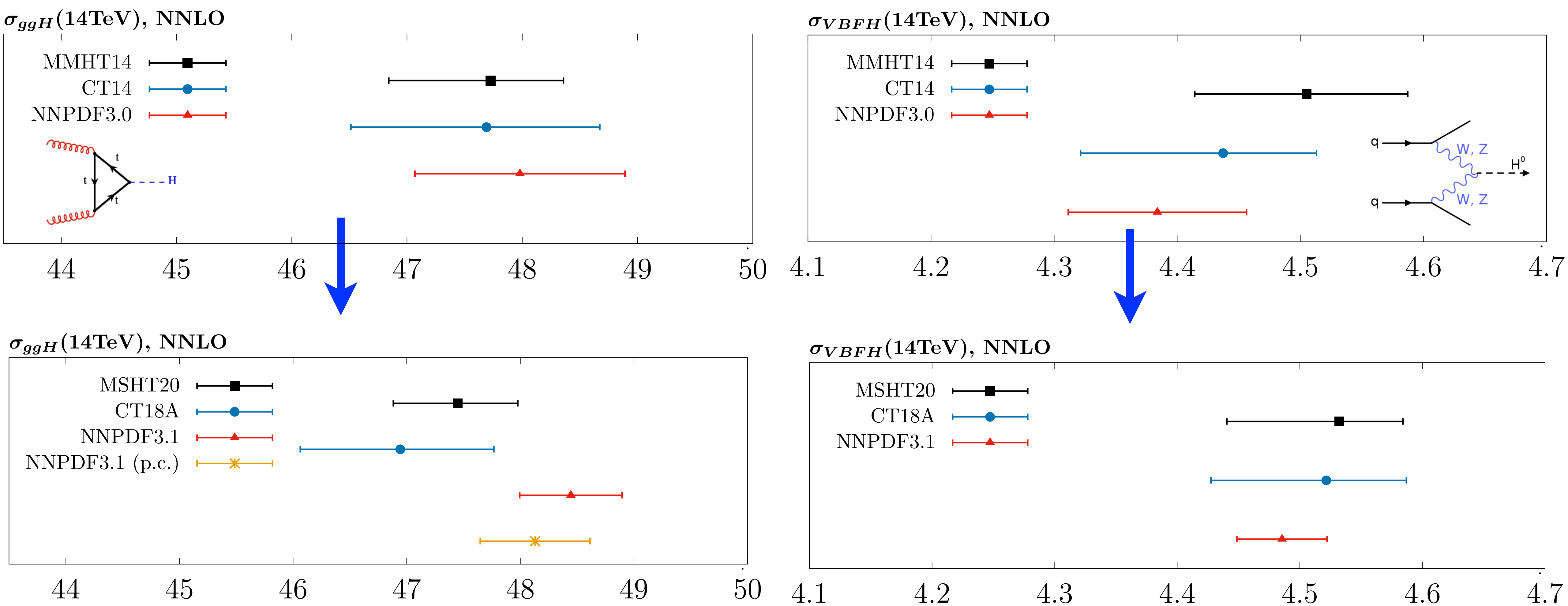

- Similarly mixed picture for light quarks.
- Impact on **Higgs physics**?

- **Some differences** in PDF uncertainties, in particular for VBF.
- NNPDF3.1 uncertainty lowest, though note LHC data entering fit is ~ comparable to CT18, and less than MSHT20.
- Broadly: **encouraging improvement** in PDF uncertainties clear, driven by new LHC data in fits.
- What about absolute values?



- **Mixed picture**: for ggH, as with gluon, the spread has increased.
- Fitted vs. pert. charm in NNPDF plays some but not entire role in this. Appears to relate to impact of newer LHC gluon sensitive data being different in various fits.
- For VBF picture is opposite: spread has decreased, as well as individual uncertainties! LHC precision W,Z will play role in this.




- Clearly cannot be end of story: further work + **benchmarking** needed.

# Benchmarking: Towards PDF4LHC20

- **Benchmarking exercise** is underway to understand as much as possible these differences between CT18, MSHT20 and NNPDF3.1. Work towards eventual "**PDF4LHC20**" combination of these.
- Many questions to get a handle on:

**See also S. Amoroso's talk**

  - ★ Tensions between datasets in fits. Different implications/interpretations of these by groups.
  - ★ Variations in choices of distributions that are fit, and their treatment.
  - ★ Differences (and possible issues/bugs!) in theory treatments.
  - ★ …

