Impact of Jet data on PDFs

Lucian Harland-Lang, University of Oxford

Jets and their Substructure from LHC data, 31 May 2021

In collaboration with Shaun Bailey, Tom Cridge, Alan Martin and Robert Thorne

With special thanks to Emanuele Nocera and Alex Huss

UNIVERSITY OF OXFORD

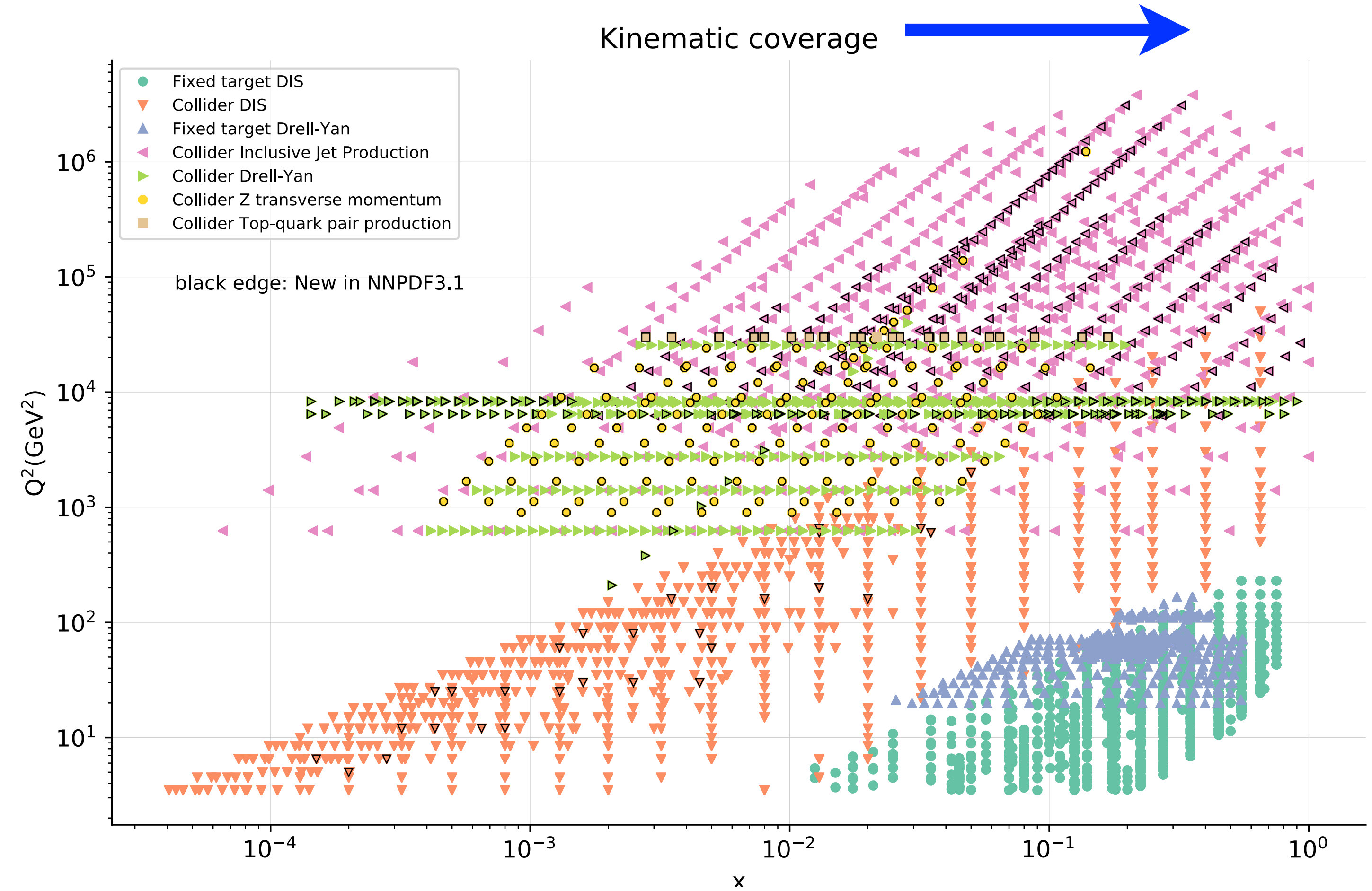
Jets for PDF fits

- Jet production a key ingredient in modern PDF fits.
- Sensitive to gluon and quark initial states:

$$gg \to gg,\ gg \to q\bar{q},\ gq \to gq,\ q\bar{q} \to gg\ ,$$
$$q\bar{q} \to q\bar{q},\ q\bar{q} \to q'\bar{q}',\ q\bar{q}' \to q\bar{q}',\ qq \to qq,\ qq' \to qq'\ ,$$

High $p_\perp$, m_{jj}

- By pushing to larger jet $p_\perp$ (dijet m_{jj}) go to larger x .
- Though quark-initiated contribution can be large, these are rather well constrained (DIS).
- Hence, jet data particularly relevant for gluon at high x (less well constrained).



• NNLO QCD (and NLO EW) theory available for both inclusive and dijet data.

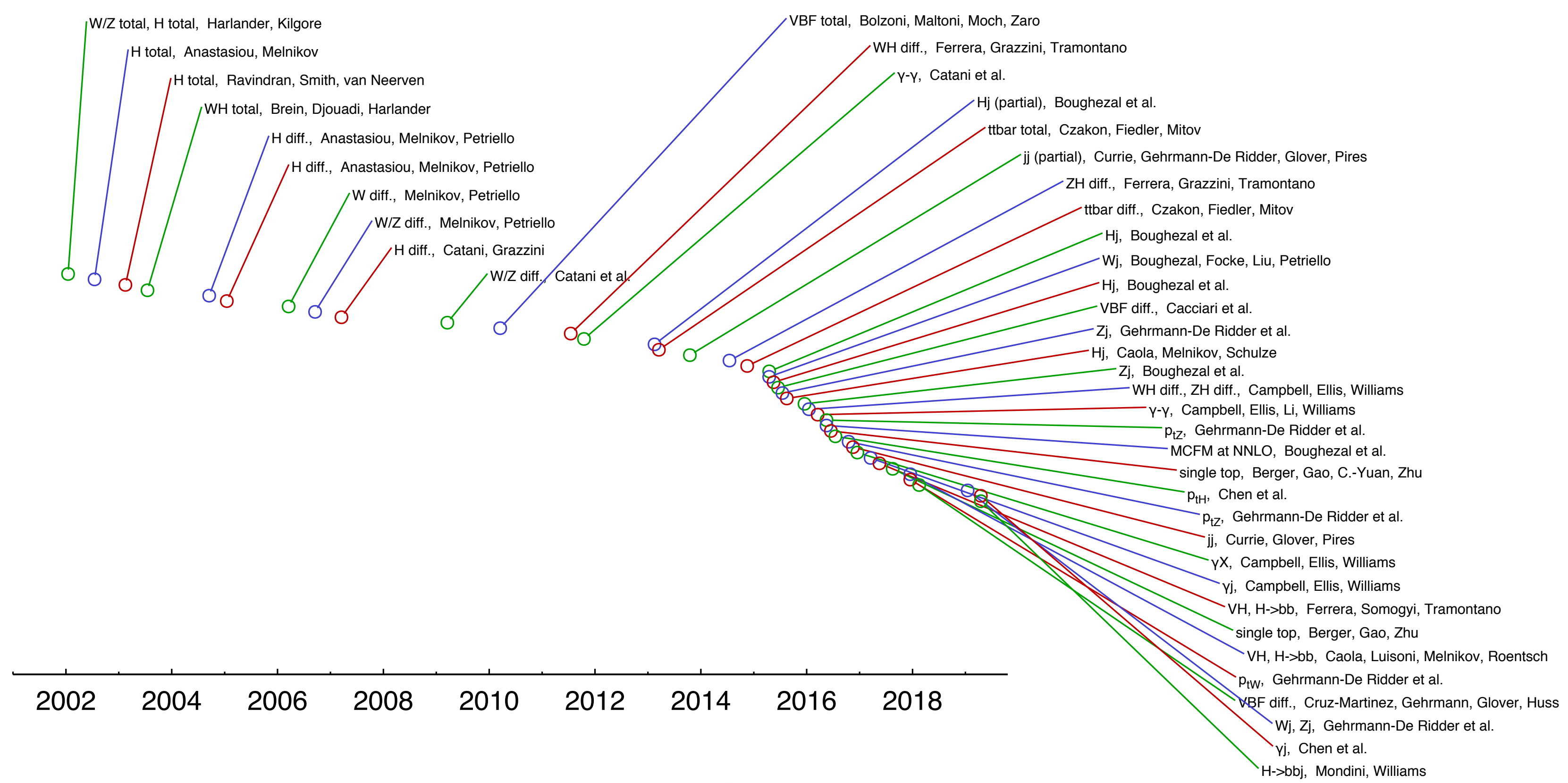


Image Credit: Gavin Salam

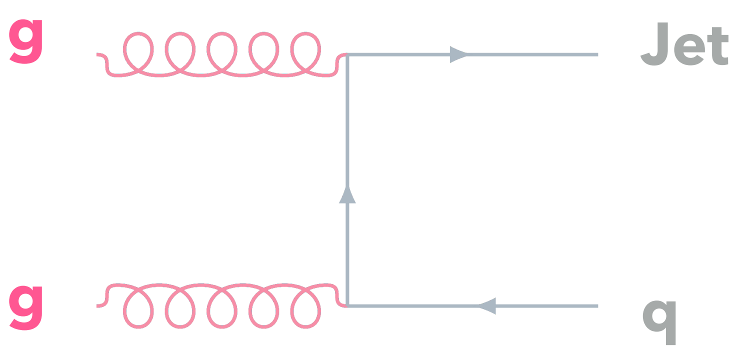
See T. Gehrmann's Talk

• As we will see a key ingredient in fitting these data (particularly dijets).

• In addition, significant amount of inclusive jet and dijet data from LHC. High precision and spanning large range of kinematic space.

See B. Bilin and P. Starovoitov talks

Jet Kinematics: Inclusive



- Inclusive jets measured in terms of jet $p_\perp$ and y_j.
- Schematically, **LO** relationship to high x parton:

$$x = \frac{p_\perp}{\sqrt{s}} \left(e^{y_j} + e^{y_{j'}} \right)$$

Observed Jet j ($y_j > 0$)
'Unobserved' Jet j'

$\longrightarrow$ Need 3 kinematic inputs to uniquely determine x.

- Inclusive jets: effectively integrate over $x \gtrsim \frac{p_\perp}{\sqrt{s}} e^{y_j}$.
- So certainly sensitive to high x region, but washed out somewhat.

Jet Kinematics: Dijets

- For dijets, both jets measured. Same schematic **LO** relationship:

$$x_{1,2} = \frac{p_\perp}{\sqrt{s}} \left(e^{\pm y_j} + e^{\pm y_{j'}} \right)$$

- Double differential measurements in terms of m_{jj} and $y^*/y_{\max}$: not sufficient to uniquely pin down LO x .

- That is, some washing out (though precise effect depends on choice of variable).

- However, also possible to measure triple differentially - expect to provide stronger, more direct constraints.

$$\mathrm{d}^3\sigma/\mathrm{d}p_{\perp,avg}\mathrm{d}y_b\mathrm{d}y^*$$

See J.Stark's/K. Rabbertz's talks

Jets in MSHT20

S. Bailey et al., *Eur.Phys.J.C* 81 (2021) 4, 341

Parton distributions from LHC, HERA, Tevatron and fixed target data: MSHT20 PDFs

S. Bailey[a], T. Cridge[b], L. A. Harland-Lang[a], A. D. Martin[c], and R.S. Thorne[b]

[a] Rudolf Peierls Centre, Beecroft Building, Parks Road, Oxford, OX1 3PU
[b] Department of Physics and Astronomy, University College London, London, WC1E 6BT, UK
[c] Institute for Particle Physics Phenomenology, Durham University, Durham, DH1 3LE, UK

Abstract

We present the new MSHT20 set of parton distribution functions (PDFs) of the proton, determined from global analyses of the available hard scattering data. The PDFs are

- The 'Post-Run I' set from the MSTW, MMHT… group: **MSHT20.**

- Focus on including significant amount of **new data**, higher **precision theory** and on **methodological improvements**.

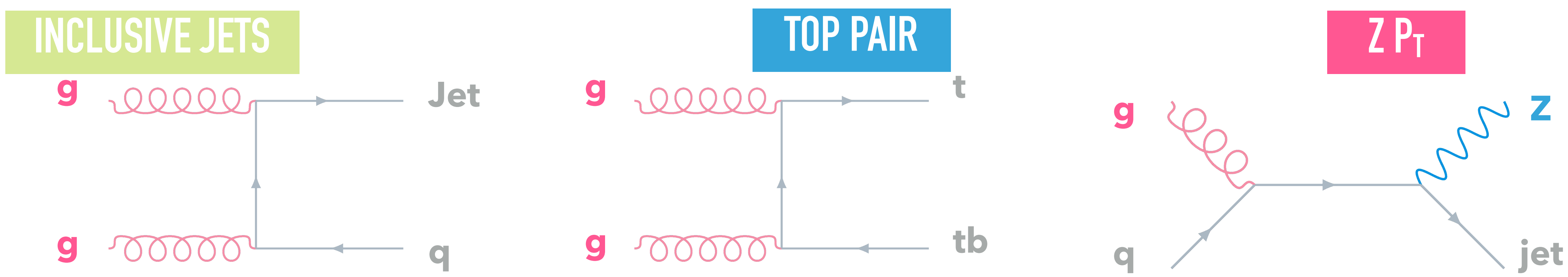
  - Range of inclusive LHC jet dat fit:

	NNLO, $\chi^2/N_{\rm pt}$
ATLAS 7 TeV jets [18]	221.6/140
CMS 2.76 TeV jet [107]	102.9/81
CMS 7 TeV jets [100]	175.8/158
CMS 8 TeV jets [101]	261.3/174

- Fit quality acceptable. Impact tied up with other high x gluon sensitive data….

High x gluon: global sensitivity

- Jet data one of three major LHC datasets with potential significant impact on high x gluon.



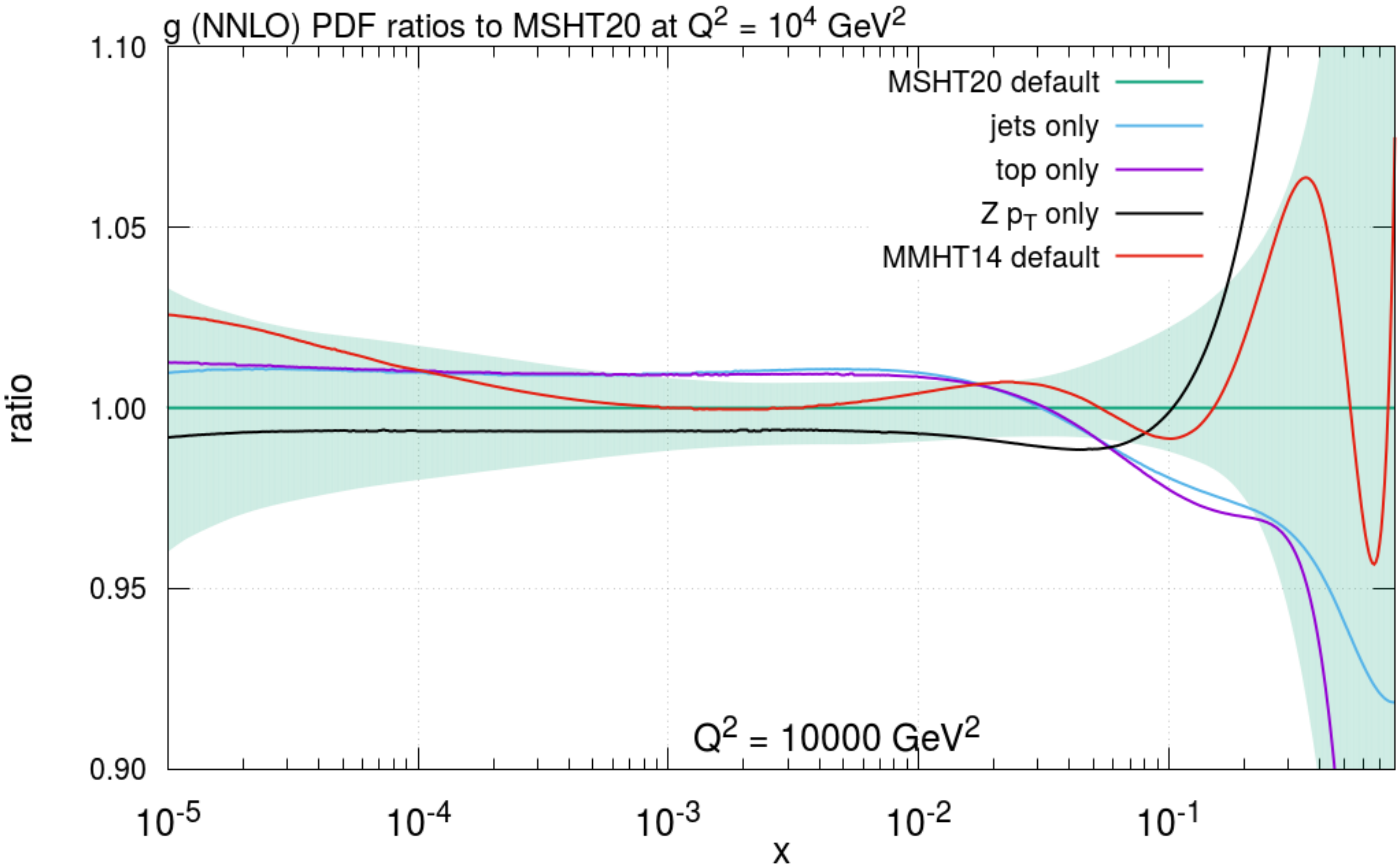
M. Ubiali, Higgs Coupling 2019

- NNLO QCD + high precision multi-differential data available.

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

S. Bailey et al., *Eur.Phys.J.C* 81 (2021) 4, 341
S. Bailey and LHL., *Eur.Phys.J.C* 80 (2020) 1, 60
LHL, R.S. Thorne and A.D. Martin., *Eur.Phys.J.C* 78 (2018) 3, 248

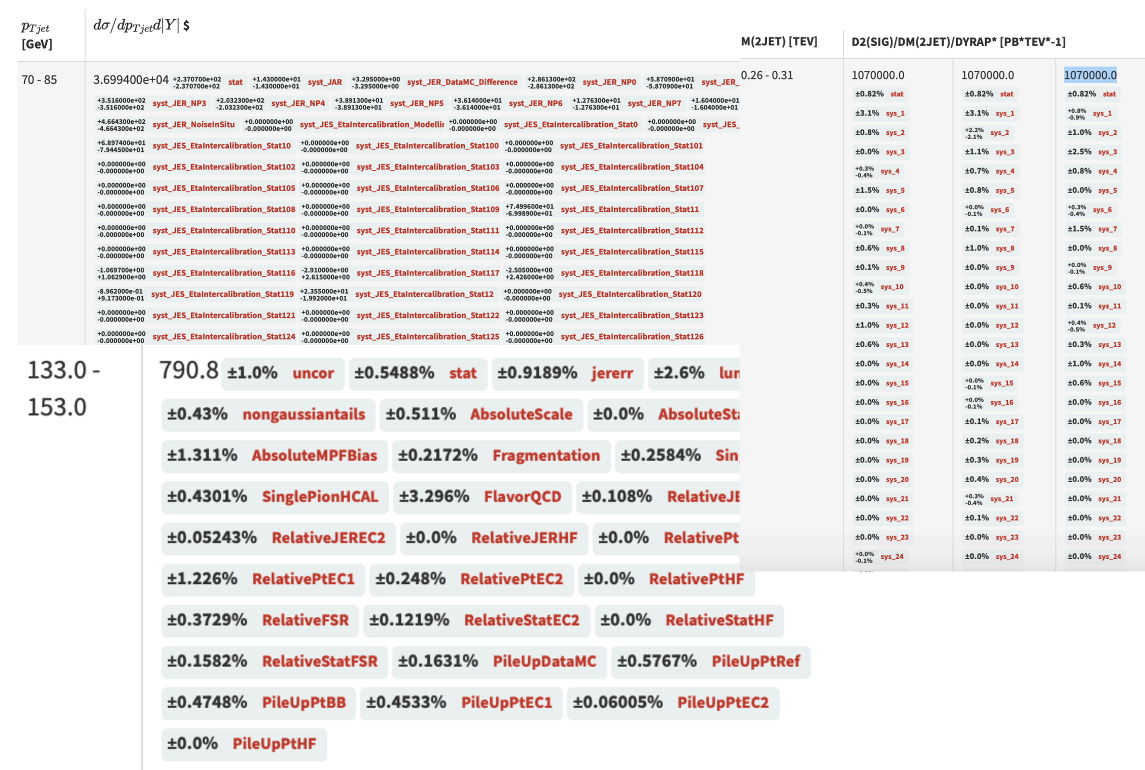
- Caveats: various issues in fit quality in all three cases.
- In terms of gluon impact, these do not pull in same direction (for MSHT).



- MSHT20: top and jet data pull in (roughly) same direction, i.e. lower high x gluon. $Z\,p_\perp$ prefers the opposite.
- Final result a balance between these. Important to bear in mind: impact on PDFs in global fit and not in isolation!

Fitting Jets

- Is fit quality good? For jets well beyond simple: $\chi^2 = \sum_{k=1}^{N_{\rm pt}} \frac{(D_k - T_k)^2}{s_k^2}$



- In particular, measurements often dominated by systematic errors and their **correlations**:

- Fit quality accounts for these, and can be dominated by them:

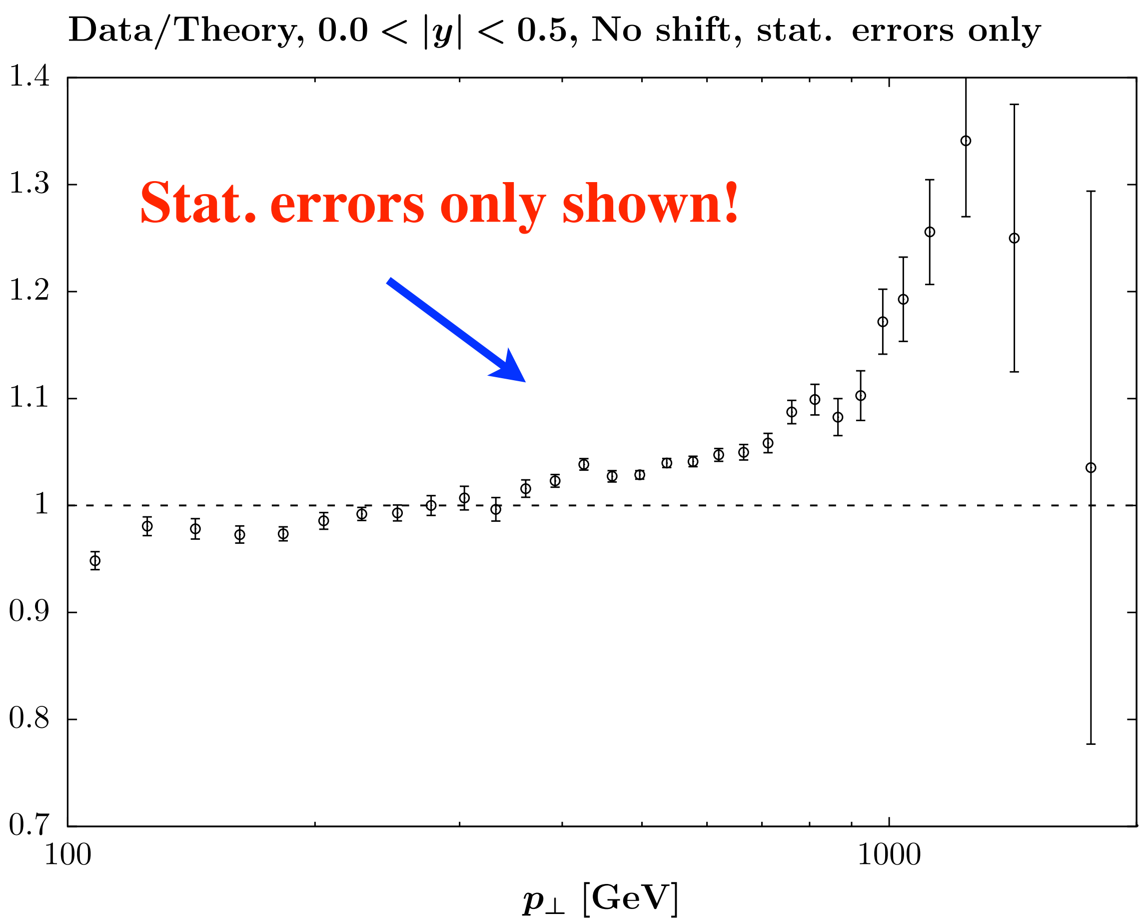
$$\chi^2(\{a\},\{\lambda\}) = \sum_{k=1}^{N_{pt}} \frac{1}{s_k^2}\left(D_k - T_k - \sum_{\alpha=1}^{N_\lambda} \beta_{k,\alpha}\lambda_\alpha\right)^2 + \sum_{\alpha=1}^{N_\lambda} \lambda_\alpha^2 \,,$$

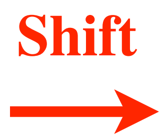
$$D_k \to D_k - \sum_{\alpha=1}^{N_\lambda} \beta_{k,\alpha}\hat{\lambda}_\alpha$$

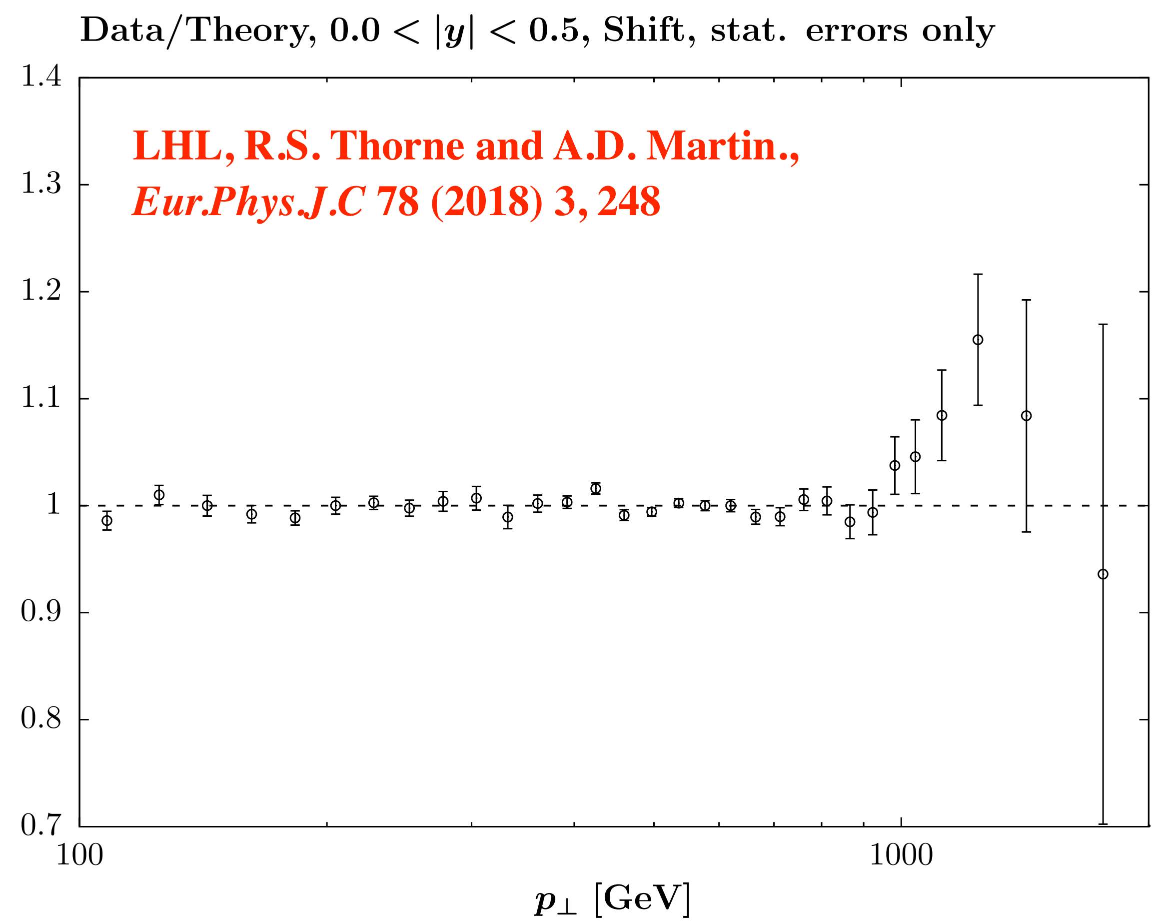
Uncorrelated errors **Correlated errors** **Penalty for shifts**

Fitting Jets

- This issue/question arises in e.g. description of ATLAS 7, 8 TeV jet data. Provided differentially in $p_\perp$ for various jet rapidity bins.
- Studied in **arXiv:1711.05757**. 7 TeV data can be v. well fit for single rapidity bin, but not all in combination.
- Can also get good fit with systematics correlated only within rapidity bins.
- This is clearly too strong a choice, but indicates sensitivity.







- Fitting to e.g. just one rapidity bin throws away information and does not resolve underlying issue.

- In some cases due to two-point model variations - can certainly loosen. Dedicated ATLAS study on this.

ATLAS Collab., JHEP 09 (2017) 020

Splitting options for $R = 0.4$	CT14	NNPDF3.0
JES Flavour Response Opt 7		
JES MJB Fragmentation Opt 17		
JES Pile-up Rho topology Opt 18		
Scale variations Opt 17		
Alternative scale choice Opt 7		
Non-perturbative corrections Opt 7	268/159	257/159

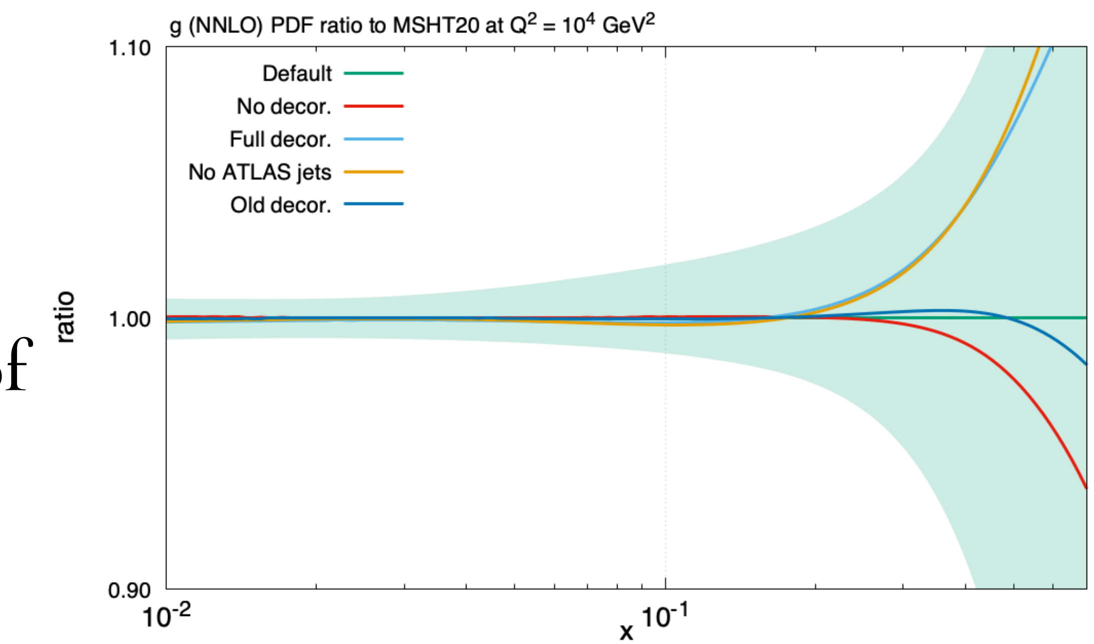
- We follow this in MSHT fits:

ATLAS 7 TeV $\chi^2/N_{\rm pt}$ **S. Bailey et al., *Eur.Phys.J.C* 81 (2021) 4, 341**

	No decor.	Ref. [15] decor.	Smooth decor.	Full decor.
$\rm LO_{EW}$	2.00	1.09	1.48	0.81
$\rm NLO_{EW}$	1.80	1.15	1.57	0.92

- Permissible decorrelation improves fit. Impact on gluon limited (though not zero).

- **However**: level of improvement sensitive to theory input. More complete theory/inclusion of MHO uncertainties may reduce this sensitivity further. Delicate interplay here…



Fitting Jets/Dijets: Recent Studies

Jet Data at the LHC

- Focussing on Run-I data (i.e. current PDF fits):

$\mathrm{d}^2\sigma/\mathrm{d}p_\perp \mathrm{d}y$

  - Inclusive jets:

$0.0 < |y| < 2.5 - 3.0$

★ CMS 2.76 TeV: 81 points – 5.43 pb^{-1} – $74 < p_\perp < 592$ GeV

★ CMS 7 TeV: 158 points – 5.0 fb^{-1} – $74 < p_\perp < 2500$ GeV

★ CMS 8 TeV: 174 points – 19.7 fb^{-1} – $60 < p_\perp < 1300$ GeV

★ ATLAS 7 TeV: 140 points – 4.5 fb^{-1} – $100 < p_\perp < 2000$ GeV

★ ATLAS 8 TeV: 171 points – 20.2 fb^{-1} – $70 < p_\perp < 2500$ GeV

$\longrightarrow$ ***724*** points in total, v.s. ~ 4500 in global MSHT fit (inc.).

- We take the larger of the jet radii available in both cases, i.e. R=0.6/0.7.

See B. Bilin and P. Starovoitov talks

- Dijets:

    - ⋆ ATLAS 7 TeV: 90 points – 4.5 fb^{-1} – $\mathrm{d}^2\sigma/\mathrm{d}m_{jj}\mathrm{d}|y_{\rm max}|$ $0.26 < m_{jj} < 5.04\,\mathrm{TeV}$

    - ⋆ CMS 7 TeV: 54 points – 5.0 fb^{-1} – $\mathrm{d}^2\sigma/\mathrm{d}m_{jj}\mathrm{d}|y^*|$ $0.25 < m_{jj} < 4.48\,\mathrm{TeV}$

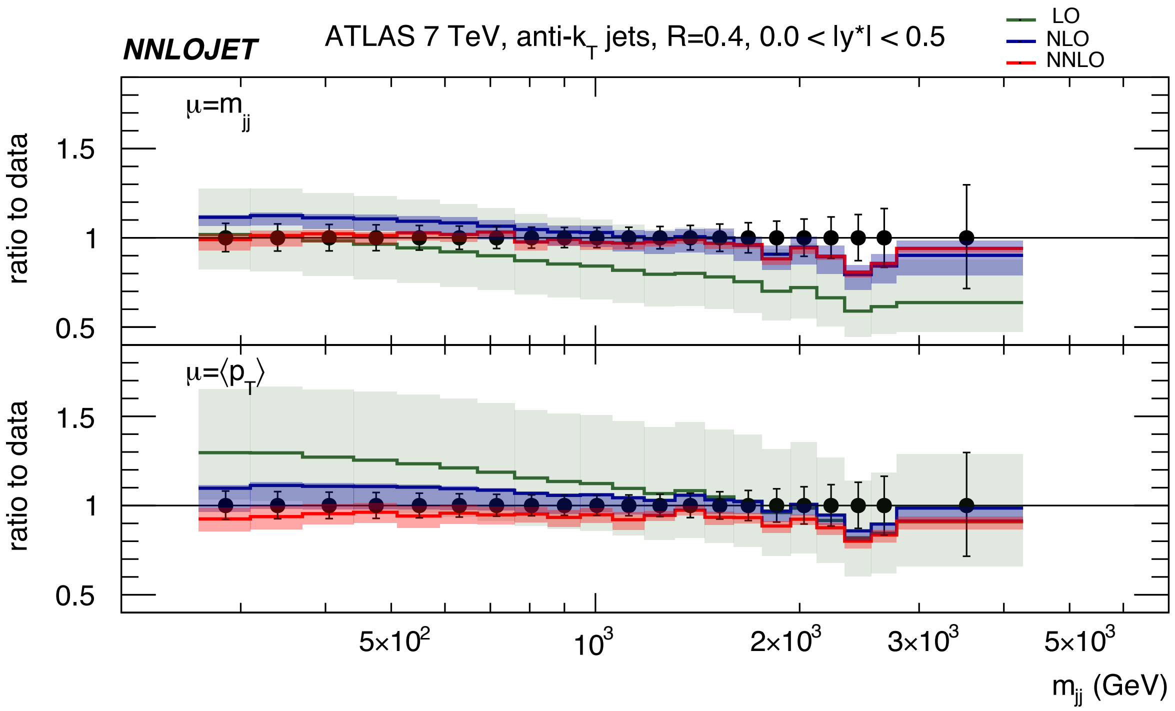
    - ⋆ CMS 8 TeV: 122 points – 19.7 fb^{-1} – $\mathrm{d}^3\sigma/\mathrm{d}p_{\perp,avg}\mathrm{d}y_b\mathrm{d}y^*$ $143 < p_{\perp,avg} < 1638\,\mathrm{GeV}$

$\longrightarrow$ **266** points in total, v.s. ~ 4000 in global MSHT fit (inc.).