Nikhef

VU VRIJE UNIVERSITEIT AMSTERDAM

**Update on the PDF4LHC20 benchmarking exercise**

**Juan Rojo**
VU Amsterdam & Theory group, Nikhef

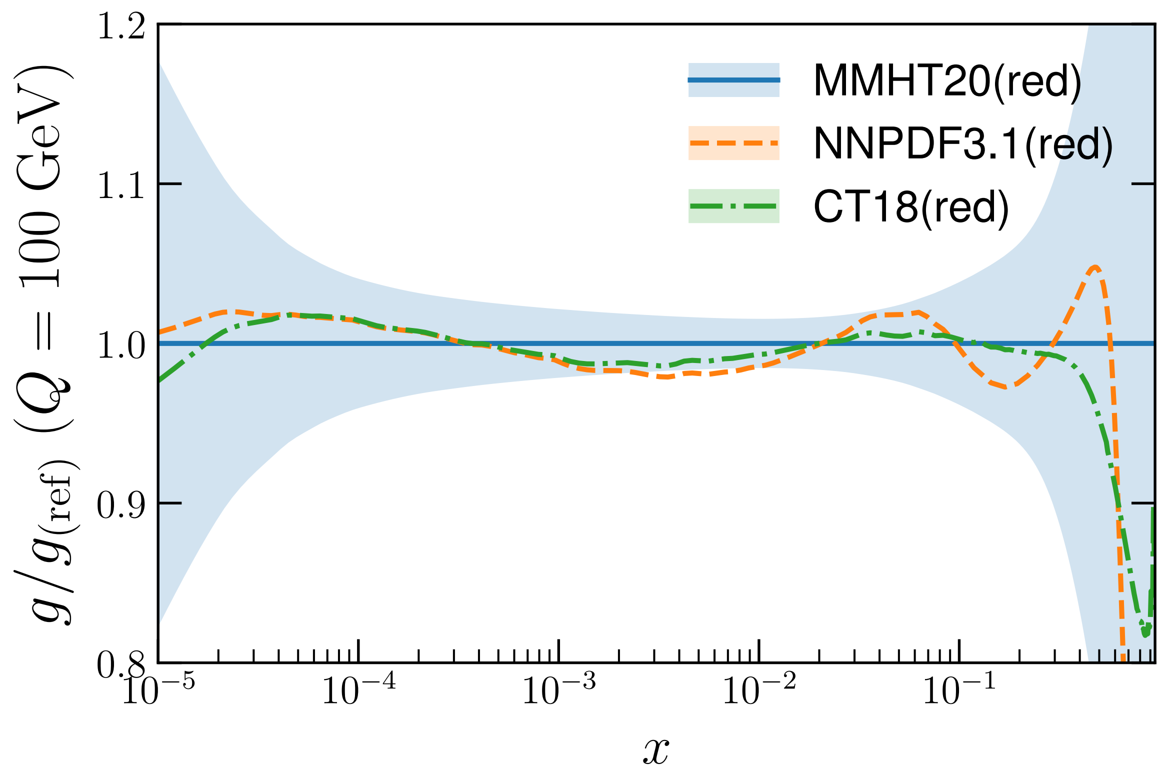
**11/08/2020, PDF4LHC meeting**

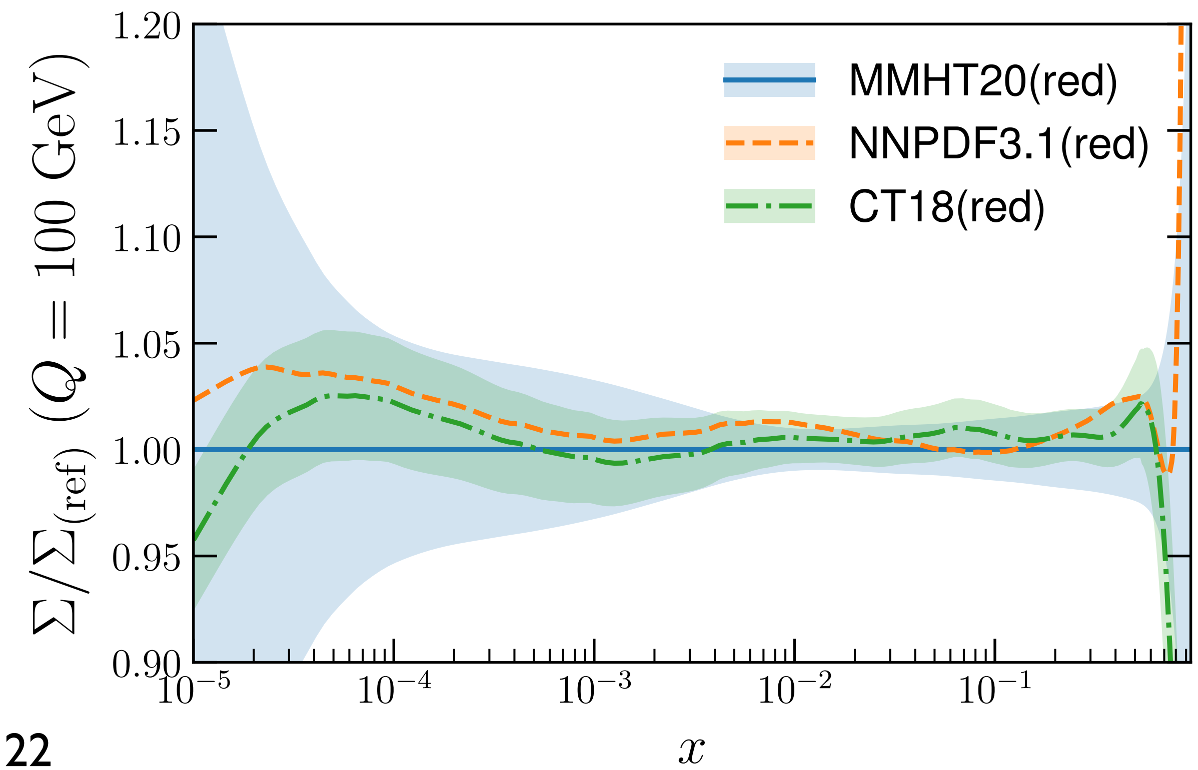
https://indico.cern.ch/event/955077/contributions/4017826/attachments/2114481/3557330/rojo-PDF4LHC20-bench.pdf

- First step: fit to **reduced dataset** that still constrains PDFs sufficiently and includes some new LHC data.
- Similar theory settings for now.

- **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

- Perform comparison, understand differences and then broaden scope.
  - ★ First results encouraging, but with some differences to be resolved. **Work in progress**.
- Proposal to evaluate correlations between PDF sets also under discussion.



# Missing higher orders + N3LO

# Missing Higher Orders

- On theory side, a consideration of the uncertainty due to the use of fixed order pQCD in PDF fits needed, i.e. the missing higher order (MHO) theory uncertainty.

$\longrightarrow$ Essential to include measure of this if we are to have reasonable/viable interpretation of **fit quality** at high precision, in particular if default poor. Without this may be biasing fit.

$\longrightarrow$ Additional motivation, to give estimation of **uncertainty** in extracted PDFs due to MHOs in fit.

**R. Abdul Khalek et al., *Eur.Phys.J.C* 79 (2019) 11, 931**

- Detailed NNPDF study: include these via scale variations.
- However: we already include MHO uncertainty by scale variation when predicting observables with PDFs. Risk of **double counting**?

**LHL and R. S. Thorne, EPJC79 (2019), no.1, 39**

- Beyond this: are scale variations even the appropriate thing?

**M. Bonvini, arXiv:2006.16293**

# Higgs and MHOs

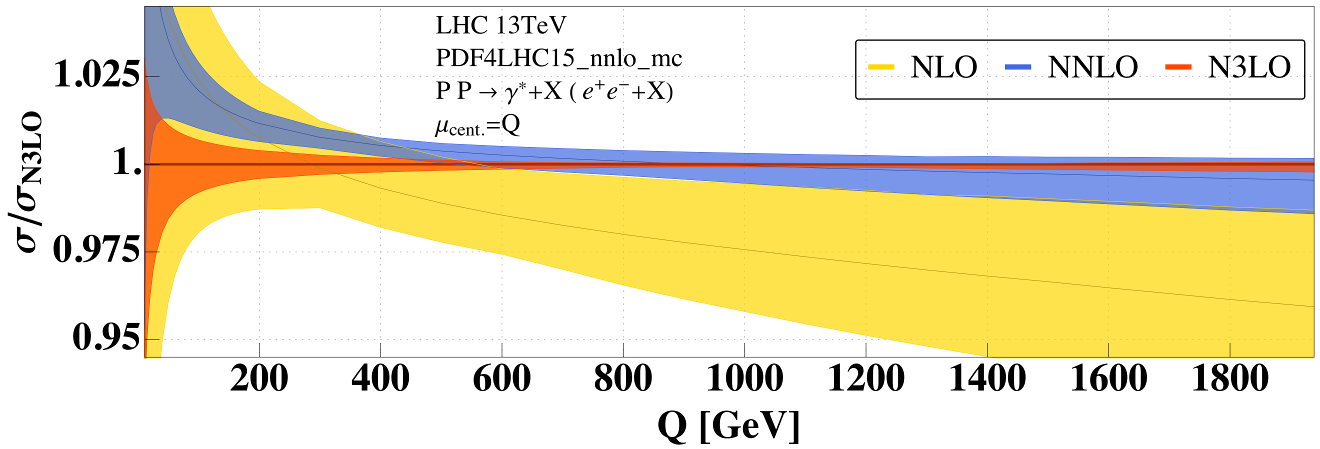
- N3LO calculations available for Higgs in gluon fusion and VBF, but no PDFs to match this level.
- Source of uncertainty:

**VBF:** $\delta_A^{\rm PDF} = \frac{1}{2}\left|\frac{\sigma^{\rm NNLO}_{\rm NNLO\text{-}PDF} - \sigma^{\rm NNLO}_{\rm NLO\text{-}PDF}}{\sigma^{\rm NNLO}_{\rm NNLO\text{-}PDF}}\right| = 1.1\%$ **ggH:** $\delta({\rm PDF-TH}) = \frac{1}{2}\left|\frac{\sigma^{(2),NNLO}_{EFT} - \sigma^{(2),NLO}_{EFT}}{\sigma^{(2),NNLO}_{EFT}}\right| = \frac{1}{2}\,2.31\% = 1.16\%$

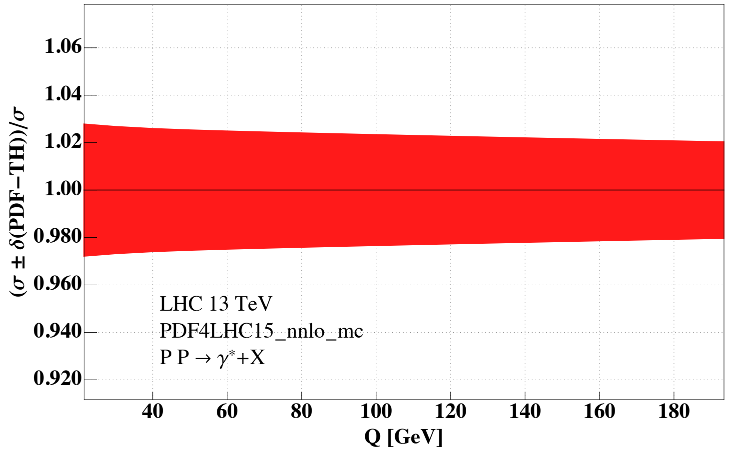
F. Dreyer and A. Karlberg, PRL 117 (2016), no.7, 072001

C. Anastasiou et al., JHEP 05 (2016) 058

- Also relevant for recent calculation of $\gamma, W^{\pm}$ at N3LO in QCD.
- In some regions scale variation bands do not overlap between NNLO/N3LO.