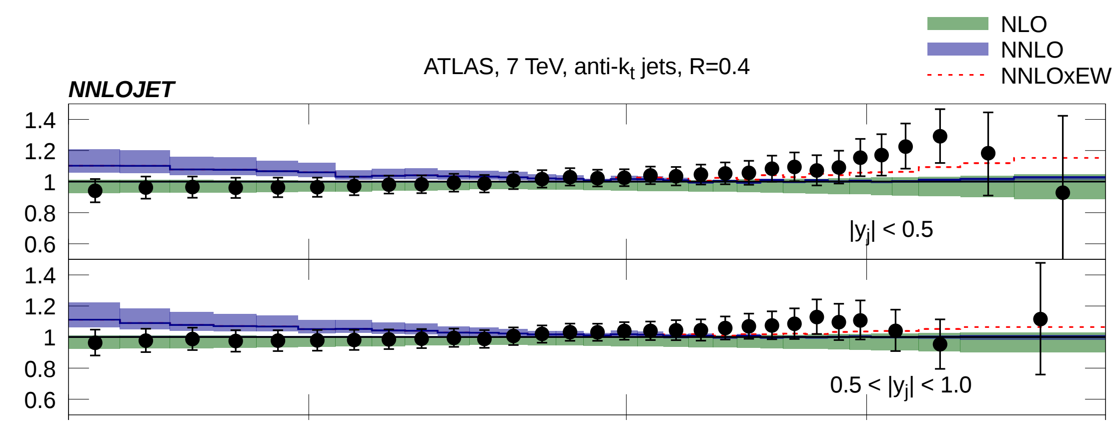
- Again take the larger of the jet radii available in both cases, i.e. R=0.6/0.7.

Fit Quality

- Will consider fits either to 7 + 8 TeV inclusive jets, or to 7 + 8 TeV dijets.
- MSHT20 baseline, and NNLO QCD + NLO EW, unless otherwise stated. Theory from **arXiv:2005.11327** + NNLOJET.
- For inclusive jets take $\mu = p_\perp^j$, for dijets $\mu = m_{jj}$.
- Various benchmarks/checks performed, but results presented here **preliminary**.



J. Currie et al., *Phys.Rev.Lett.* 119 (2017) 15, 152001



J. Currie, E.W.N. Glover, J. Pires, *Phys.Rev.Lett.* 118 (2017) 7, 072002

- We find:

Dijet fit:

	$N_{\rm pts}$	$\chi^2/N_{\rm pt}$
ATLAS 7 TeV dijets	90	1.05
CMS 7 TeV dijets	54	1.43
CMS 8 TeV dijets	122	1.04
Total Dijets	266	1.12

	$N_{\rm pts}$	$\chi^2/N_{\rm pt}$
ATLAS Z $p_\perp$	104	1.65
Diff. top	54	1.24
7 + 8 TeV Jets	643	[1.62]

Prediction

Jet fit:

	$N_{\rm pts}$	$\chi^2/N_{\rm pt}$
ATLAS 7 TeV jets	140	1.53
ATLAS 8 TeV jets	171	1.45
CMS 7 TeV jets	158	1.22
CMS 8 TeV jets*	174	1.80
Total Jets	643	1.50

	$N_{\rm pts}$	$\chi^2/N_{\rm pt}$
ATLAS Z $p_\perp$	104	1.85
Diff. top	54	1.12
7 + 8 TeV Dijets	643	[1.32]

★ Fit quality to dijet data very good (1.12), clearly worse for jets (1.50).

★ No signs of significant inconsistency in fit vs. predicted χ^2, though some difference in pull implied.

★ Fit quality to top (Z $p_\perp$) data better in jet (dijet) fit. Latter particularly notable.

★ (Not shown) - fit quality to other data in global fit v. similar.

*NB we use stat. correlations here. Not included by other groups, and leads to deterioration in fit quality.

Impact of HO corrections

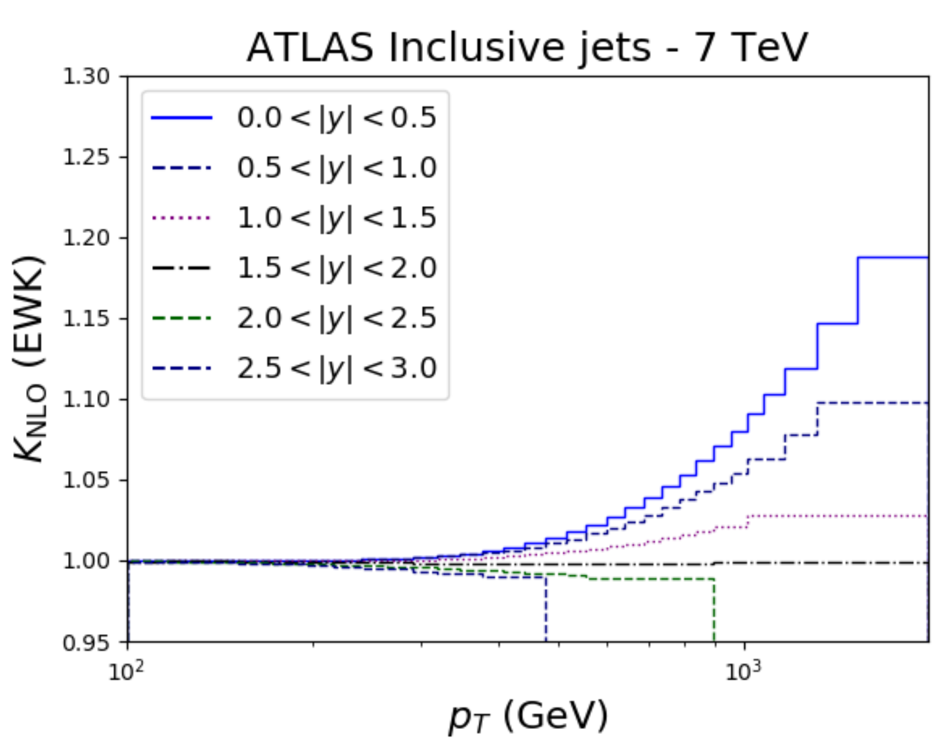
★ **EW** corrections:

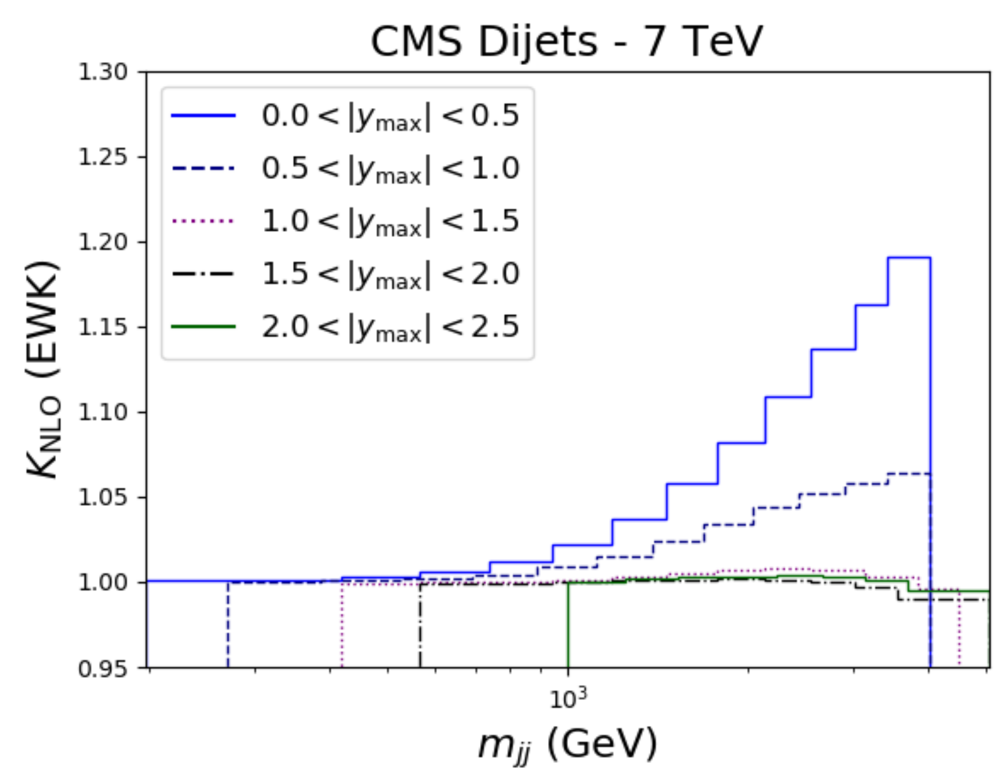
Dijet fit: $\chi^2(\text{no EW}) \to \chi^2(\text{EW}) : 1.34 \to 1.12$

Jet fit: $\chi^2(\text{no EW}) \to \chi^2(\text{EW}) : 1.39 \to 1.50$

(NNLO QCD)

i.e. we find dijet fit quality improved (driven by CMS 8 TeV), but inclusive (uniformly) deteriorates! Unclear why, but clearly impacts on discussion of relative fit quality.





R. Abdul Khalek et al., *Eur.Phys.J.C* 80 (2020) 8, 797

★ **NNLO QCD** corrections.

Jets fit:

	N_{pts}	NLO	NNLO
ATLAS 7 TeV jets	140	1.69	1.53
ATLAS 8 TeV jets	171	2.37	1.45
CMS 7 TeV jets	158	1.38	1.22
CMS 8 TeV jets	174	1.65	1.80
Total Jets	643	1.78	1.50

Dijet fit:

	N_{pts}	NLO	NNLO
ATLAS 7 TeV dijets	90	1.10	1.05
CMS 7 TeV dijets	54	1.71	1.43
CMS 8 TeV dijets	122	5.30	1.04 ← **Not a typo!**
Total Dijets	266	3.15	1.12

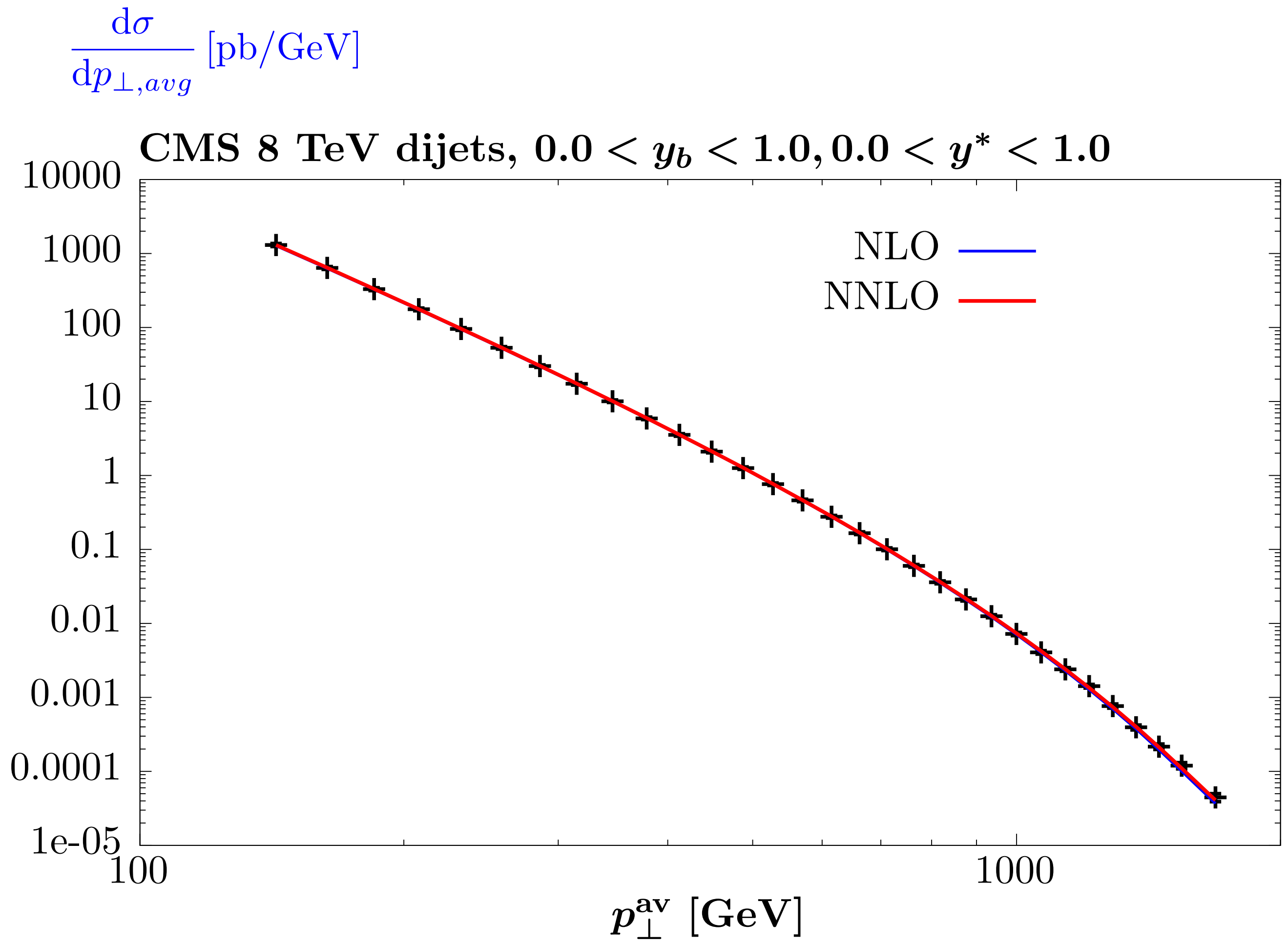
★ Clear trend in both cases for QCD corrections to improve fit quality. pQCD working as it should!

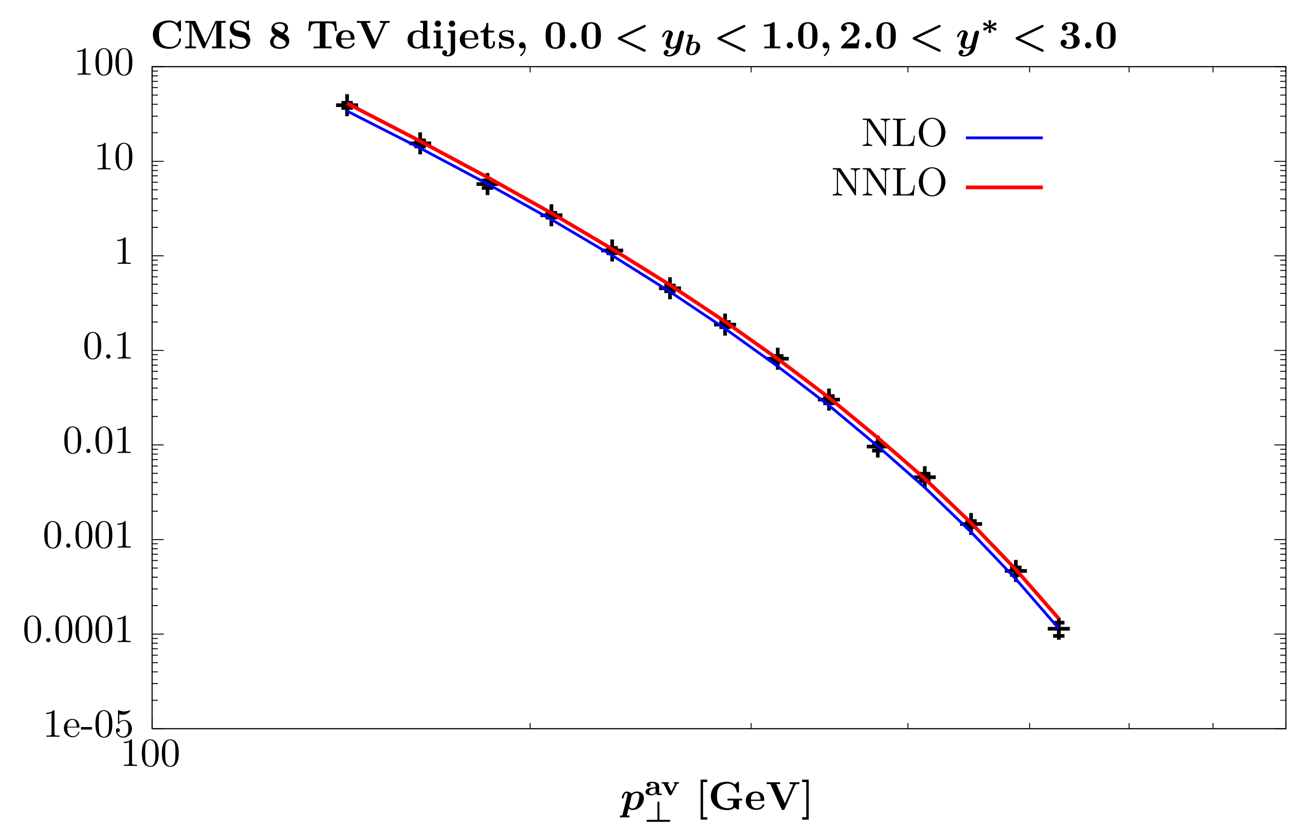
★ For jets, this is different to arXiv:2005.11327 trend (though same as in MSHT20), but note scale different ($p_\perp^j$ rather than H_T).

★ Improvement in CMS 8 TeV dijets particularly remarkable. Clear need for NNLO QCD at high precision LHC (c.f. e.g. ATLAS 7 TeV, W,Z, which gives 5.0 $\rightarrow$ 1.9). In more detail…

CMS 8 TeV Dijets

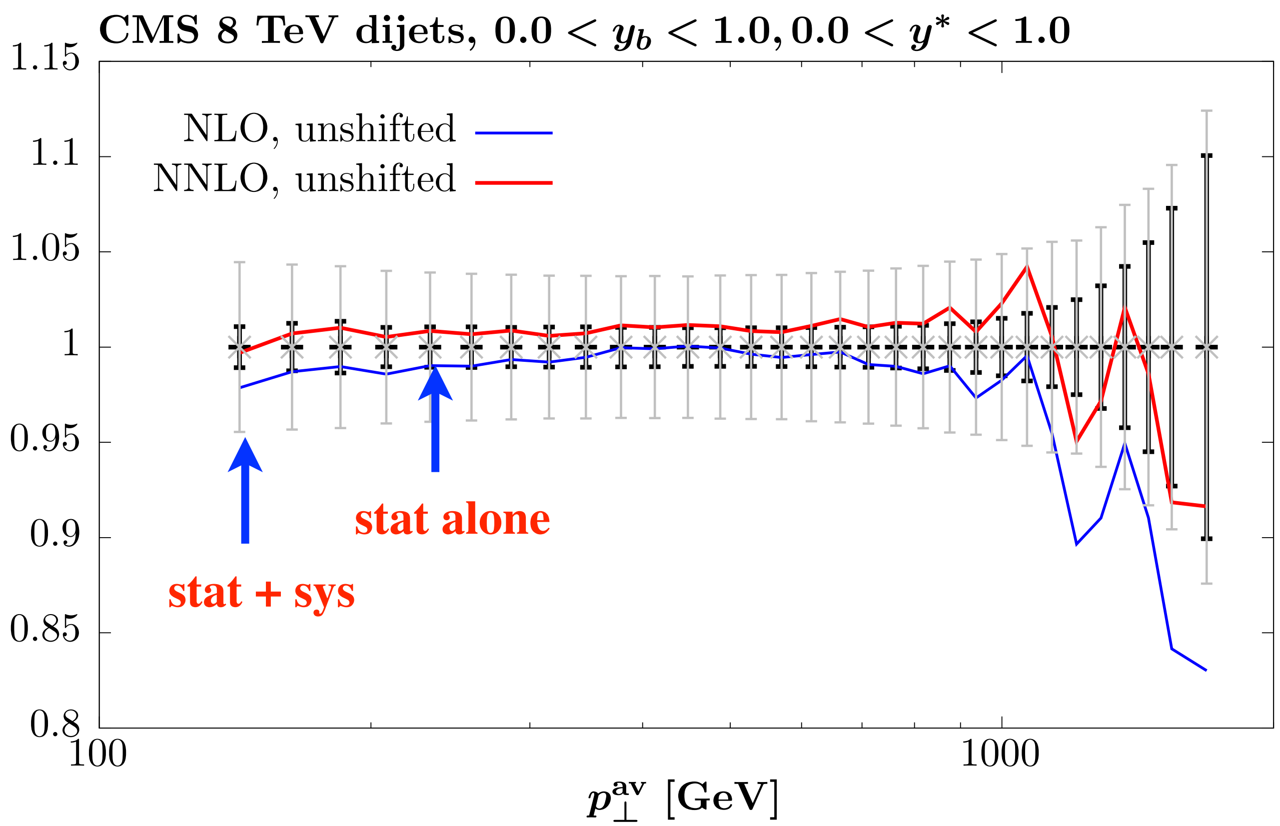
- What is driving this improvement?

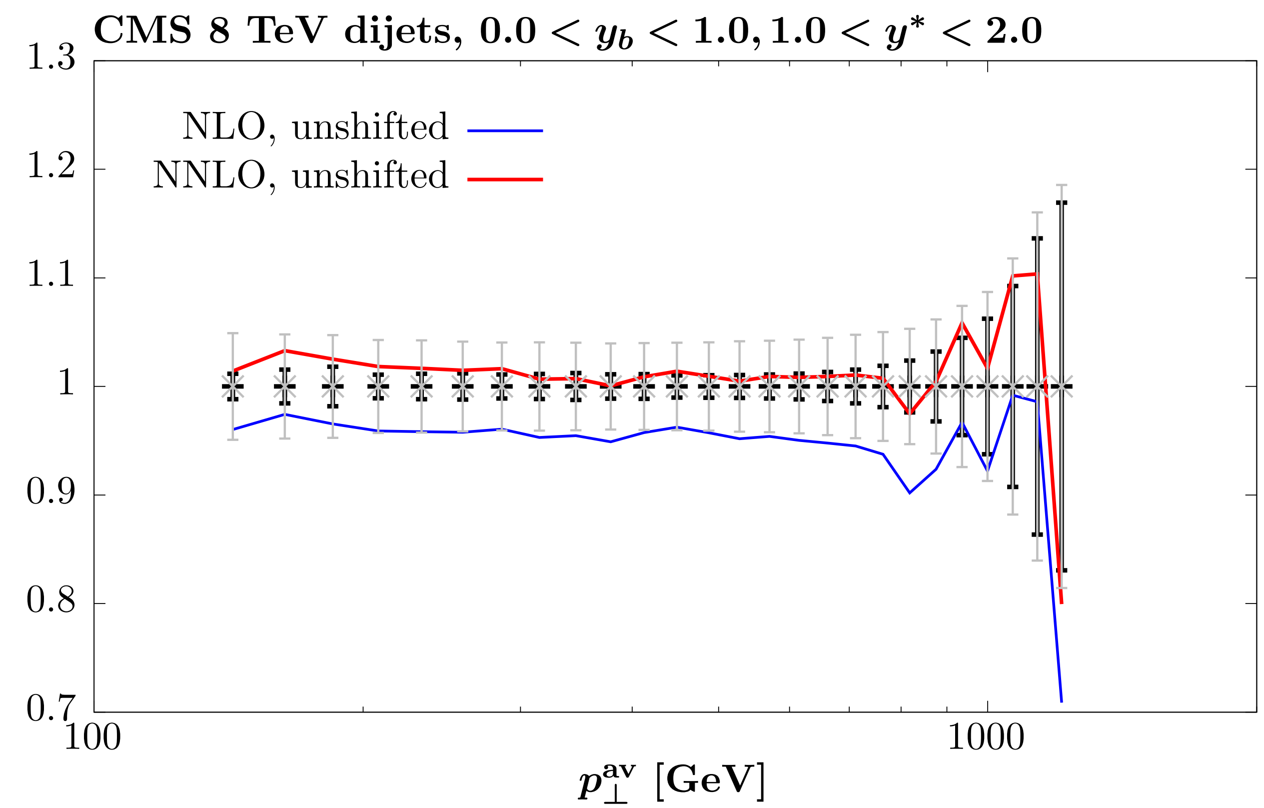


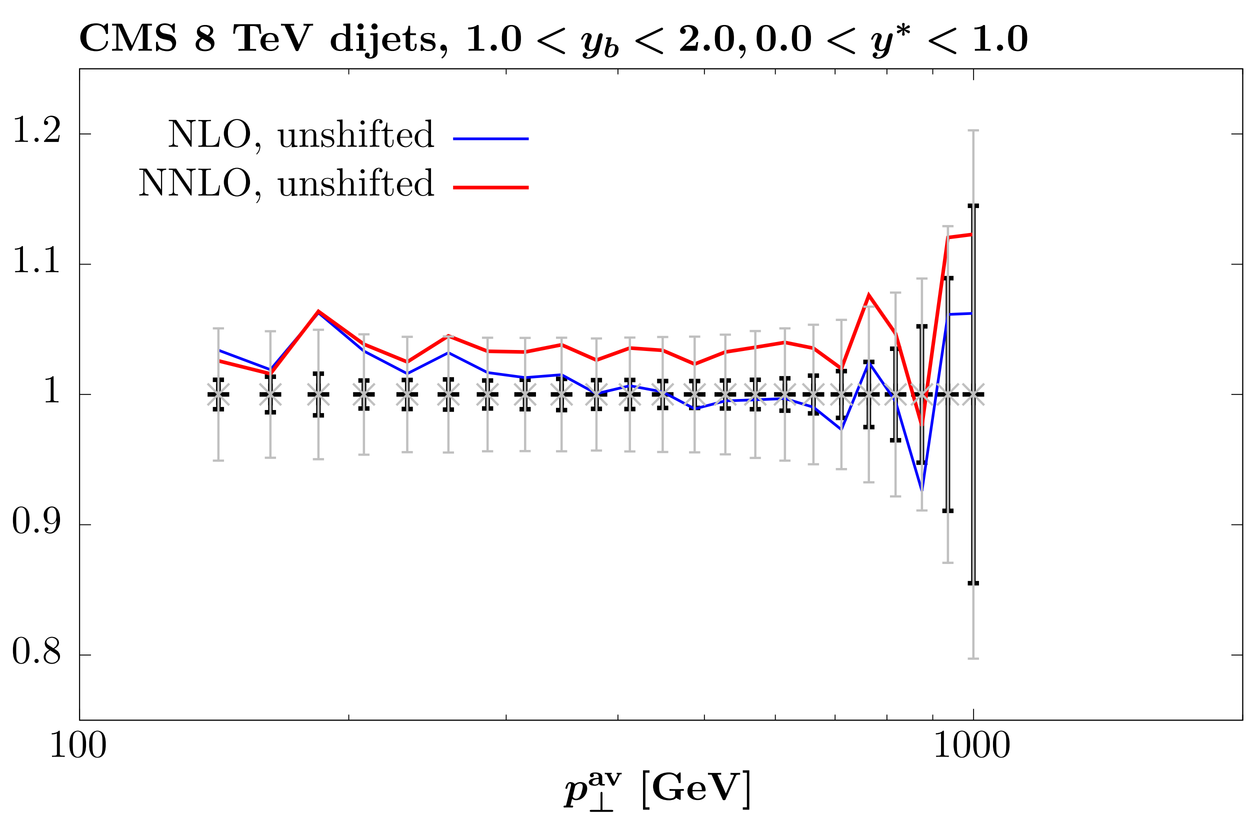


- Focus in on data/theory comparison…

- Overlaying data/theory no clear, by eye, trend for better description at NNLO.

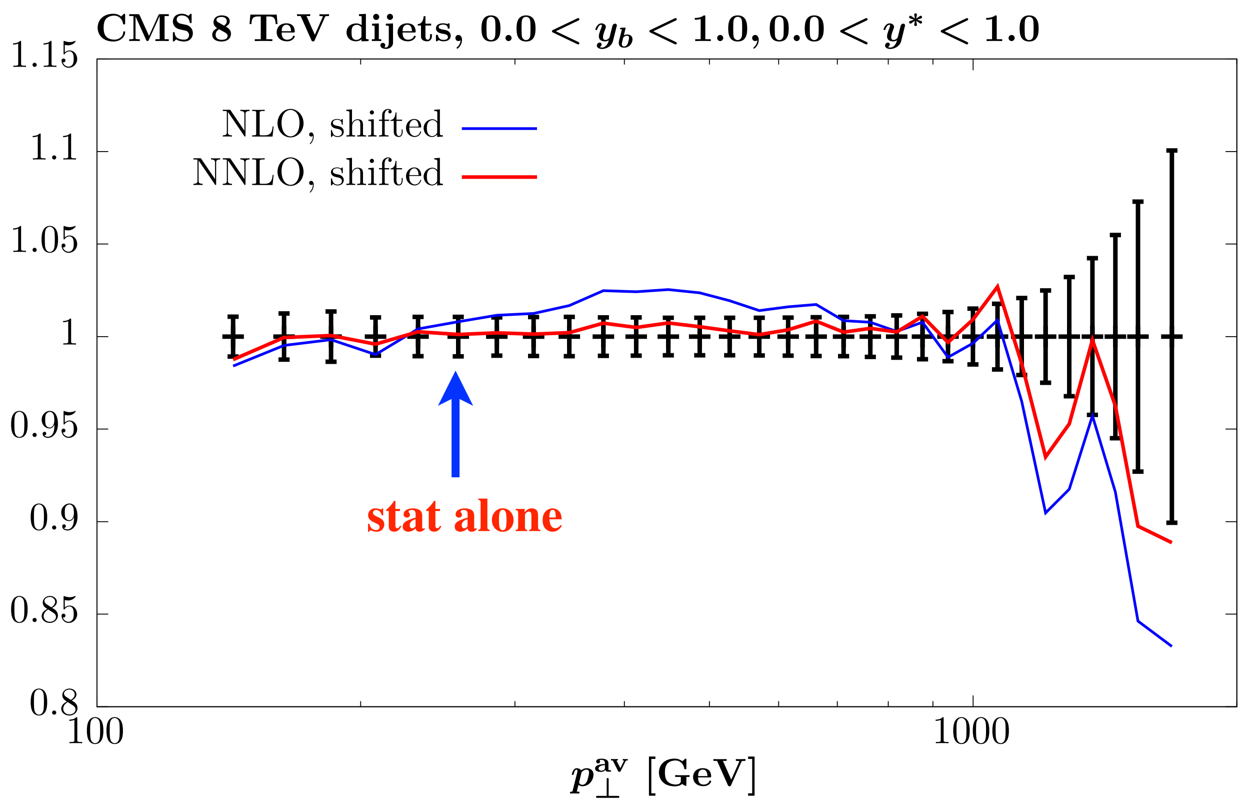


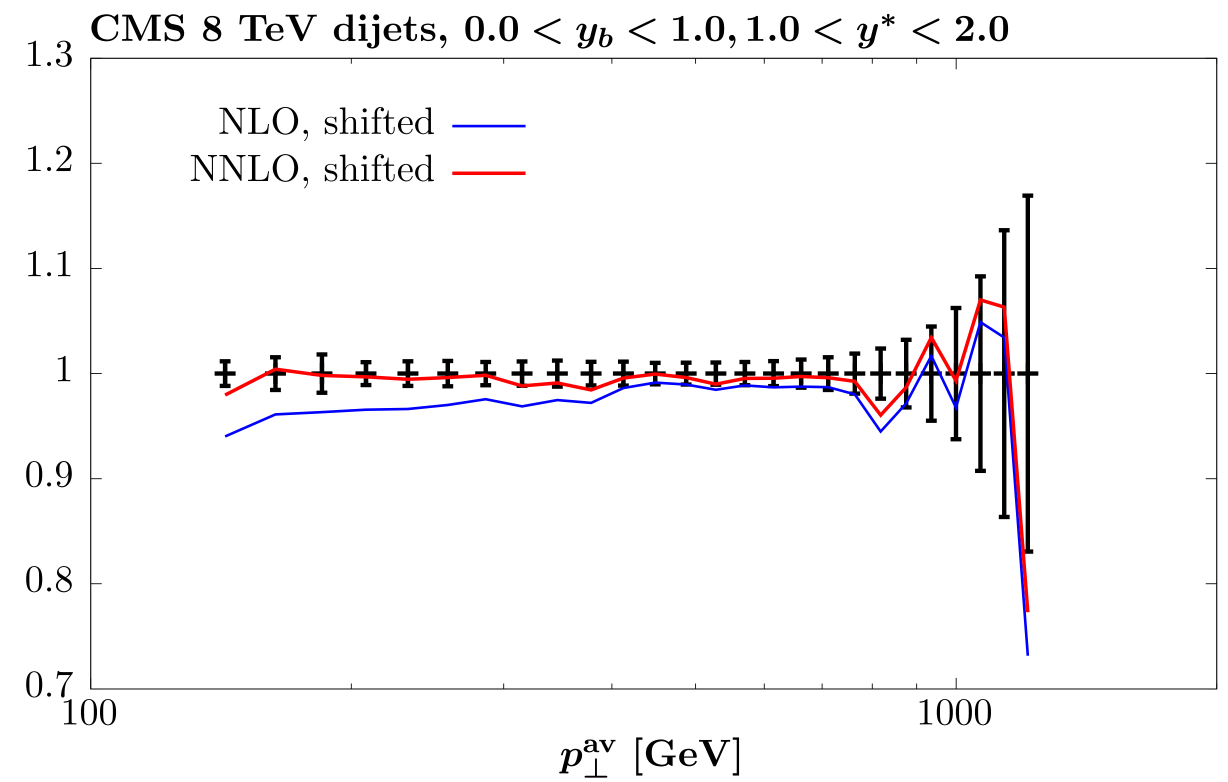


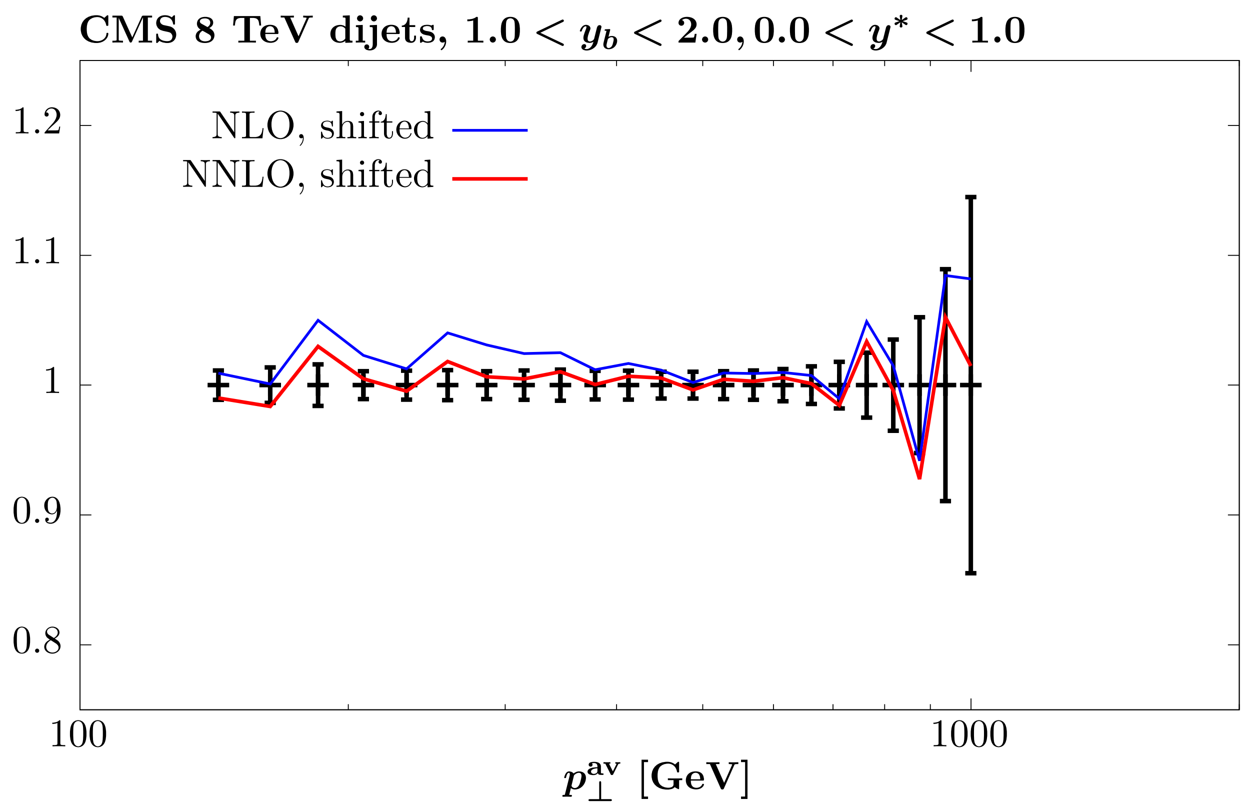


- However this is **before** shifting by correlated systematics.