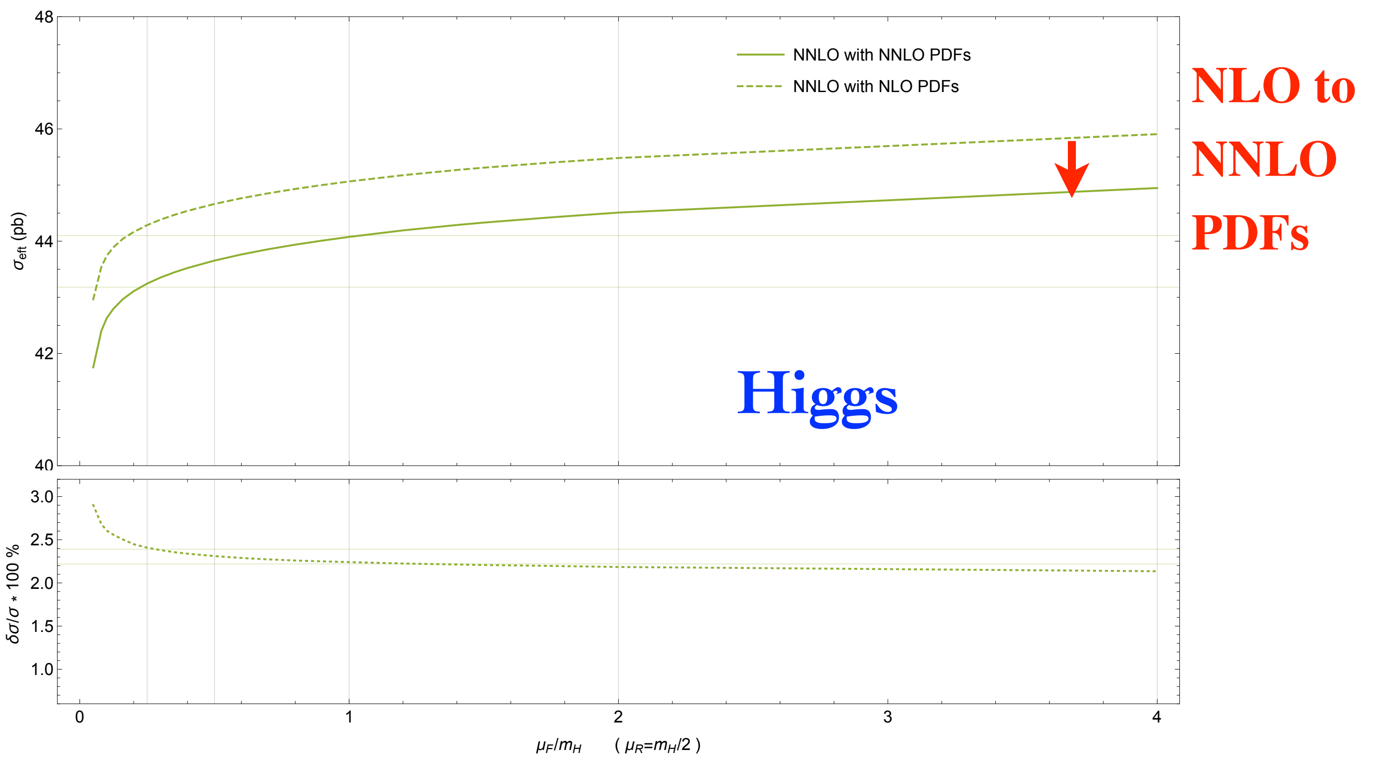



- The non-overlap is at the same level as the difference one might get from using N3LO PDFs.



# N3LO PDFs

- N3LO PDFs in principle needed to match this precision. Currently no full N3LO evolution and limited calculations for processes that enter fits, but approximations available.
- Important **future** milestone, but advances/time needed. Took ~ 10 years from first NNLO contributions in global fits to benchmark NNLO fits!
- **More immediate goal**: systematic inclusion of MHO uncertainties at NNLO with correlations.
- Higgs: the NLO to NNLO PDF difference at NNLO ~ constant with $\mu$.
- Suggests due to difference in PDF inputs from fit precision.
- **Work ongoing** towards NNLO MHO uncertainties!



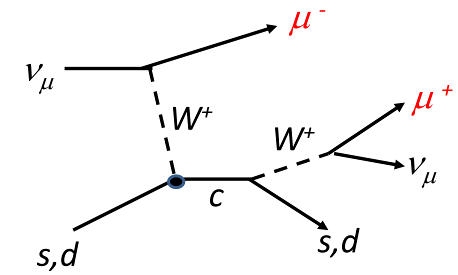
# Summary/Outlook

- ★ Precision LHC era: **significant opportunity** for Higgs physics and for PDF determination.
- ★ Precise data from LHC + precise theory already having **significant impact** on PDF fits and resulting Higgs cross sections. Multiple 'Post Run-I' sets available.
- ★ But **significant challenges** before us: benchmarking, confronting high precision data in fits, dealing with tensions, poor fit quality, including theory uncertainties effectively…
- ★ LHC data playing key role in global fits. But not only question of adding ever more data to PDF fits. Much work ahead to make sense of what we are seeing…

Thank you for listening!

# Backup

# Precision W,Z

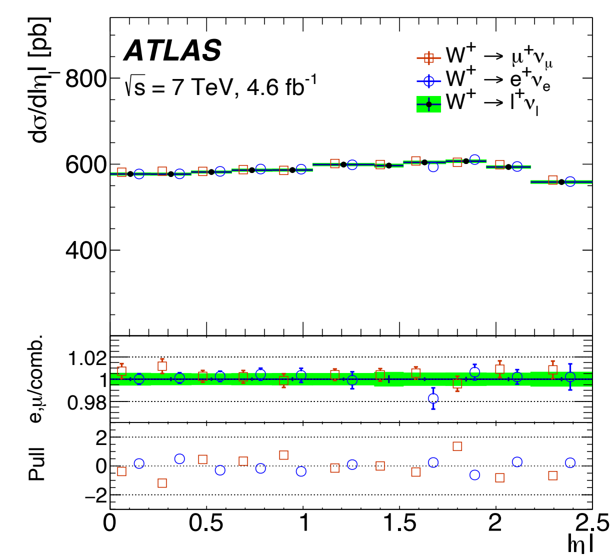



- Precision W, Z has significant impact on proton quark content.
- ATLAS 7 TeV W,Z: larger strangeness required than previous global fits, driven by neutrino induced dimuon production ($\overline{\nu} s \to lc$).
- **Updates** since original analysis:

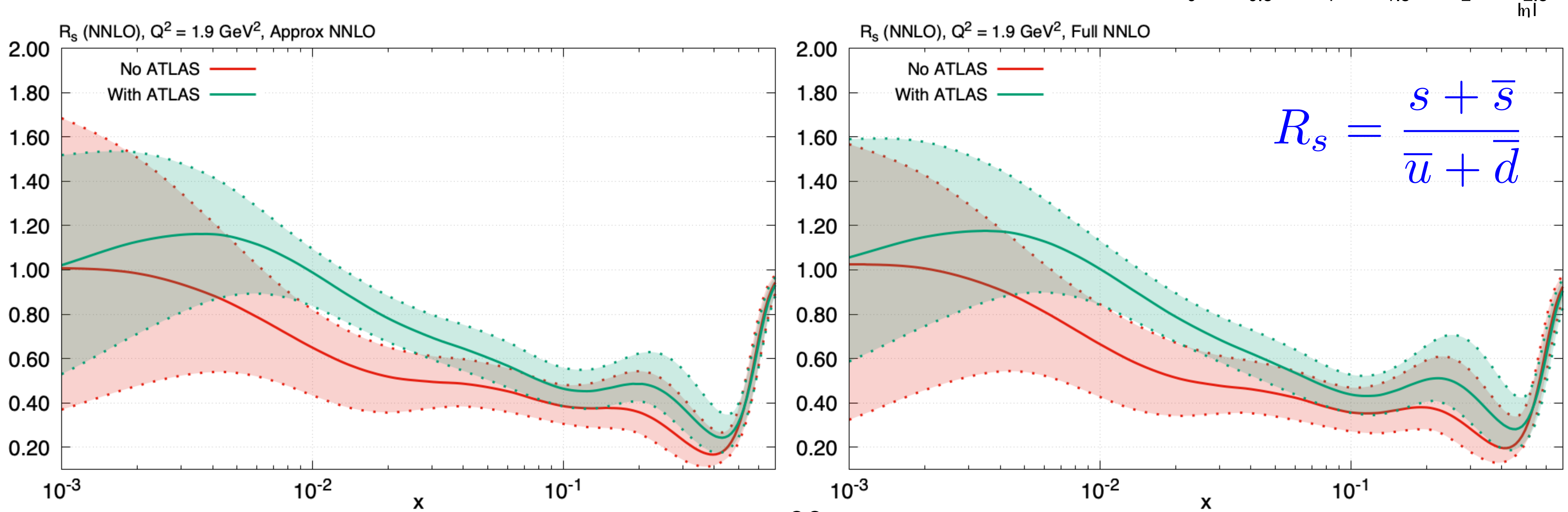
★ In global fit can account for both datasets, with mild tension. Strangeness increased.

★ Full NNLO corrections for dimuon data now included. Alleviates tension somewhat.

**ATLAS collab., Eur. Phys. J C77 (2017) 367**

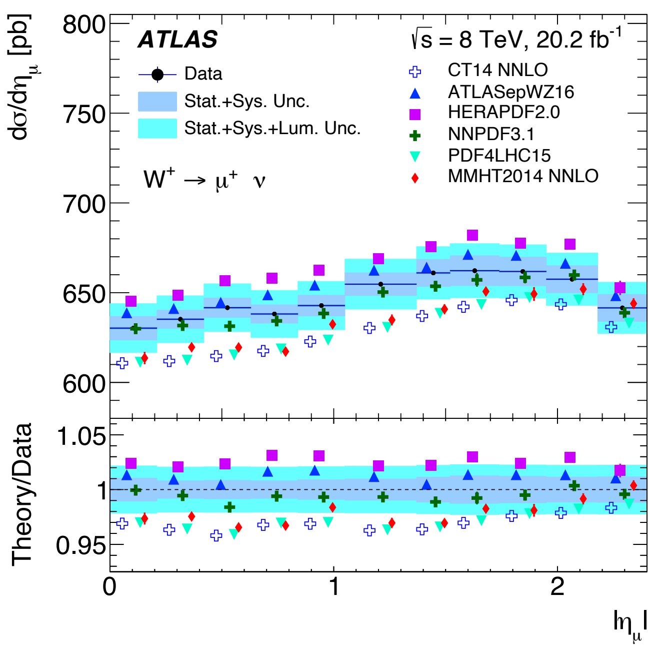



**J. Gao, JHEP 1802 (2018) 026**

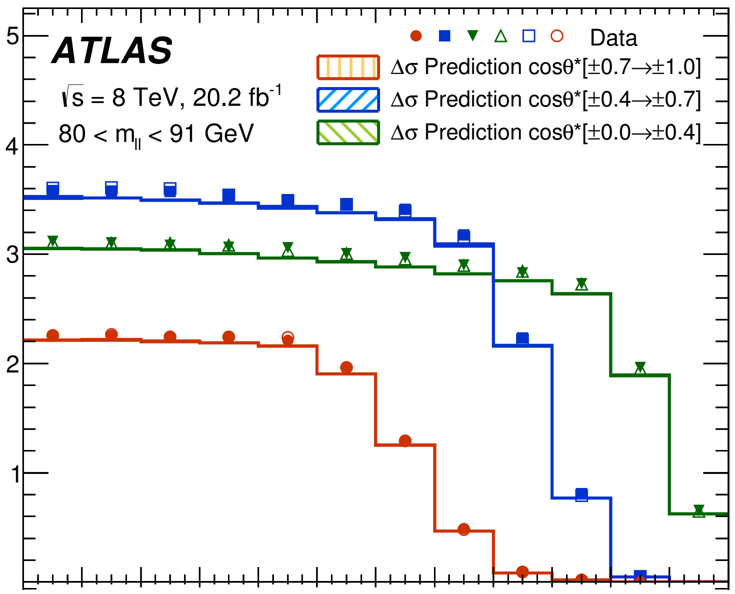


★ ATLAS 8 TeV W and triple differential lepton pair production available. Do they have the same impact?

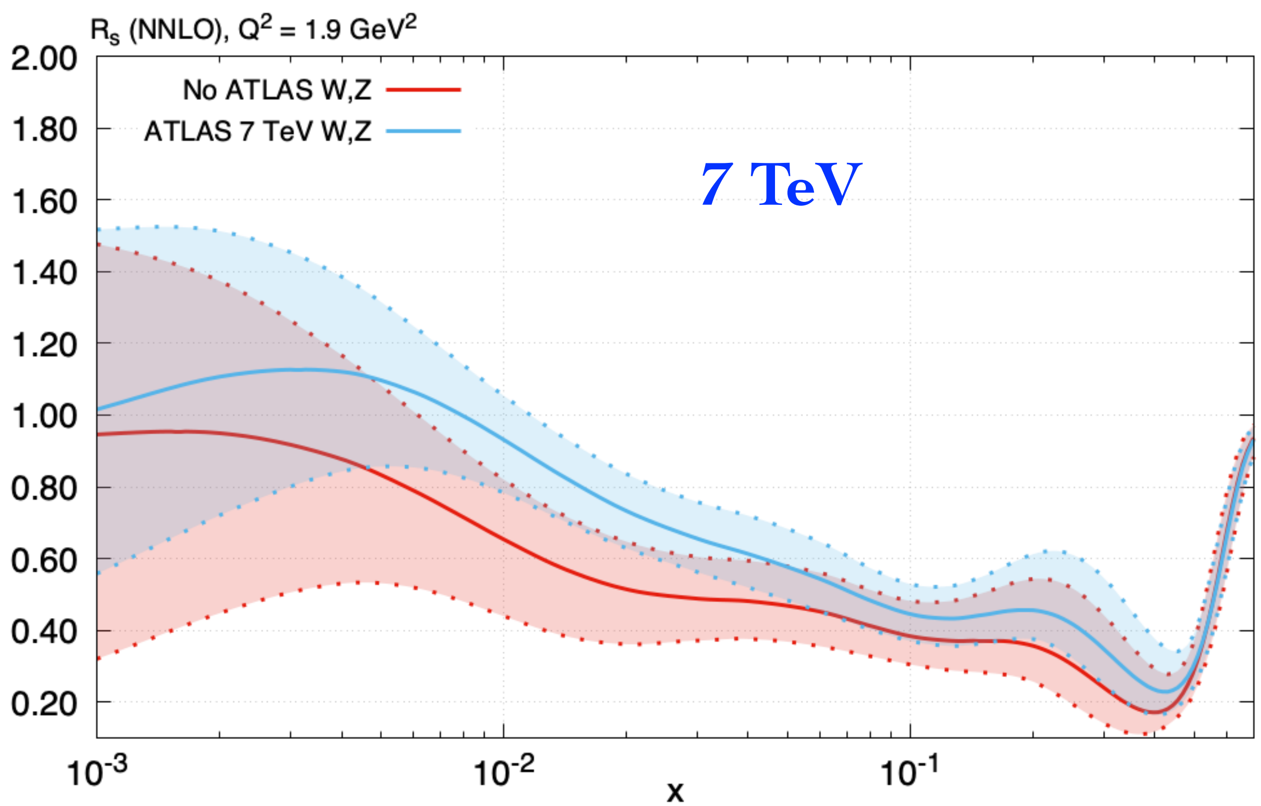
**ATLAS collab., Eur. Phys. J. C 79 (2019) 760**

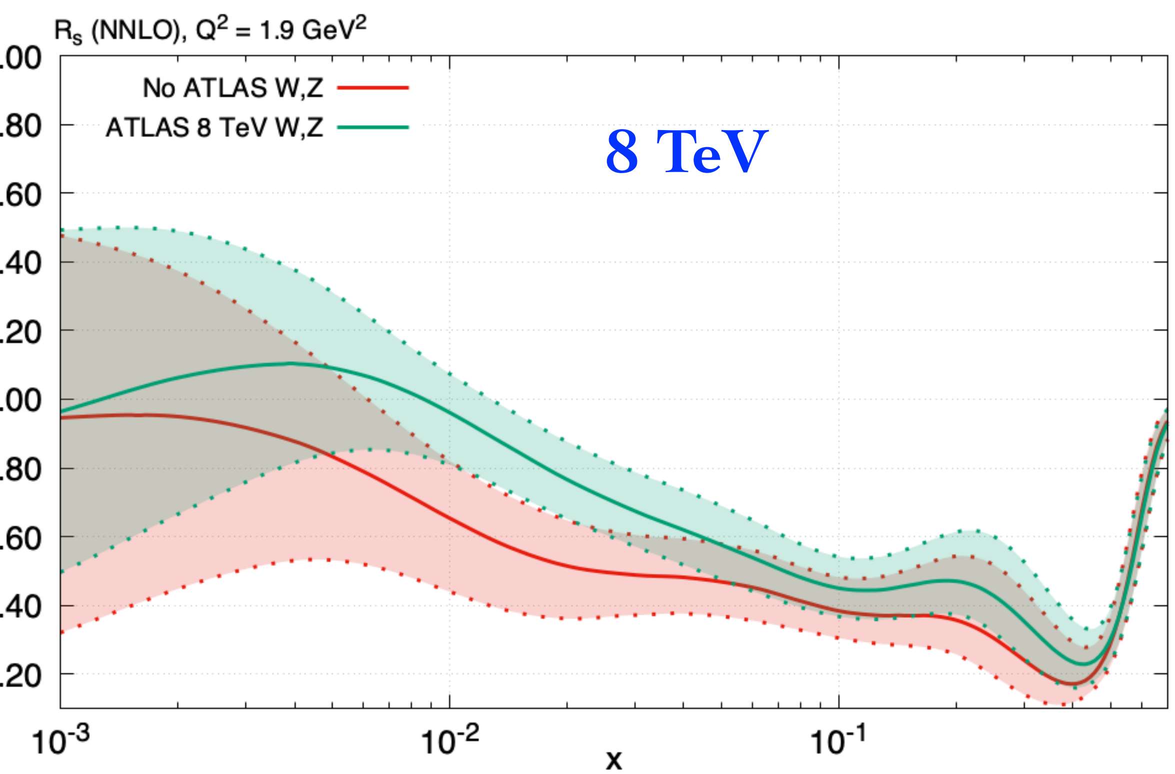



**ATLAS collab., JHEP 12 (2017) 059**




★ Yes! Nice consistent picture develops.