- After including shifts from correlated systematics:



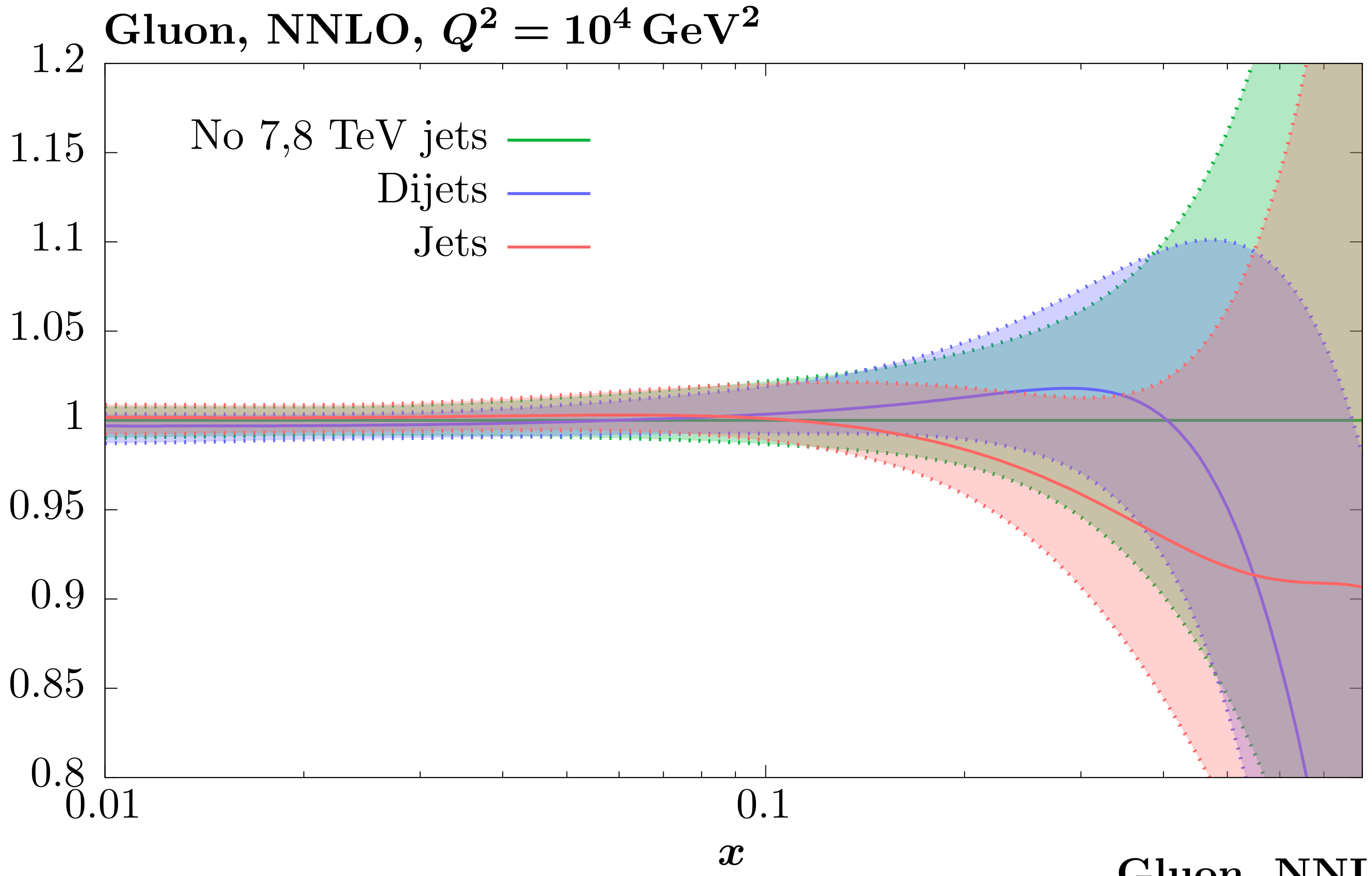




- Now improvement of NNLO is clear.
- Impact on shape of distributions in 3D kinematic space and interplay with correlated systematics drives this.

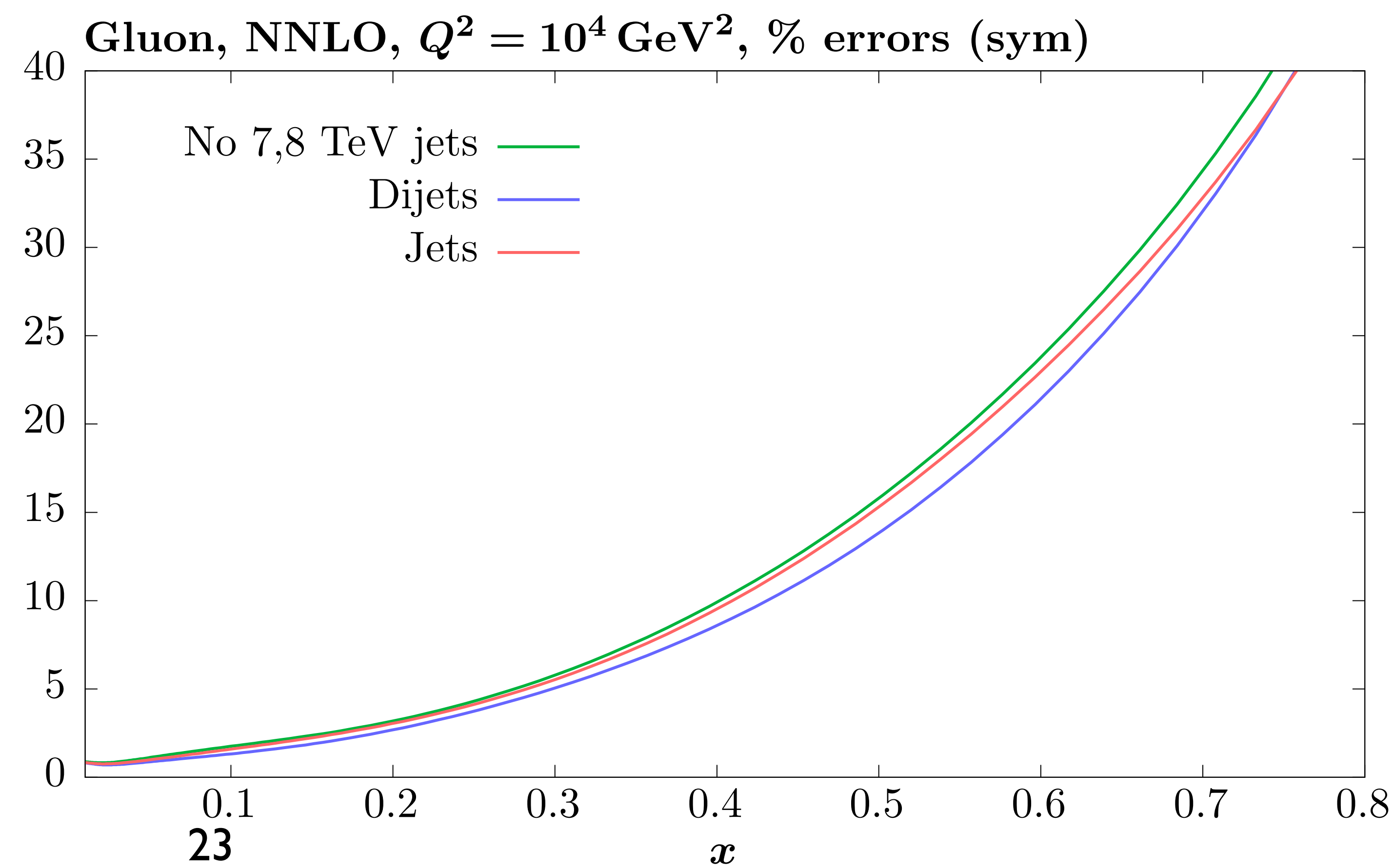
Impact on PDFs

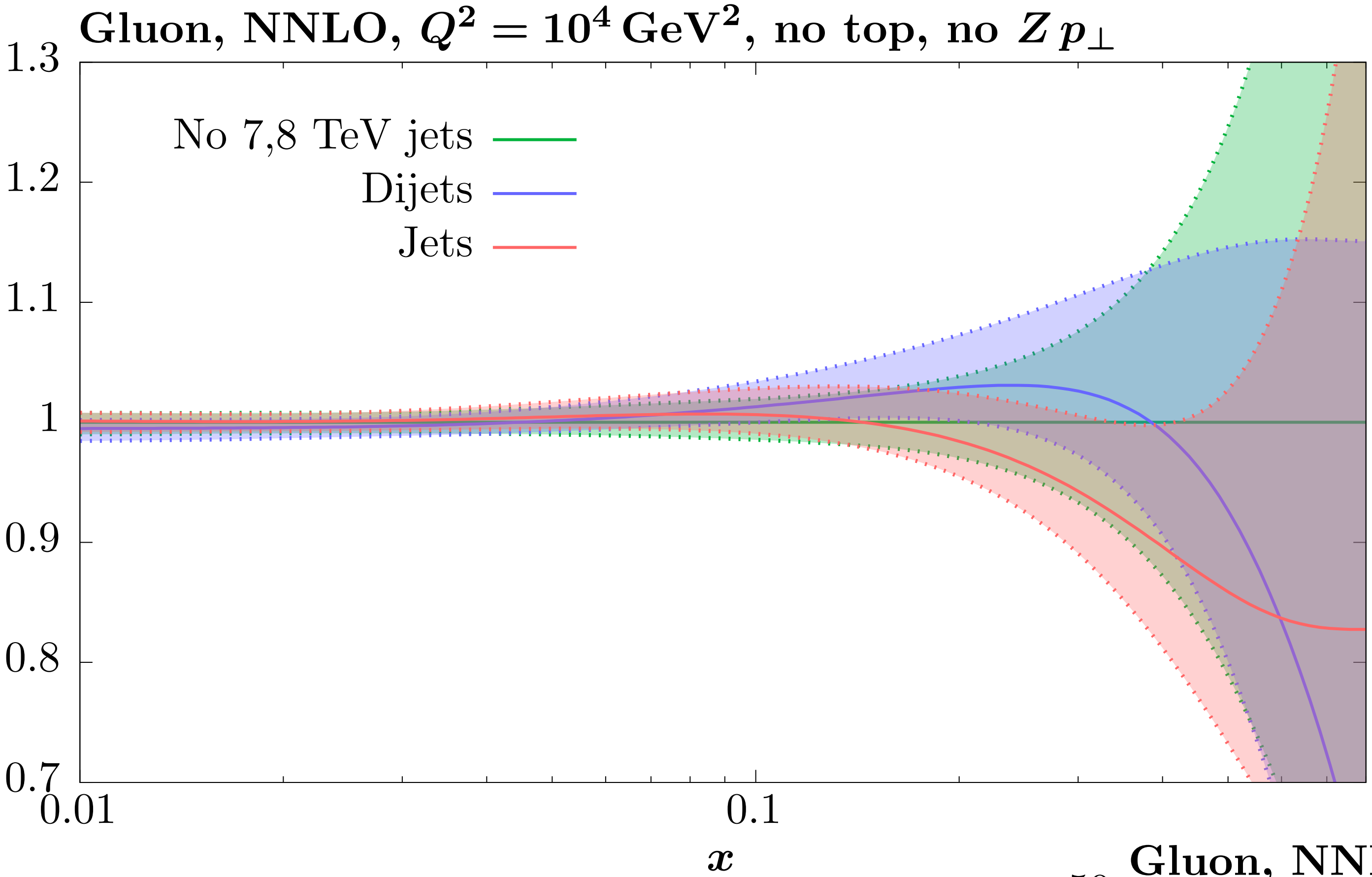
Dijets vs. Jets



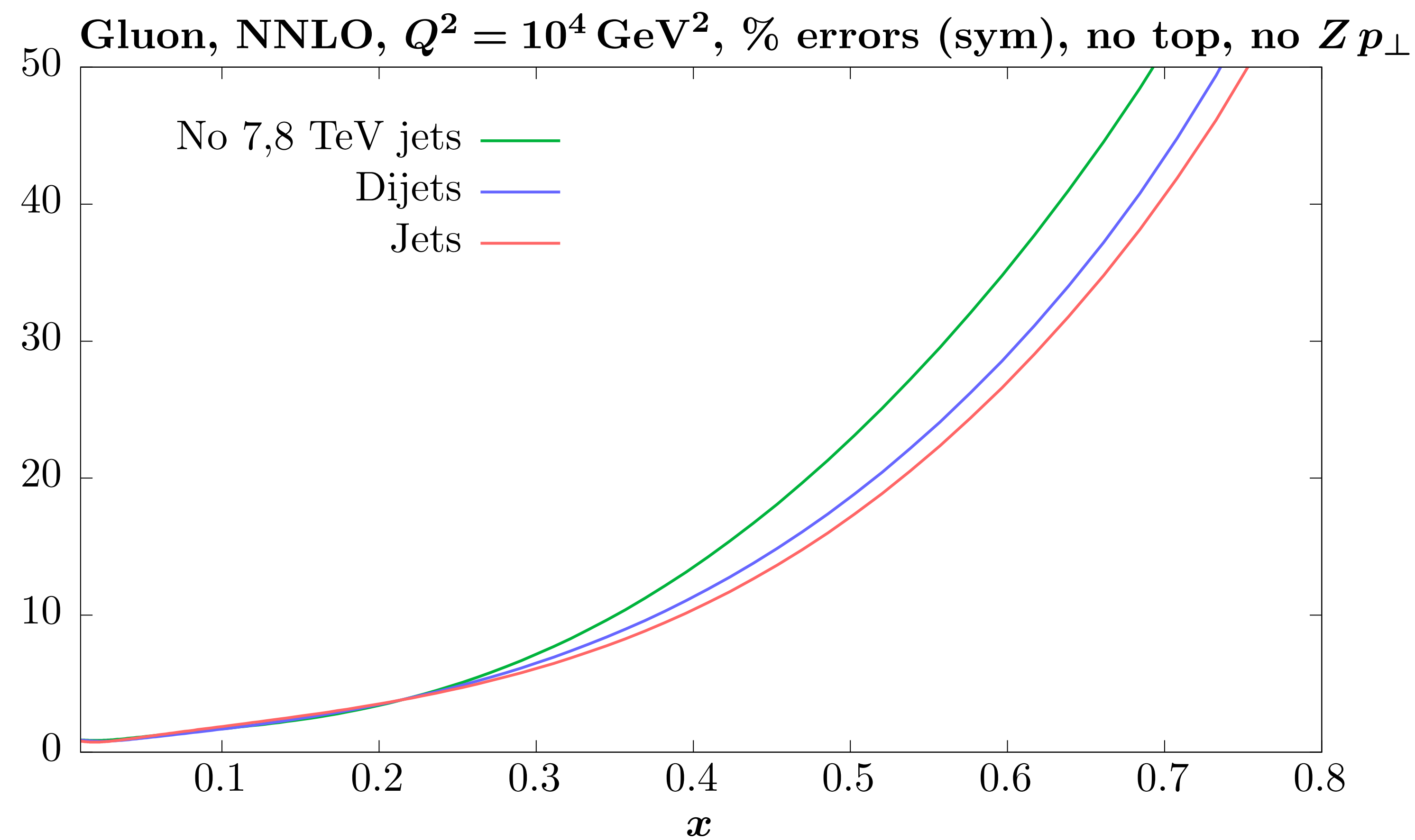
- Add dijets or jets to MSHT20 (no jets) baseline. Focus on gluon here.
- Overall consistency between two cases.
- But pull **rather different**.

- Impact of jets data on gluon uncertainty very mild. Larger for dijets.



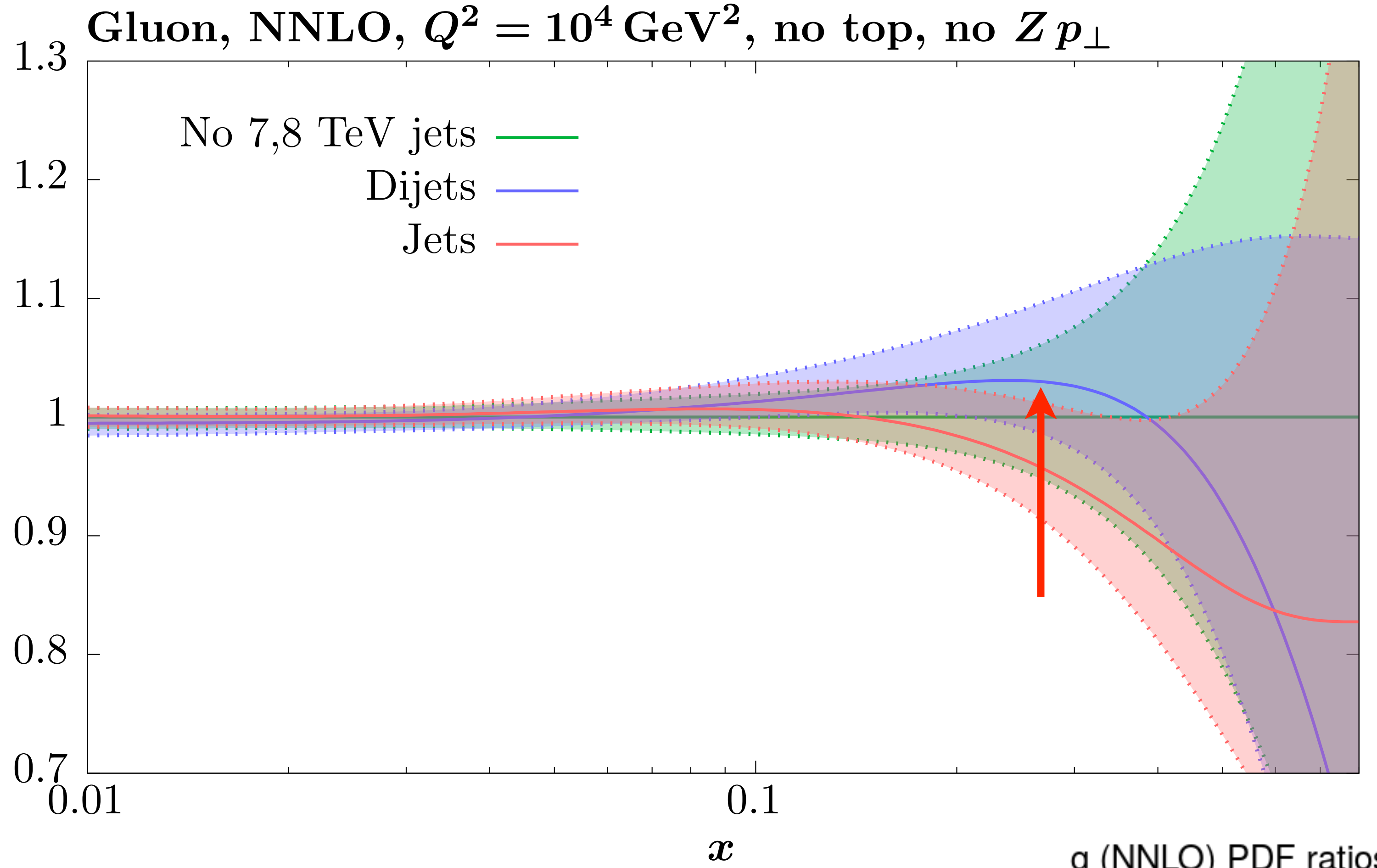


- Now consider fits but added to MSHT20 baseline with no diff. top or $Z\,p_\perp$.
- Basic pulls as before.