# Jet (and dijet) production

- Recent **NNPDF + NNLOJET** study: impact of inclusive jet and **dijet** data at NNLO QCD + NLO EW on global PDF fit.

**R. Abdul Khalek et al., 2005.11327**

| Experiment | Measurement | $\sqrt{s}$ [TeV] | $\mathcal{L}$ [fb$^{-1}$] | $R$ | Distribution | $n_{\rm dat}$ | Reference |
|---|---|---|---|---|---|---|---|
| ATLAS | Inclusive jets | 7 | 4.5 | 0.6 | $d^2\sigma/dp_T d\|y\|$ | 140 | [14] |
| CMS | Inclusive jets | 7 | 4.5 | 0.7 | $d^2\sigma/dp_T d\|y\|$ | 133 | [16] |
| ATLAS | Inclusive jets | 8 | 20.2 | 0.6 | $d^2\sigma/dp_T d\|y\|$ | 171 | [15] |
| CMS | Inclusive jets | 8 | 19.7 | 0.7 | $d^2\sigma/dp_T d\|y\|$ | 185 | [17] |
| ATLAS | Dijets | 7 | 4.5 | 0.6 | $d^2\sigma/dm_{jj} d\|y^*\|$ | 90 | [18] |
| CMS | Dijets | 7 | 4.5 | 0.7 | $d^2\sigma/dm_{jj} d\|y_{\rm max}\|$ | 54 | [16] |
| CMS | Dijets | 8 | 19.7 | 0.7 | $d^3\sigma/dp_{T,{\rm avg}} dy_b dy^*$ | 122 | [19] |

- Scale $\mu = \hat{H}_T$ taken, based on earlier work.

**J. Currie et al., JHEP 10 (2018) 155**

- Detailed study, considering consistency of datasets, perturbative stability, impact on PDFs…

**See A. Huss's talk**

- **Intriguing picture** found. For **dijet** data, clear preference for NNLO:

**NLO :**

$$\chi^2/N_{\rm pts} = 2.44$$

**NNLO :**

$$\chi^2/N_{\rm pts} = 1.65$$

- Whereas for **inclusive** jets, the trend is opposite (!):

**NLO :**

$$\chi^2/N_{\rm pts} = 1.25$$

**NNLO :**

$$\chi^2/N_{\rm pts} = 1.88$$

- NNLO fit quality slightly better for dijets (1.65 vs. 1.88) and moreover global fit quality is better (driven by top data):

**Fitting inclusive:**

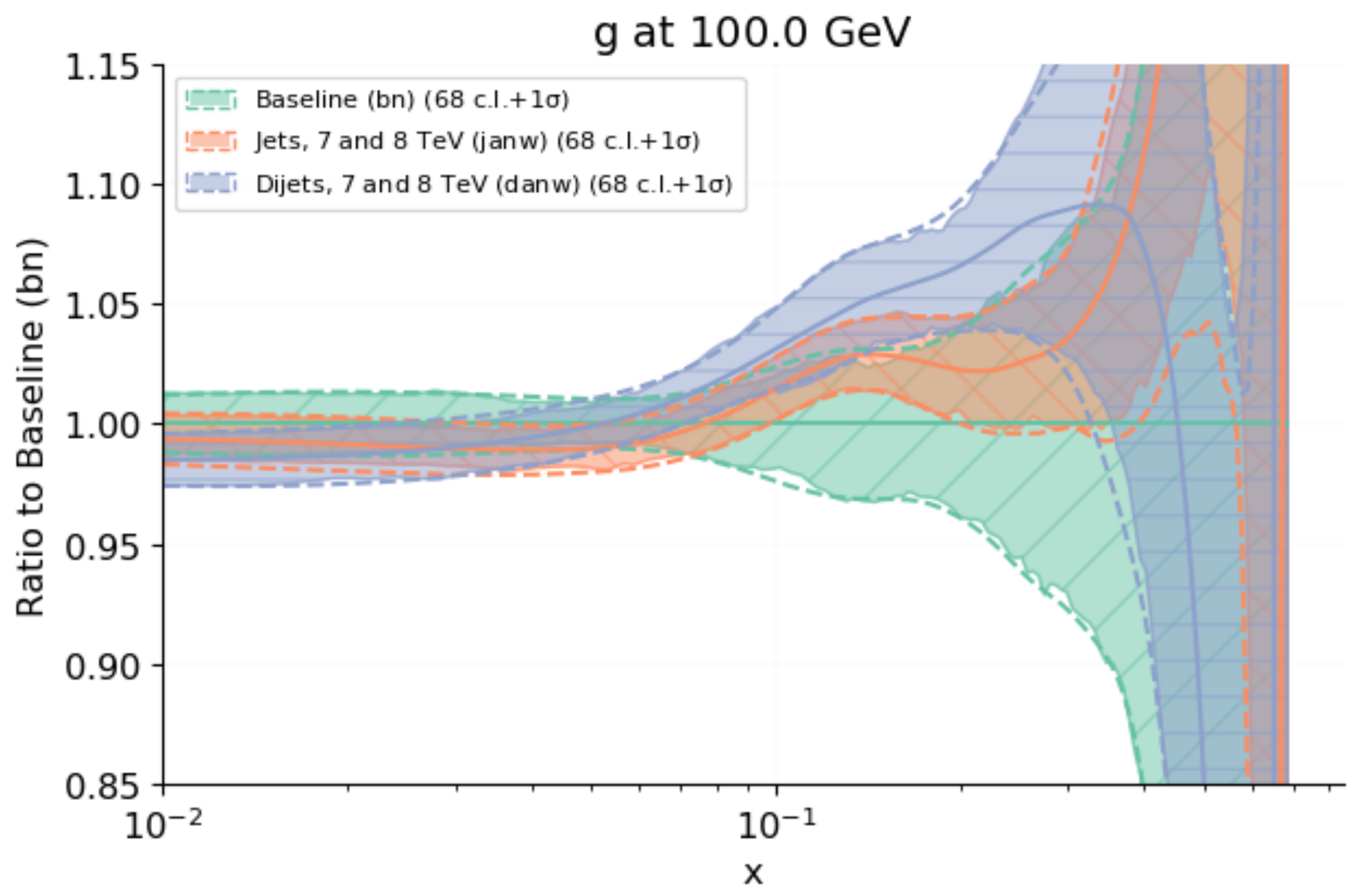
$$\chi^2_{\rm global}/N_{\rm pts} = 1.28$$

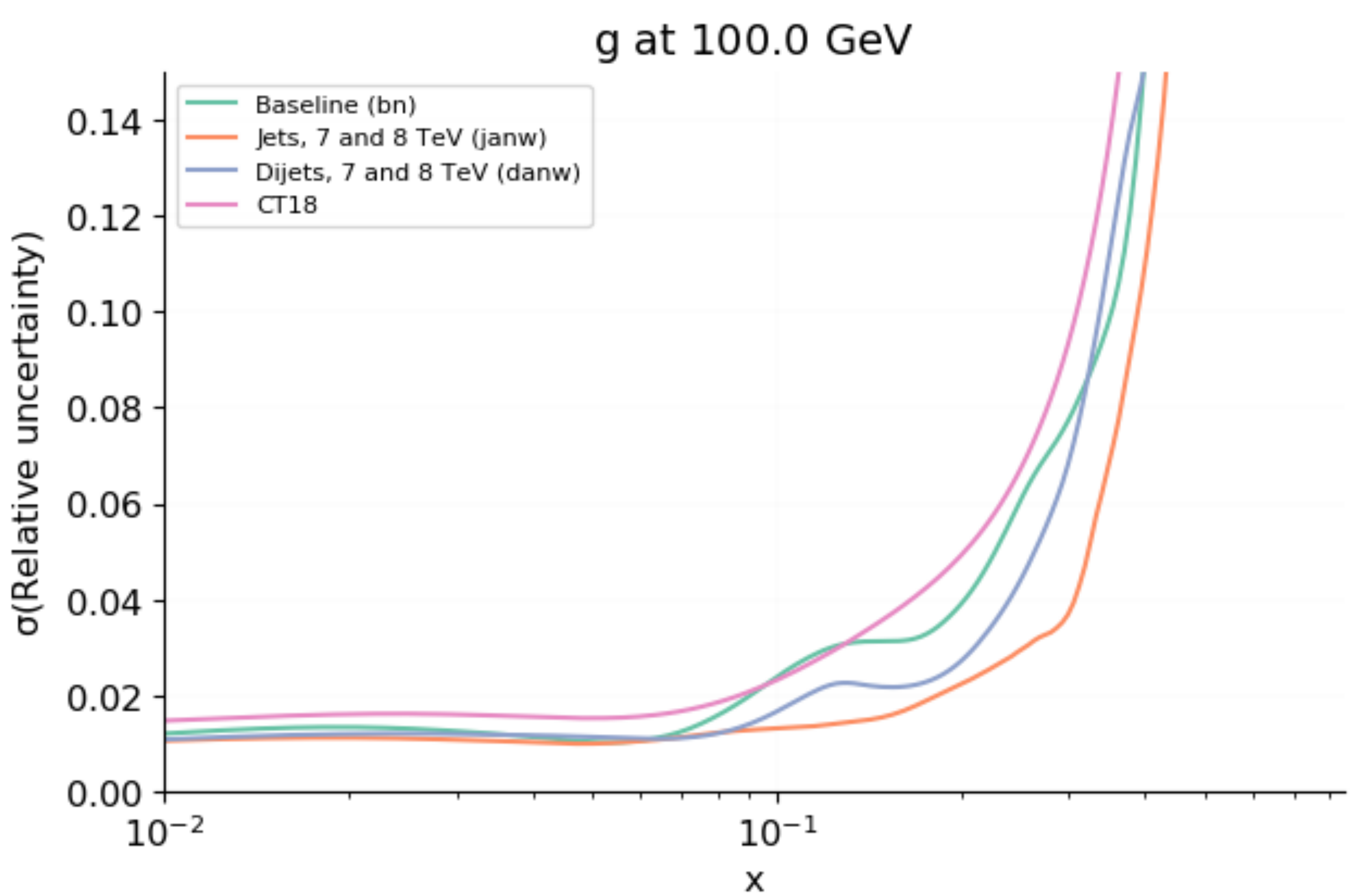
**Fitting dijets:**

$$\chi^2_{\rm global}/N_{\rm pts} = 1.22$$

- Prediction for inclusive/dijets reasonable when fitting dijets/inclusive - no big tensions between them.
- Interesting to see results for other groups (impact of $t\bar{t}$ not always same, MSHT see improvement at NNLO vs. NLO in inclusive...).

- Relatively **consistent impact** on gluon when fitting inclusive v.s dijets.



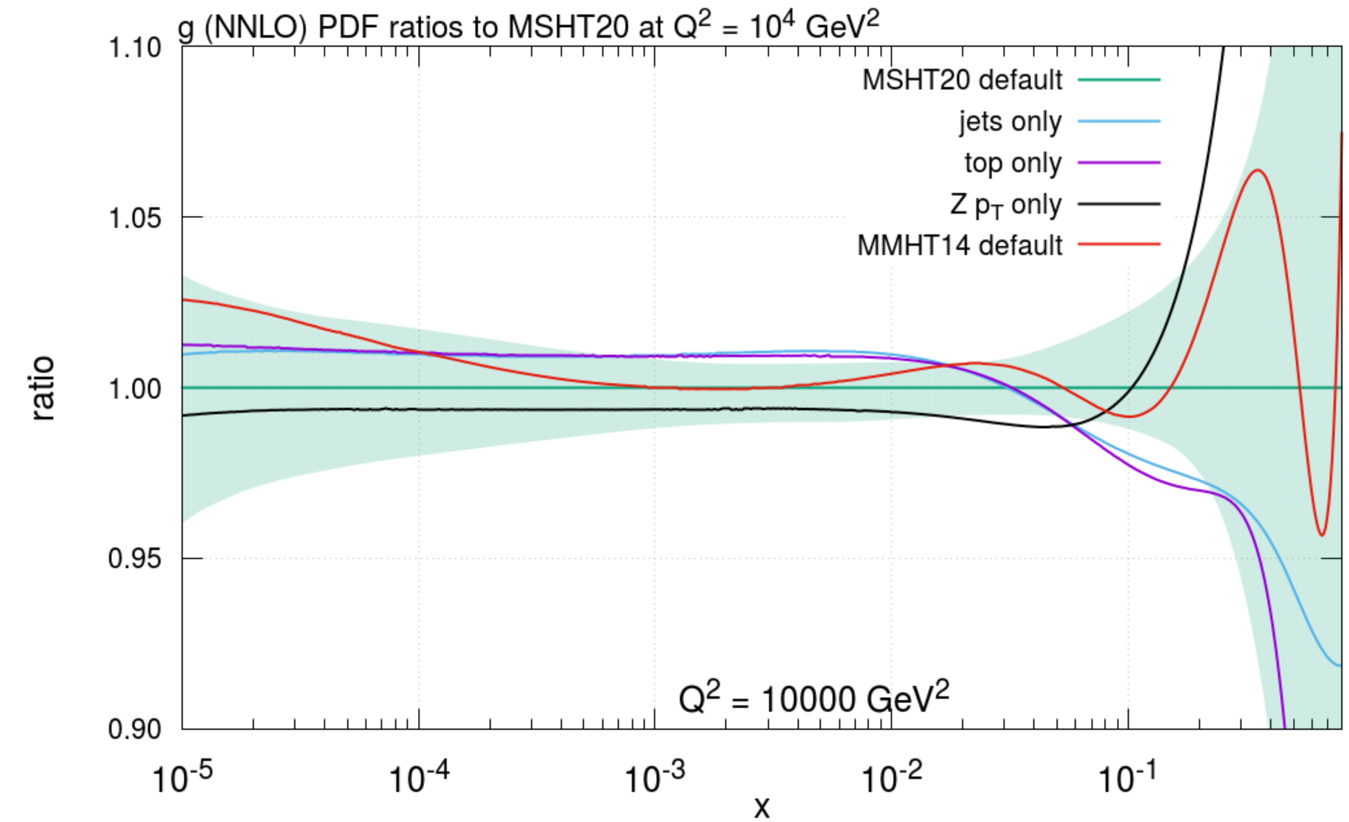




- Sizeable **uncertainty reduction** at high $x$, impact somewhat more in inclusive case (more data).