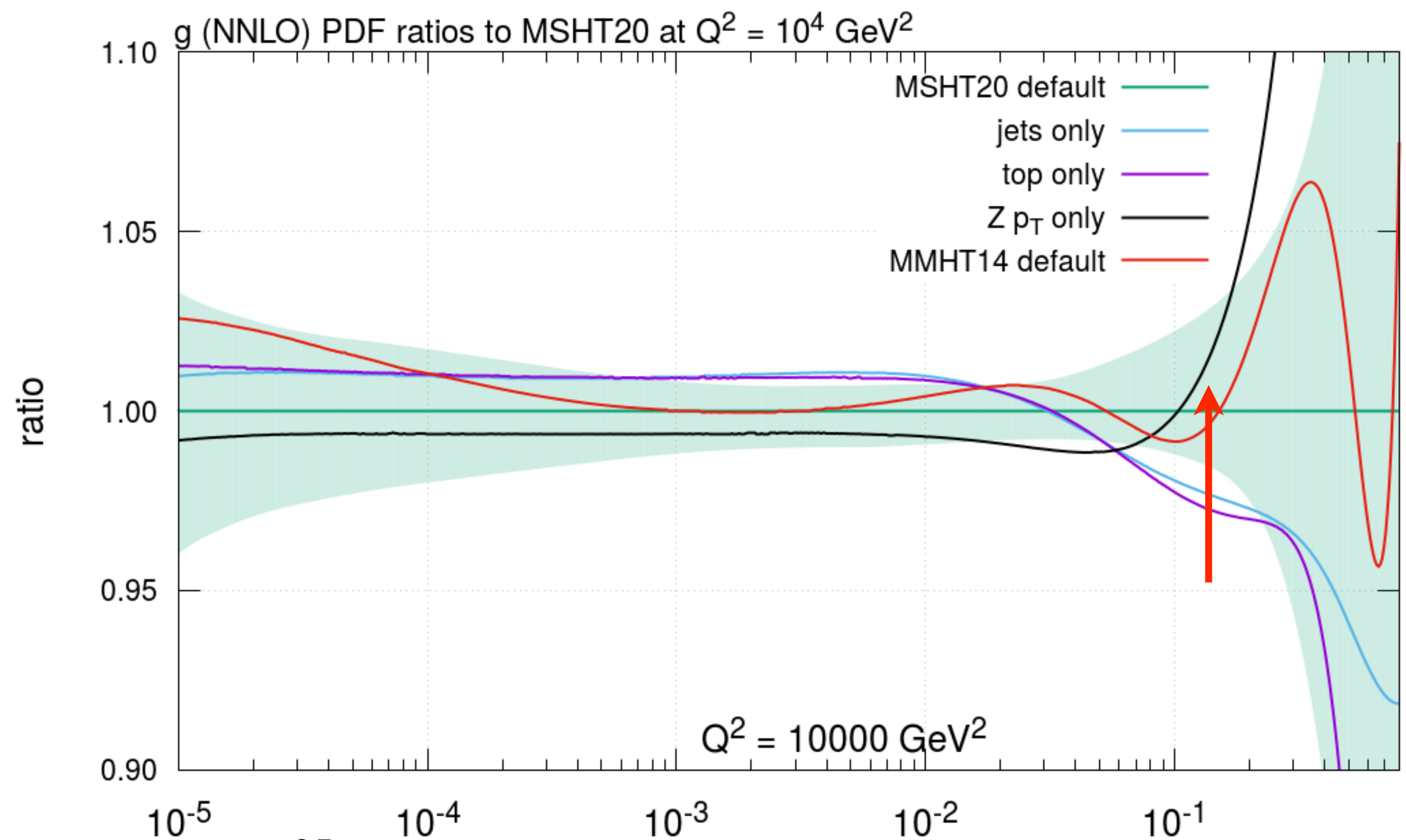


- However relative impact on uncertainty different. Now jets more significant at high x .

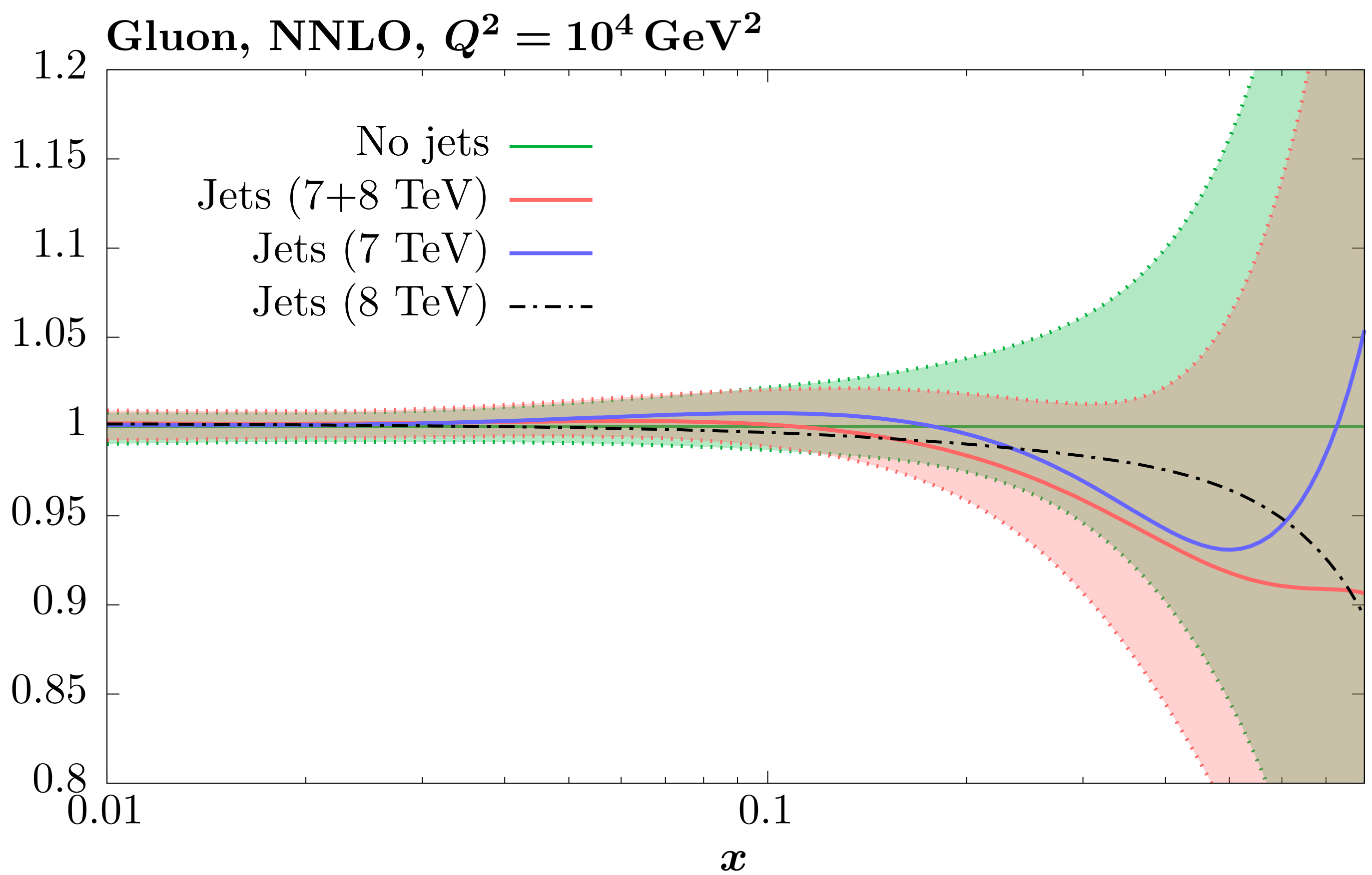
$\longrightarrow$ Clear that interplay with other data in fit drives impact. In particular greater consistency between dijets and $Z\,p_\perp$.



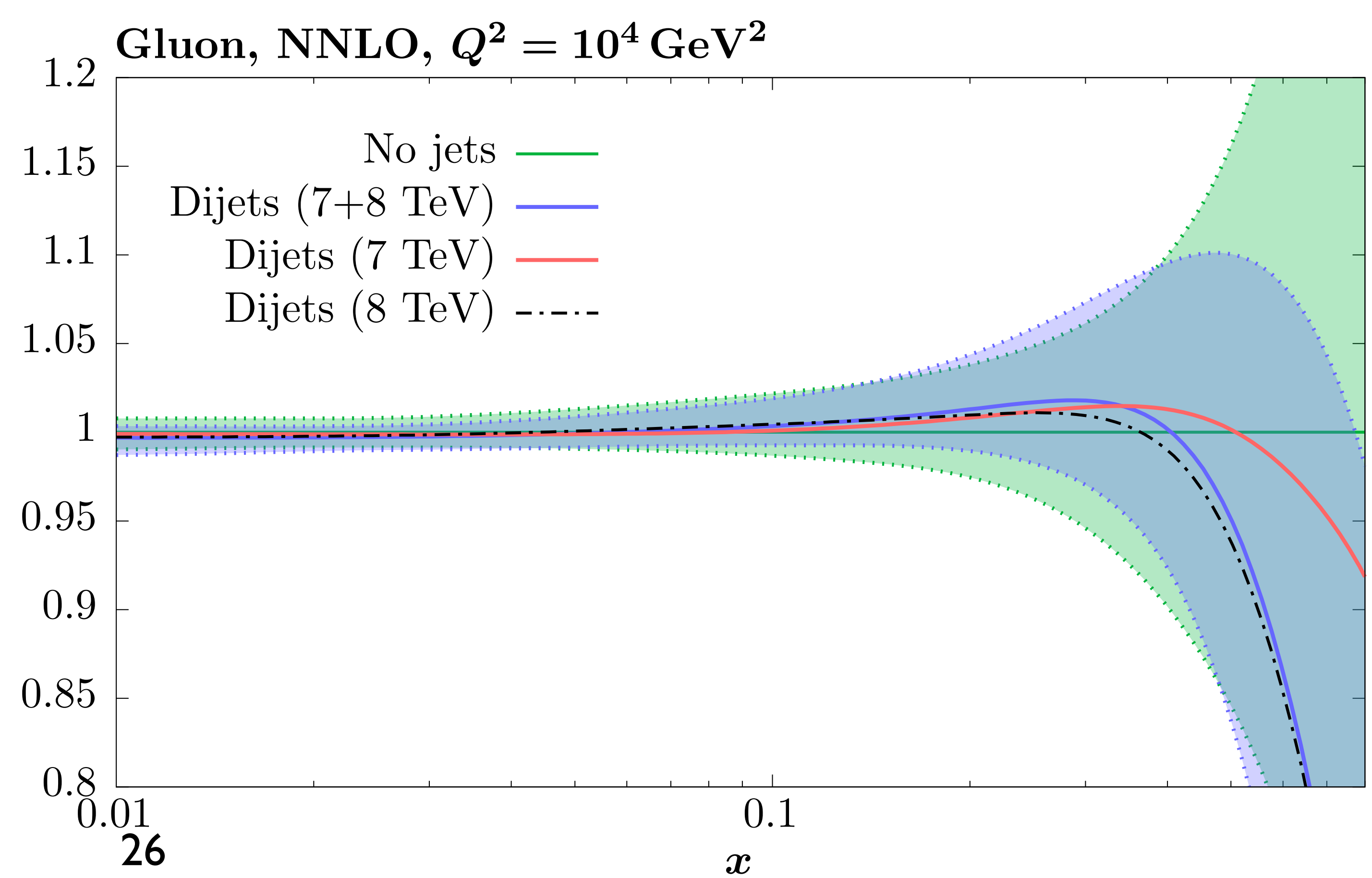
- In more detail, Dijets and both prefer rather higher $Z\,p_\perp$ gluon in $0.1 \lesssim x \lesssim 0.4$ region ($Z\,p_\perp$ out to higher x).



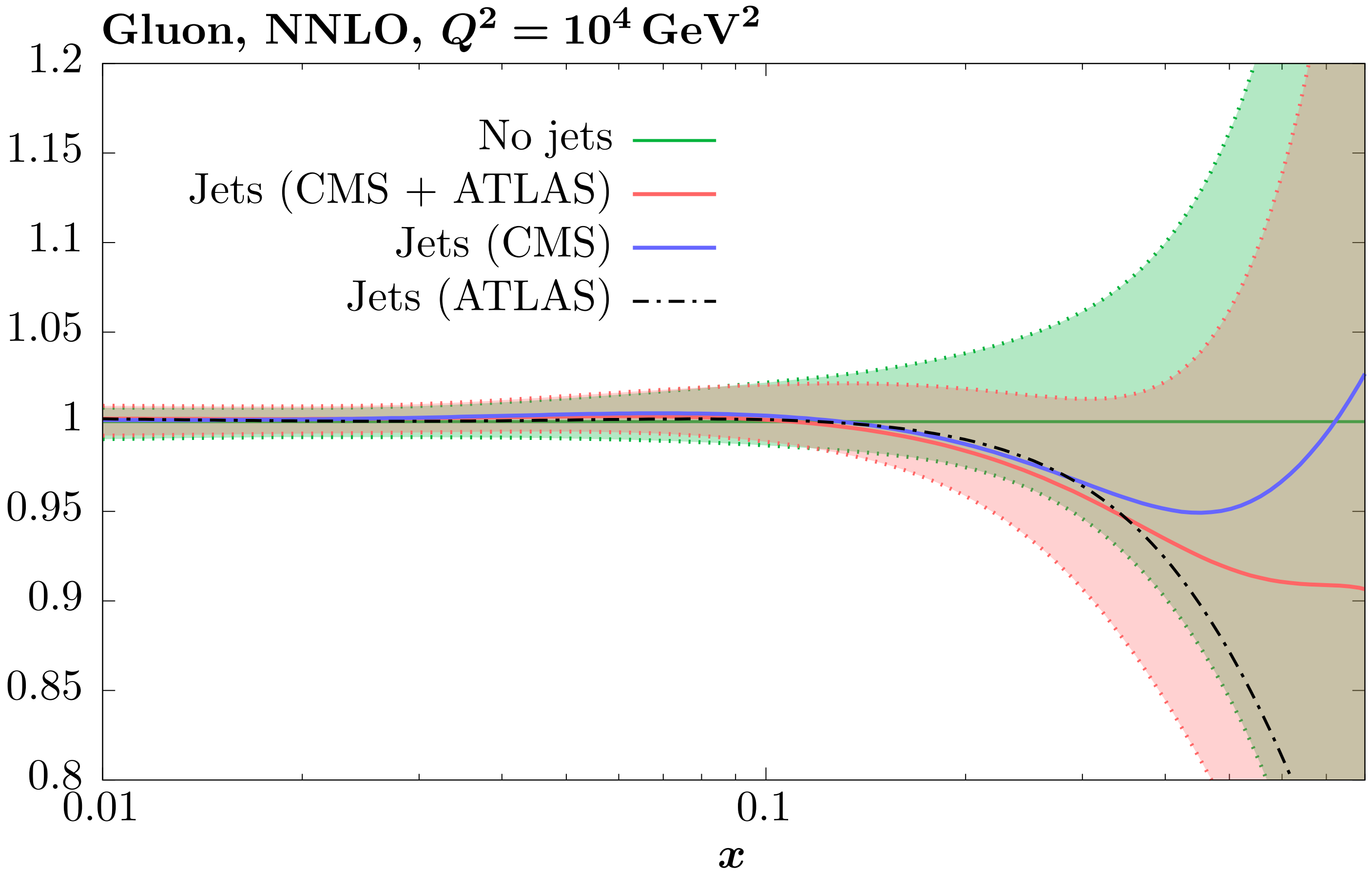
Consistency within datasets



- *7* & 8 TeV data ~ consistent for inclusive jets.