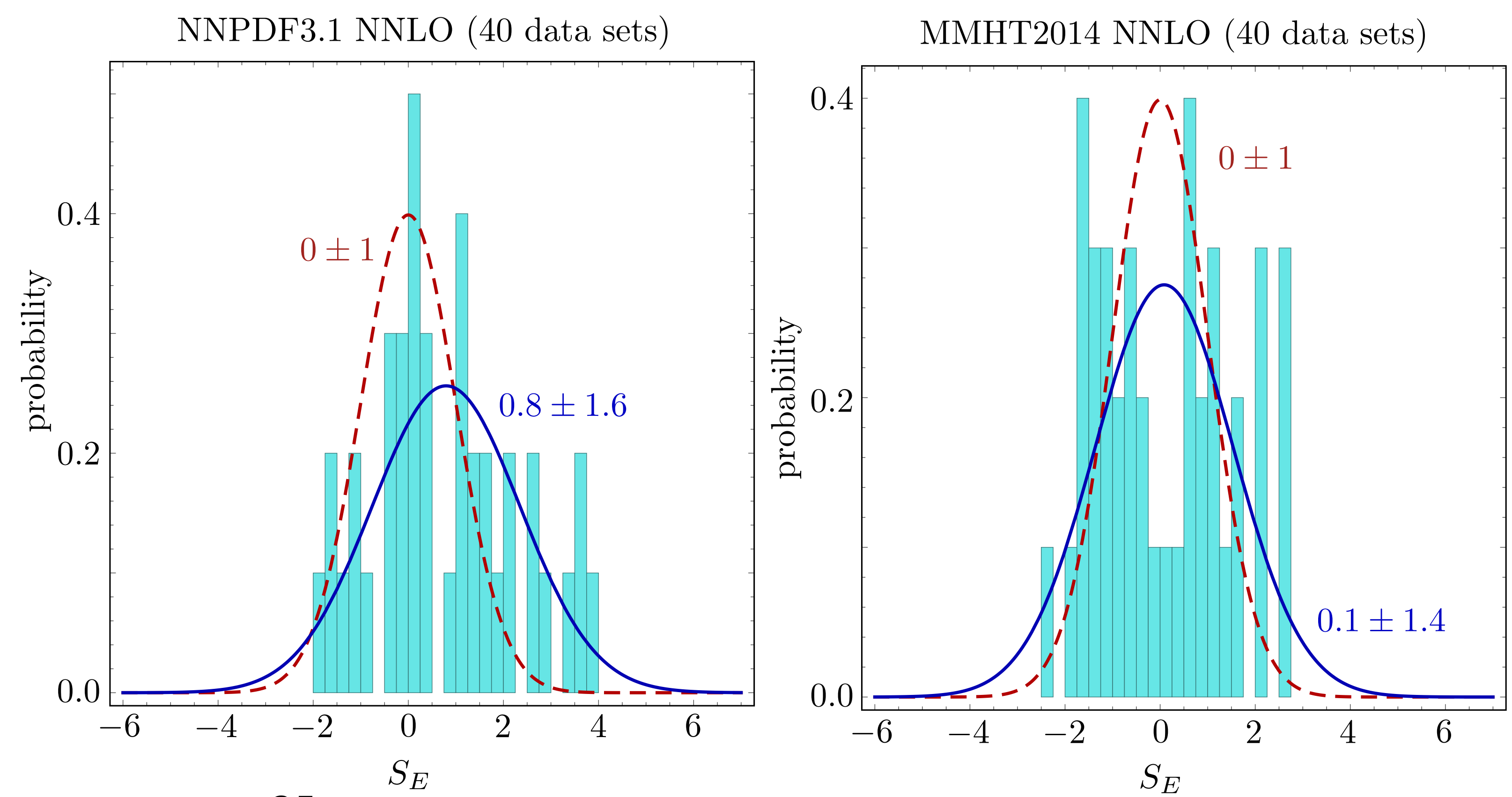
# Challenges in reaching high precision

# A Global Fit

- Global PDF fits delivers both robustness but also challenges.
- **Robustness**: many datasets entering, insensitivity to any particular experiment/assumption.
- **Challenges**: data often in tension, fit quality far from idealised textbook case.




- Various ways data/ theory may not follow textbook expectations: **experimental** and **theoretical**.

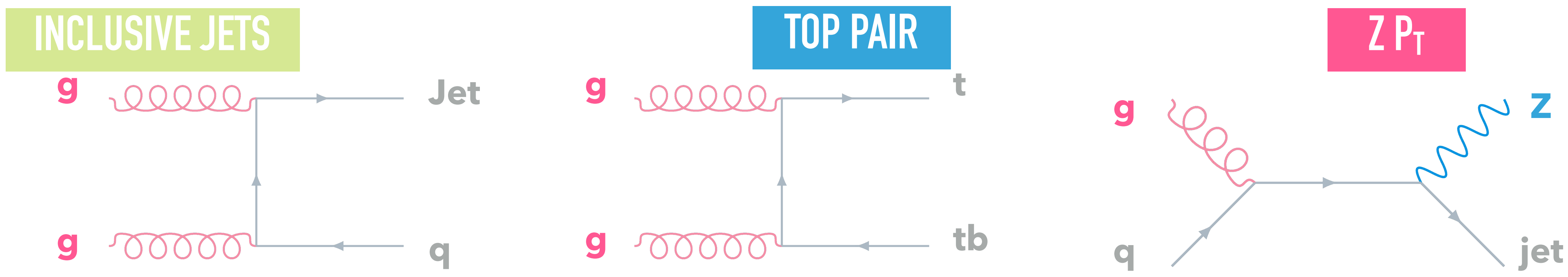





K. Kovarik, P. Nadolsky, D. Soper, arXiv:1905.06957

# Dealing with high precision data

- LHC data in fits often systematics dominated, which significant correlations between these errors.
- Difficulties in achieving good fit quality to various 'textbook' processes for PDF determination:




M. Ubiali, Higgs Coupling 2019

- Detailed studies performed, with various solutions proposed for modifying systematics in physically reasonable way. Bottom line: care needed!

**(Backup)**

**S. Bailey & LHL, *Eur.Phys.J.C* 80 (2020) 1, 60**

**ATL-PHYS-PUB-2018-017**

**LHL, R.S. Thorne, A.D. Martin, EPJC78 (2018) no.3, 248**

**R. Boughezal et al., JHEP 1707 (2017) 130**

# Confronting Precise Data

- LHC data playing increasingly important role in PDF fits.

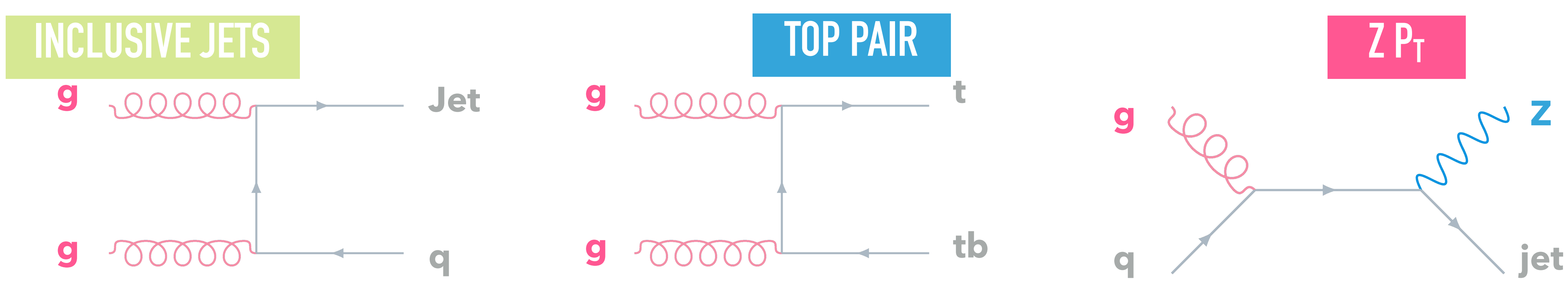
Basic motivation:

★ High precision, multi-differential data.

★ High precision theory: NNLO QCD 'standard'.

$\longrightarrow$ **High precision** PDF determination.

- However in a number of cases we are seeing **difficulty** in confronting such high precision data in PDF fits.
- Occurs in three 'textbook' LHC processes for PDF determination:




M. Ubiali, Higgs Coupling 2019

# Top Quark Pair Production

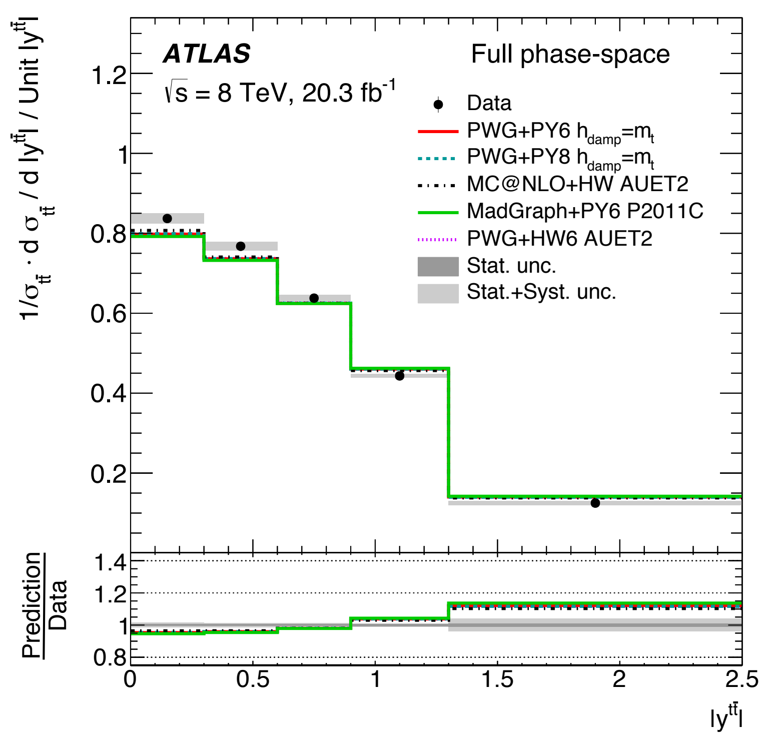
- ATLAS 8 TeV top quark pair production*: presented differentially in four relevant kinematic variables w. all cross-correlations $\Rightarrow$ **simultaneous fits**.
- Significant issues in fit quality…

$$\chi^2/N_{\rm pts}\ (N_{\rm pts}^{\rm tot} = 25)$$

| | |
|---|---|
| $p_T$ | 0.53 |
| $y_t$ | 3.12 |
| $y_{tt}$ | 3.51 |
| $M_{tt}$ | 0.70 |
| $p_T + M_{tt}$ | 5.73 |
| Combined | 7.00 |

**S. Bailey & LHL, *Eur.Phys.J.C* 80 (2020) 1, 60**

$$y_t,\ y_{t\bar{t}},\ p_\perp^t,\ M_{t\bar{t}}$$




**ATL-PHYS-PUB-2018-017**

***In addition issues with fit quality in CMS 8 TeV lepton + jet data for certain individual distributions seen by both CT and MMHT.**