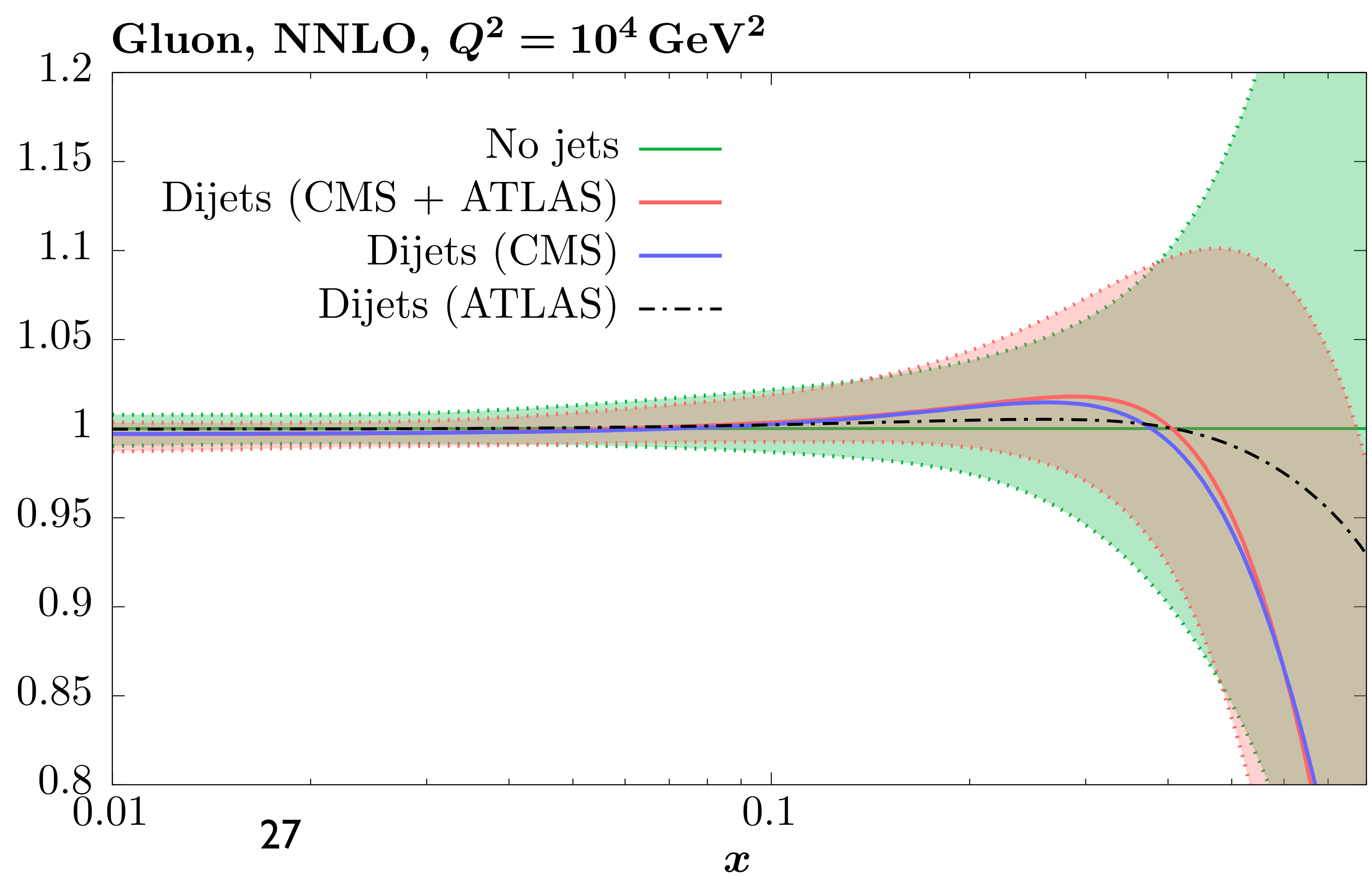


- *7* & 8 TeV data consistent for dijets, but this is due to broader result.
- That is: all dijet fits completely driven by CMS 8 TeV data.



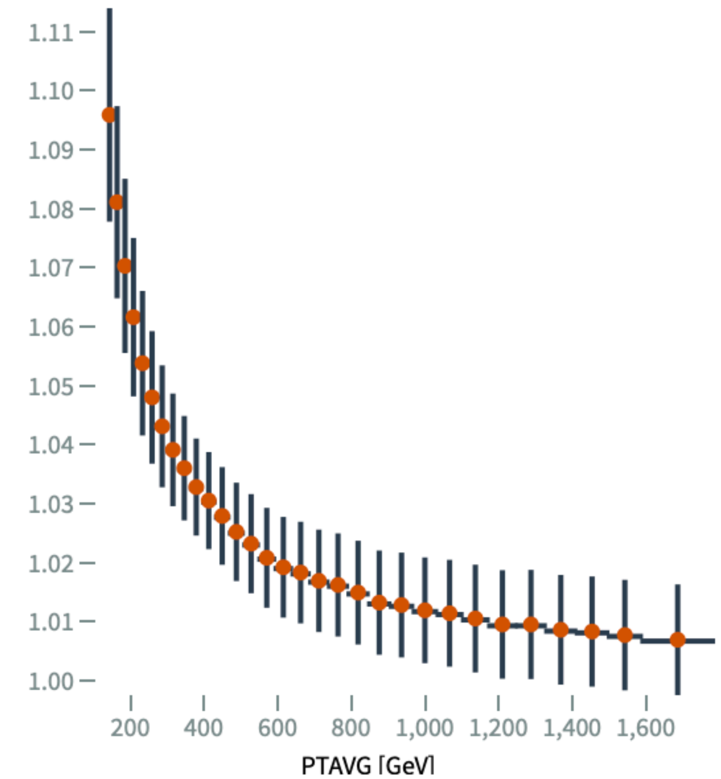
- At higher x clear difference between pulls of ATLAS and CMS (also seen in MSHT20).
- Final result compromise between these.

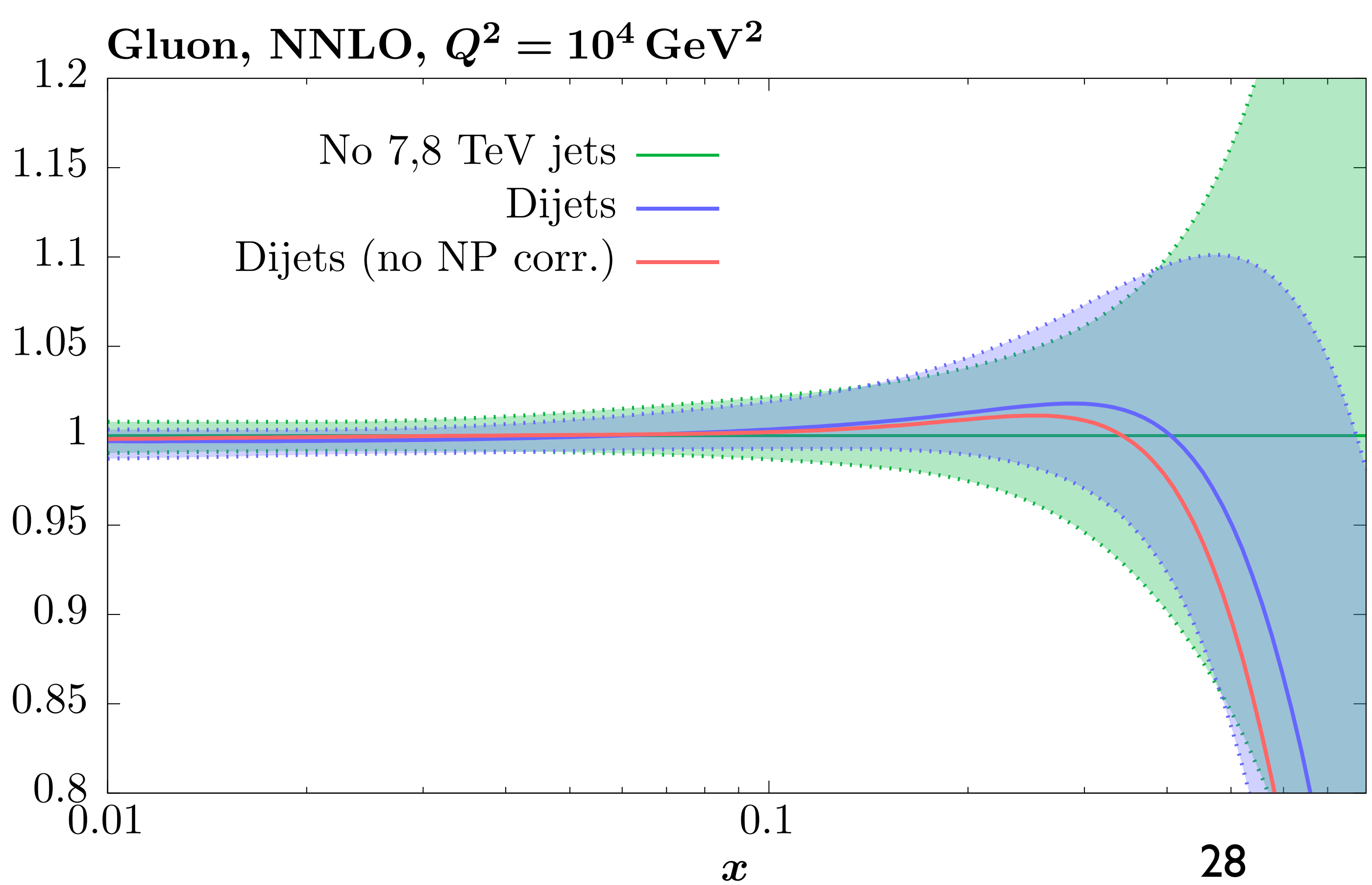
- Consistency between CMS and ATLAS, but latter has very little impact alone.
- Again CMS 8 TeV driving fit.

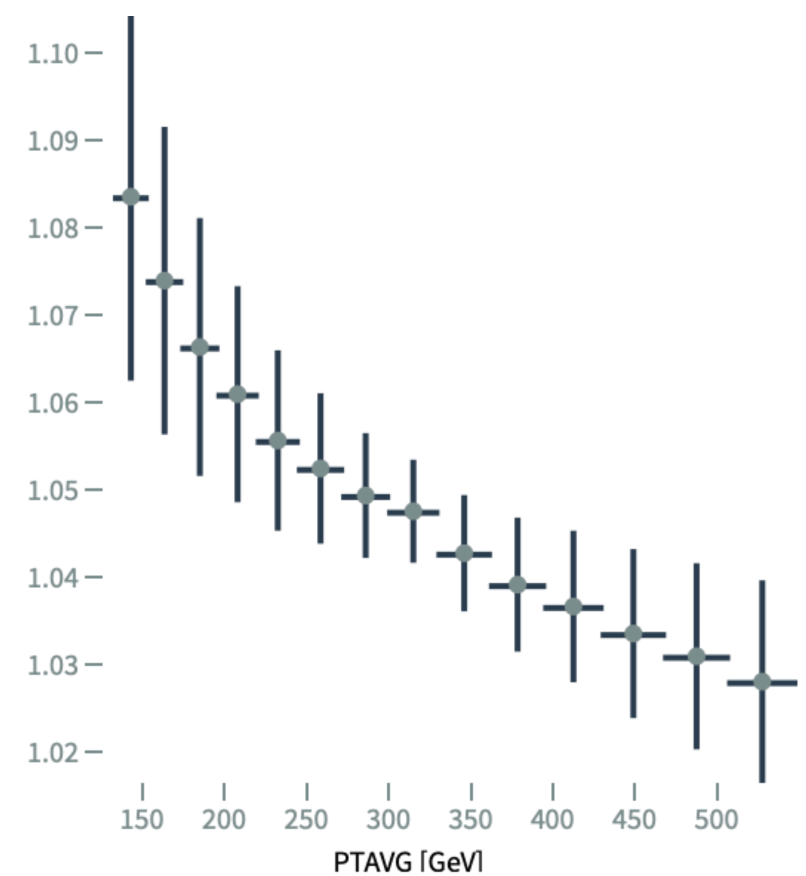


Technical aside (1) - NP corrections

- Not just high precision QCD/EW needed.
- Nonperturbative corrections enter at same level ~ NNLO QCD corrections. Percent level (two-point) uncertainties.
- Turning off in dijet fit, χ^2 deteriorates from 1.12 to 1.40. Some impact on gluon.

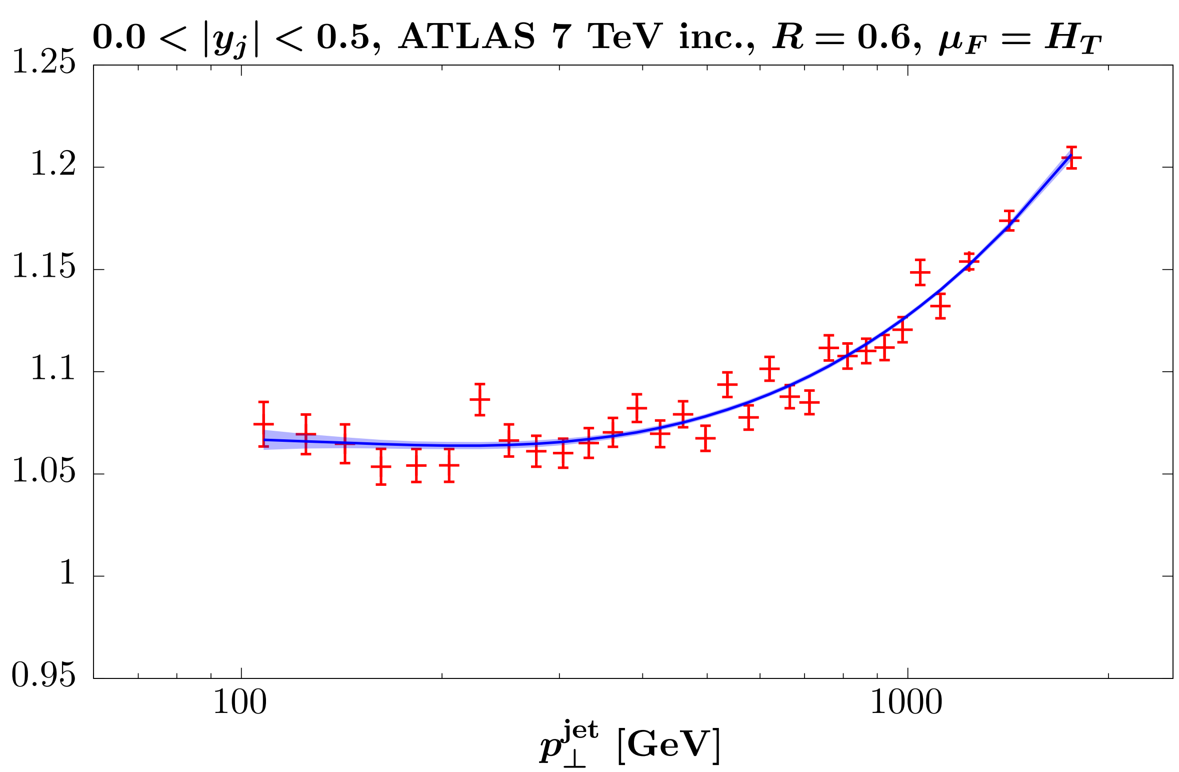


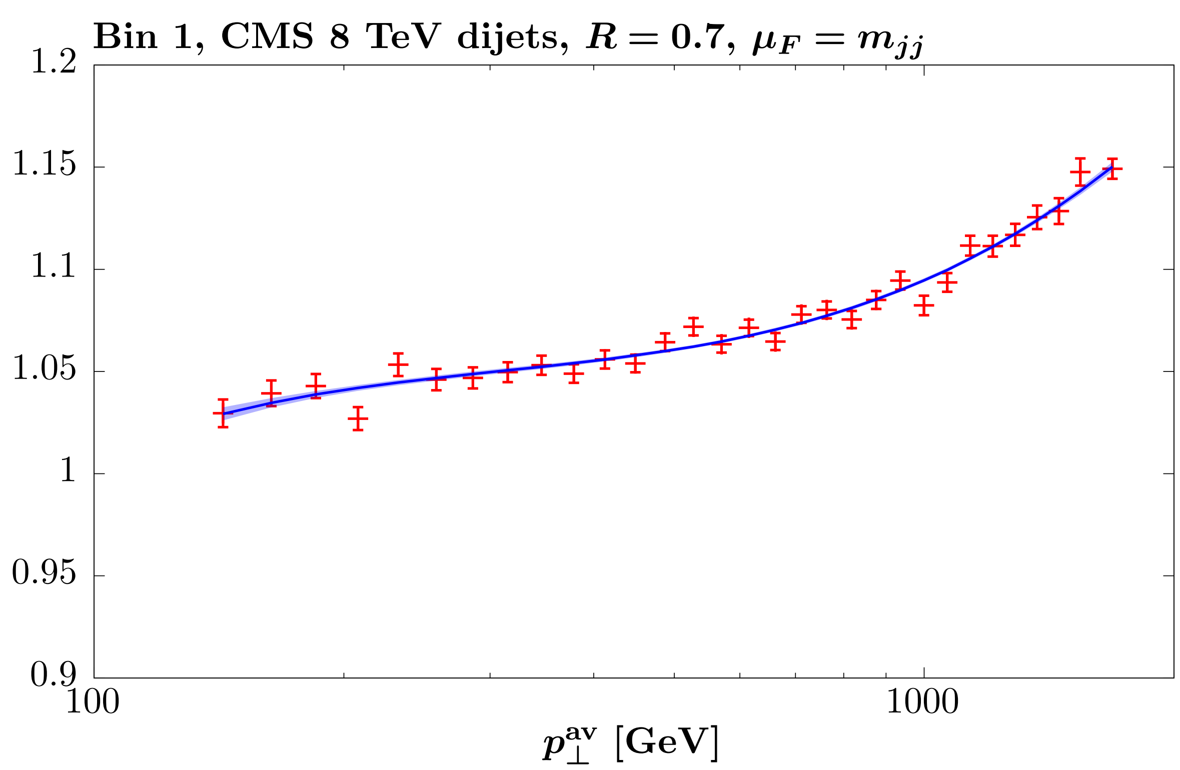




Technical aside (2) - K-factors

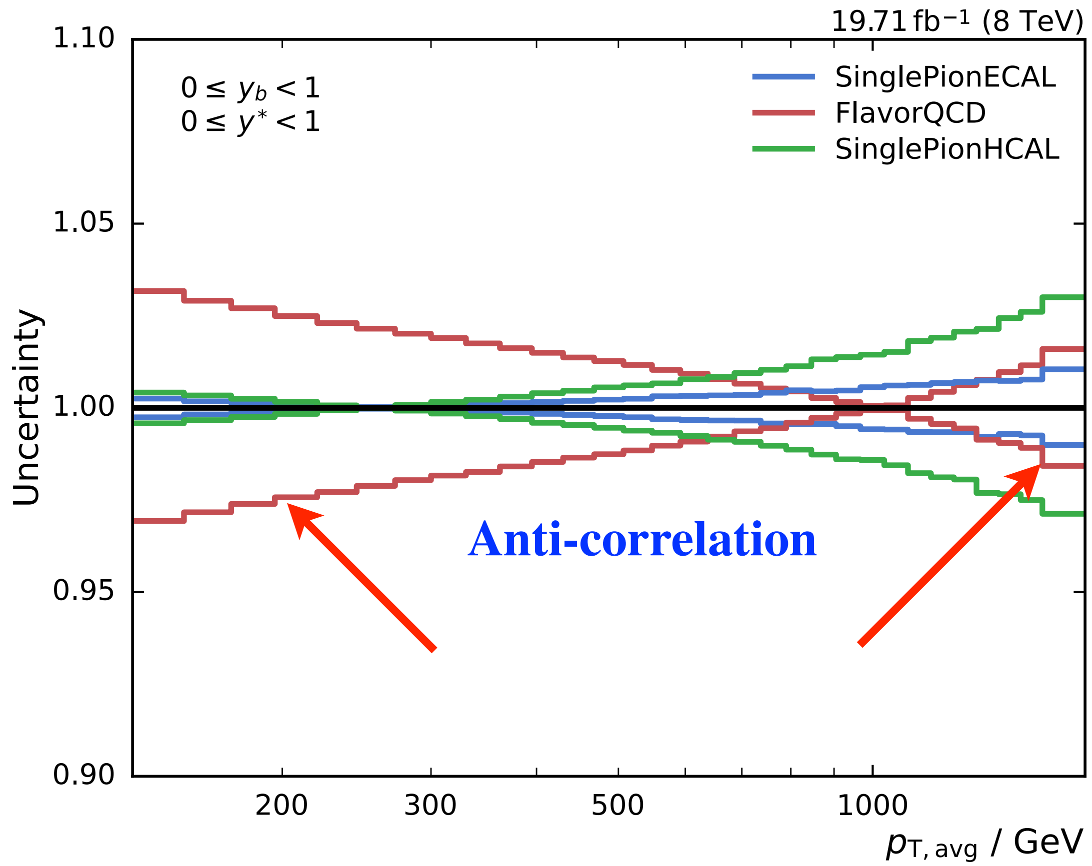
- NNLO QCD corrections included via K-factors. MC uncertainties on these not negligible.
- We argue better to fit these to smooth functions. Can impact on fit quality at the ~ 0.1-0.2 per point level, though PDFs very stable.
- Provides cleaner idea of improvements from NLO to NNLO etc.