- Data dominated by **systematics**. In particular, **two-point** MC uncertainties in correction back to top quark level by far the largest:
  - ★ Parton Shower: `POWHEG` + `Herwig` vs. `POWHEG` + `Pythia`
  - ★ Hard Scattering: `MC@NLO` + `Herwig` vs. `POWHEG` + `Herwig`
  - ★ ISR/FSR: `POWHEG` + `Pythia`(1) vs. `POWHEG` + `Pythia`(2)

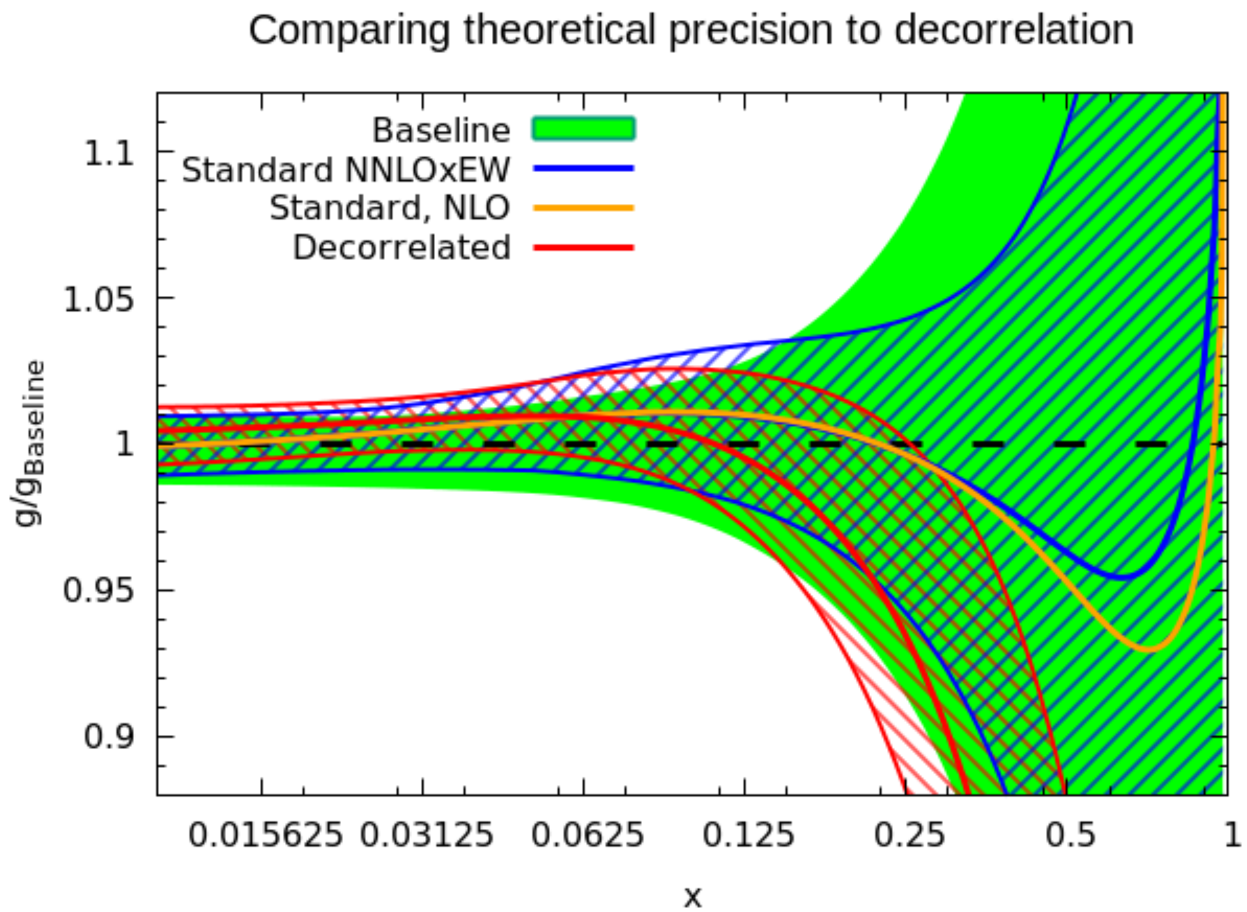
**ATL-PHYS-PUB-2018-017**

- Uncertainty **and** correlation effectively given by envelope of two MCs. Clear this method will not capture the true degree of correlation.
- Our study*: try some reasonable **loosening** of the assumed **correlation.**

- Decorrelation improves fit quality a lot, but gluon sensitive to it!
- Dependence larger than on e.g. NLO vs. NNLO theory input: **Care needed!**

| Distribution | p.s. correlated | p.s. decorrelated |
|---|---|---|
| Combined | 7.00 | 1.80 |
| $p_\perp^t + M_{tt}$ | 5.73 | 0.66 |

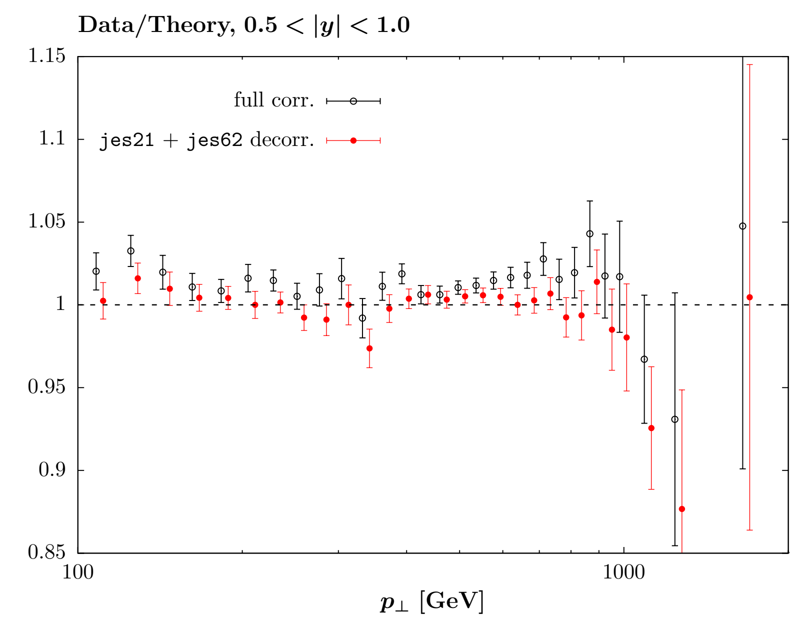
***Similar more limited study in ATL-PHYS-PUB-2018-017**



# Other Examples: Jets and Z pt

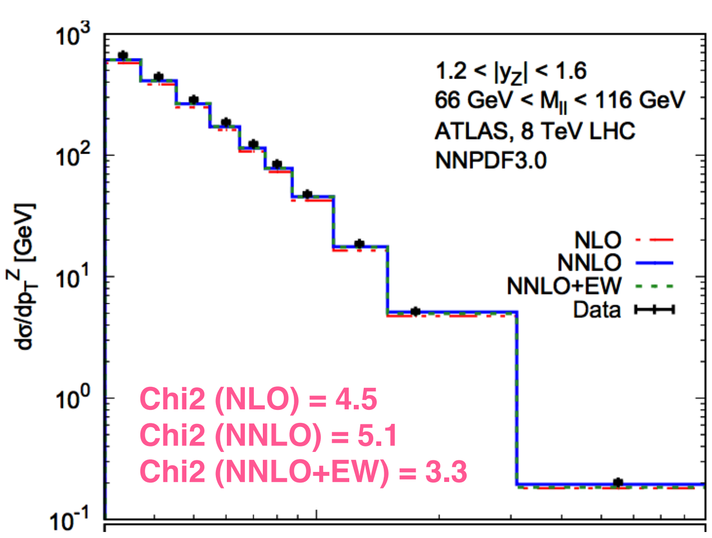
- ATLAS inclusive **Jet data**: again systematics dominated, and fit quality highly sensitive to correlations.
- Decorrelation improves things again. **NNPDF**: dijets potentially less sensitive to these effects? Aside: decorrelated inclusive + dijet gluon agree better (**Backup**).

**LHL, R.S. Thorne, A.D. Martin, EPJC78 (2018) no.3, 248**




- **Z boson** transverse momentum distribution: similar situation.
- **NNPDF**- improve by adding in uncorrelated source of uncertainty (missing theory, underestimated experimental errors…).

**R. Boughezal et al., JHEP 1707 (2017) 130**




$\longrightarrow$ **Care needed** in interpreting what precision data is telling us. Best dealt with case by case and with full breakdown of experimental systematics provided.