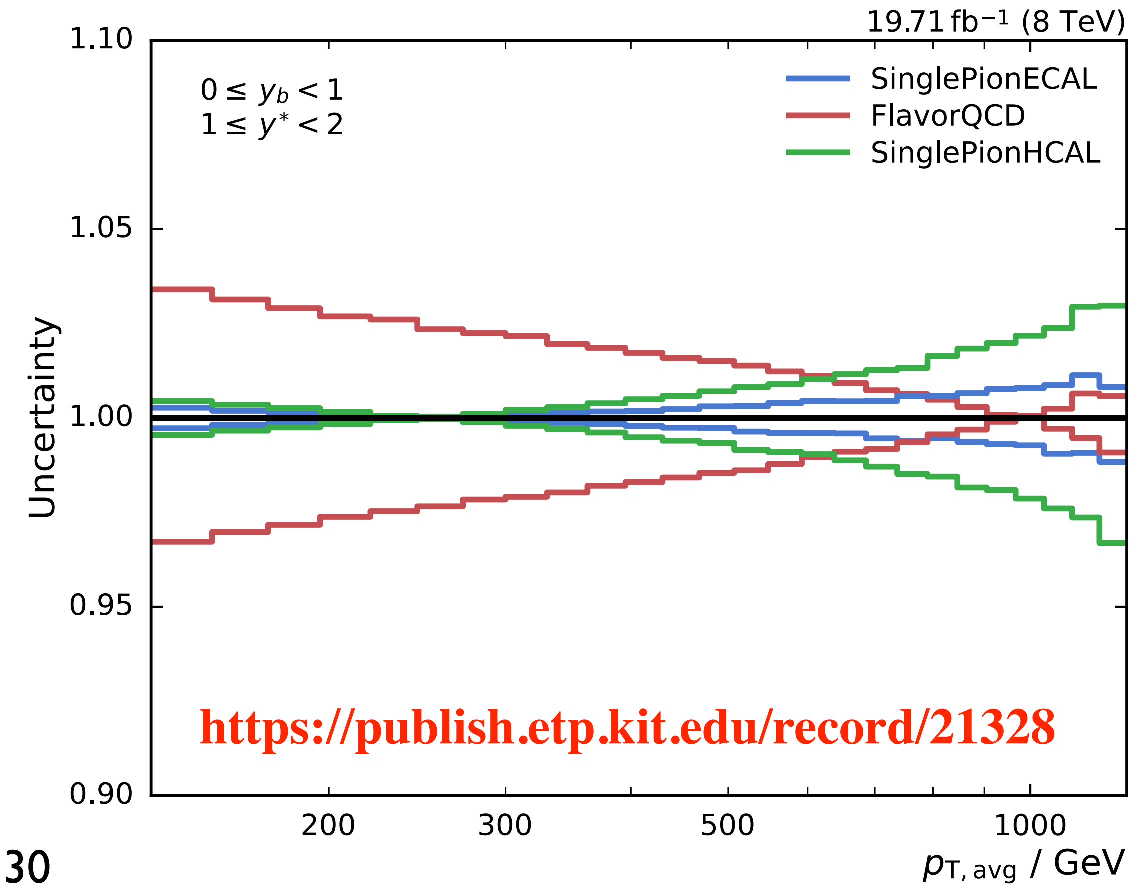




Technical aside (3) - CMS 8 dijets

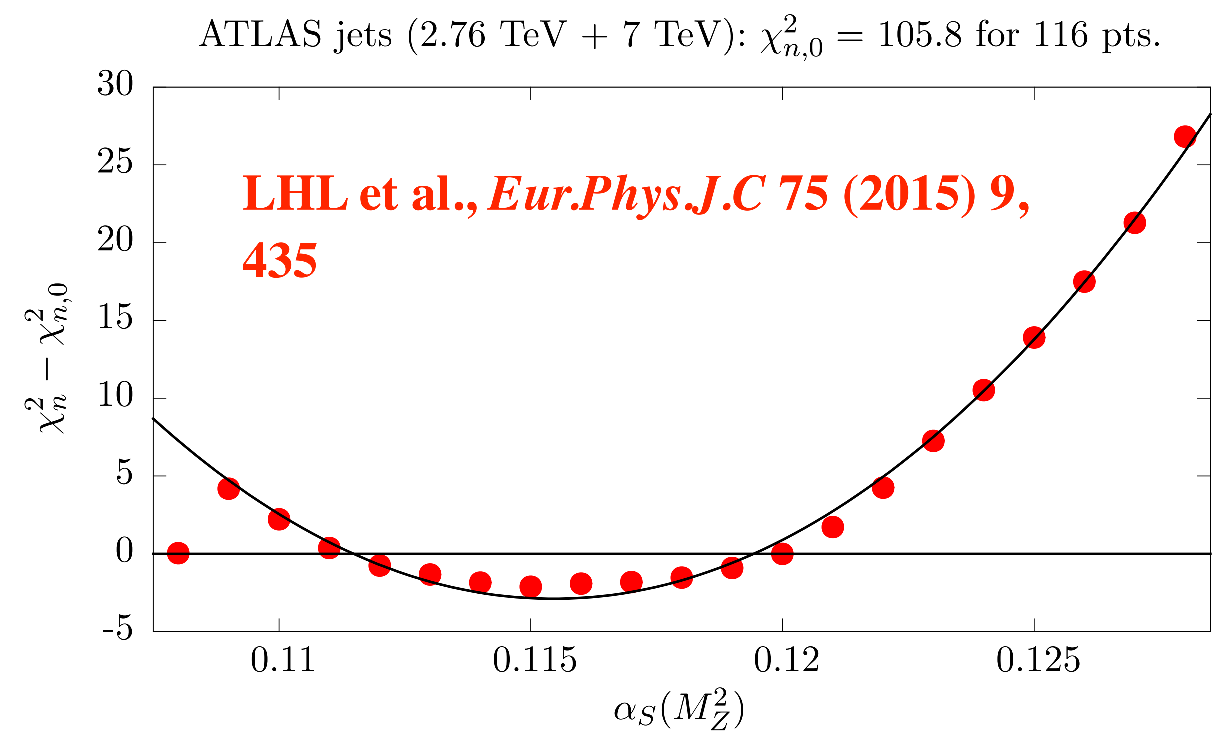
- Systematic uncertainties related to jet calibration correlated across kinematic (rapidity/$p_\perp$) space. Shape of these indicates anti-correlation between certain regions. However hepdata entries entirely positive.
- Through discussion with CMS colleagues have changed sign to more 'natural' (anti)-correlation.
- In the end this makes very little difference: improves χ^2 by ~ 1-2 points and gluon very stable. But more by chance than design.
- Detailed understanding/bookkeeping of systematic correlations key.





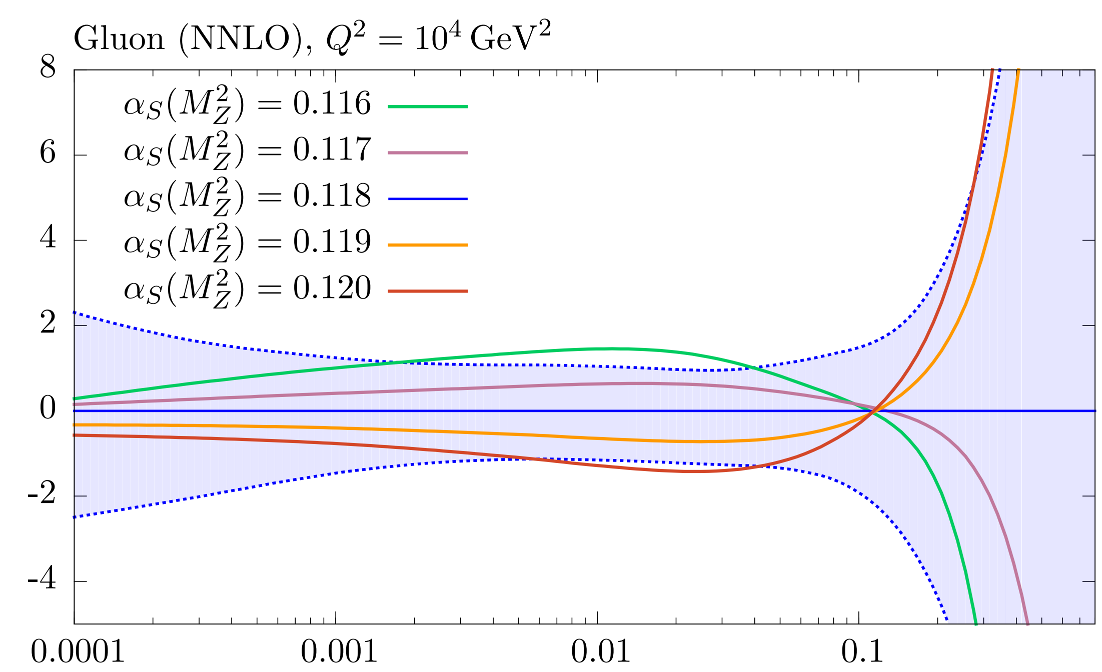
The strong coupling

- All results so far performed with fixed value of $\alpha_S(M_Z^2) = 0.118$
  - Jet data can clearly be sensitive to value of strong coupling.
  - However this is strongly correlated to extracted gluon.
  - Work in preparation.



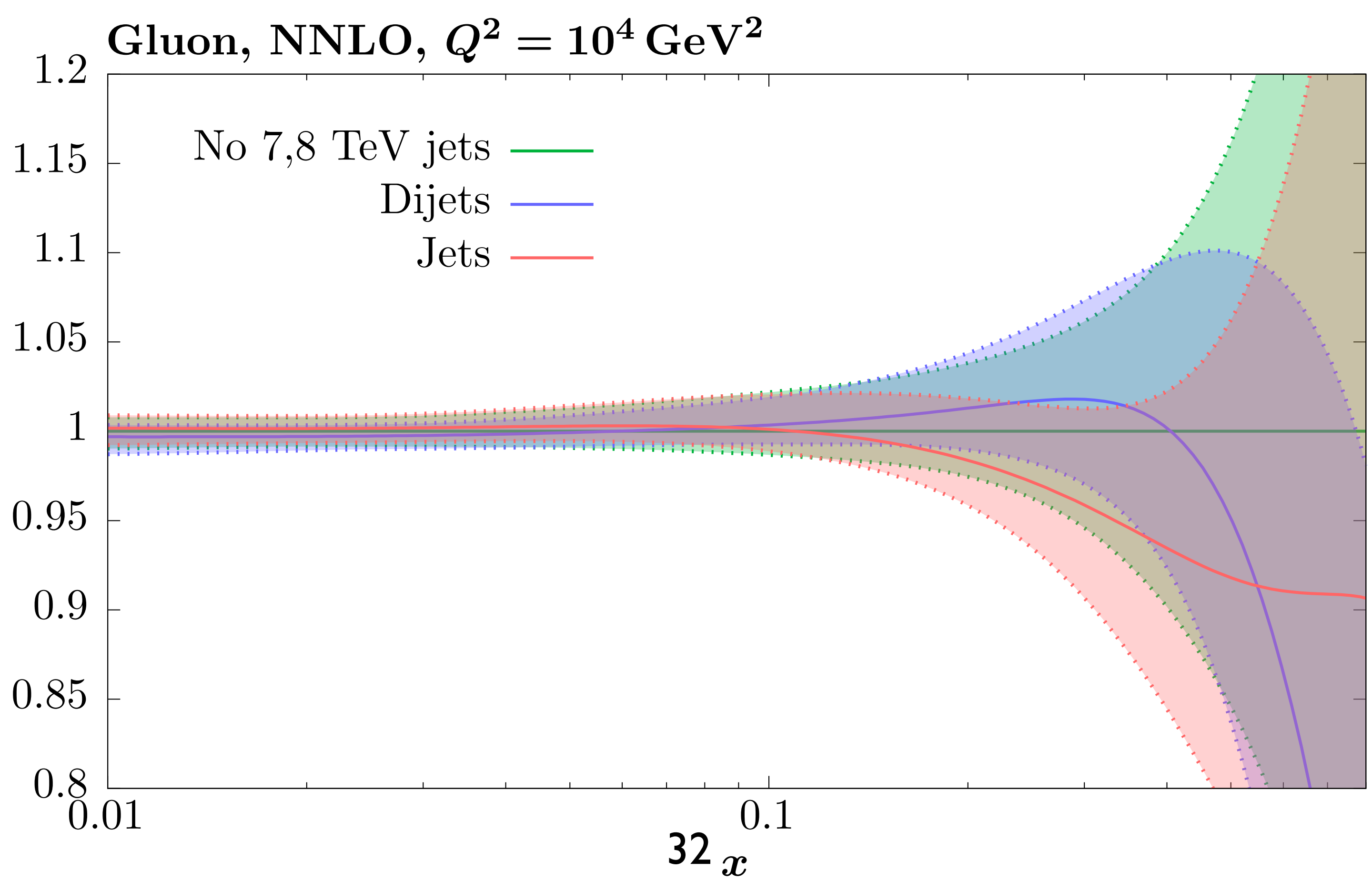
  - Care needed in interpretation of preferred value of α_S in PDF sensitive observables outside of global fit.

S. Forte and Z. Kassabov, *Eur.Phys.J.C* 80 (2020) 3, 182



Summary/Outlook

- We find:
  - ★ Impact of 7 + 8 TeV inclusive jets and dijets consistent (within uncertainties). However precise pulls not identical.
  - ★ Fit quality better for dijets and impact on gluon uncertainty larger in MSHT baseline fit.
  - ★ Great impact driven by 3D nature of CMS 8 TeV data, but also interplay with other data in fit. Global view needed.

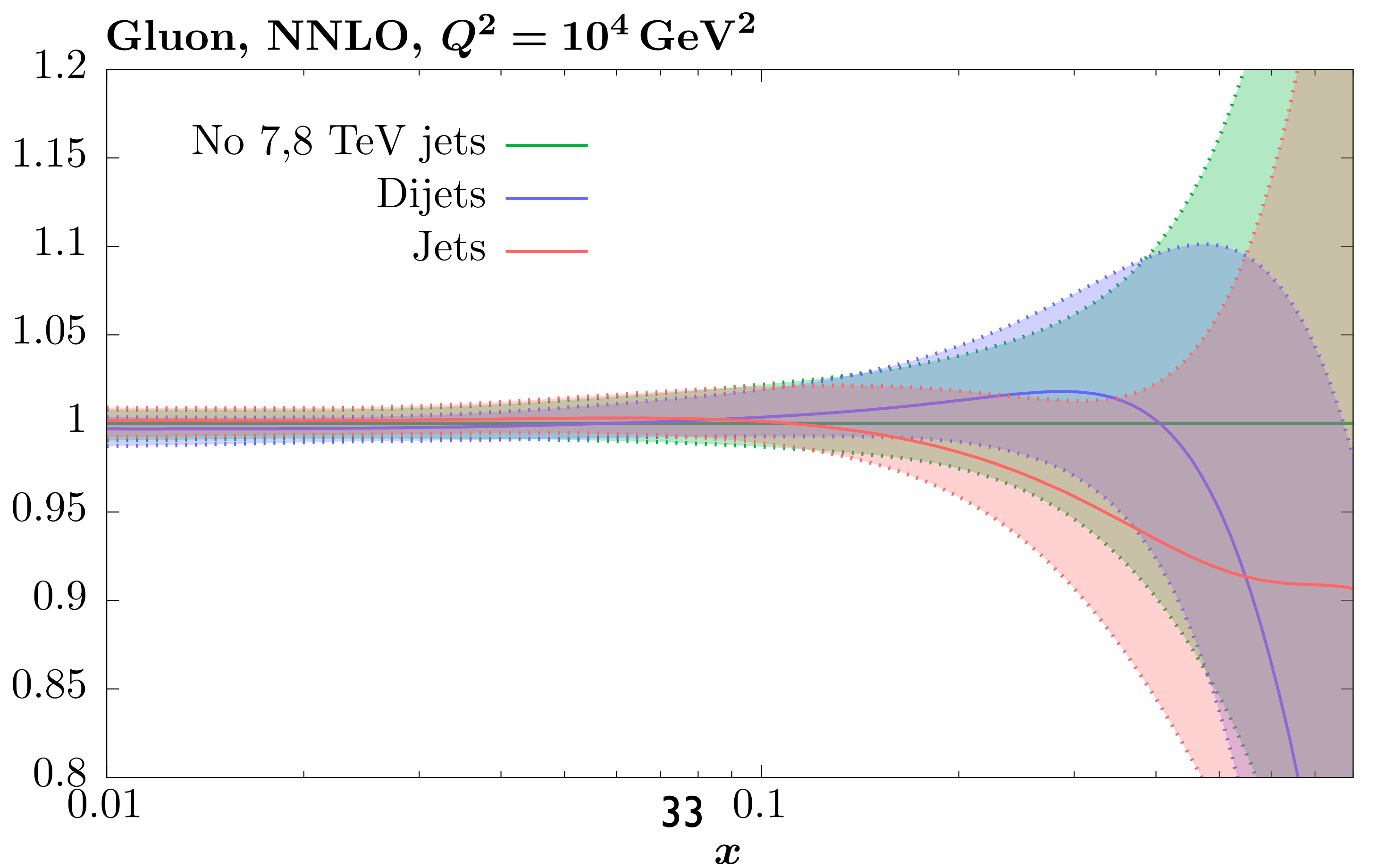


★ NNLO QCD corrections improve fit quality for inclusive jets and absolutely essential for dijets.

★ Some difference in pulls between ATLAS/CMS for inclusive jets. Not seen in dijet data.

★ However, for dijets this is essentially a trivial statement, as here completely driven by single CMS 8 TeV dataset.

★ Our results qualitatively similar, though not identical in all cases, with original NNPDF study. See Tommaso's talk (now).

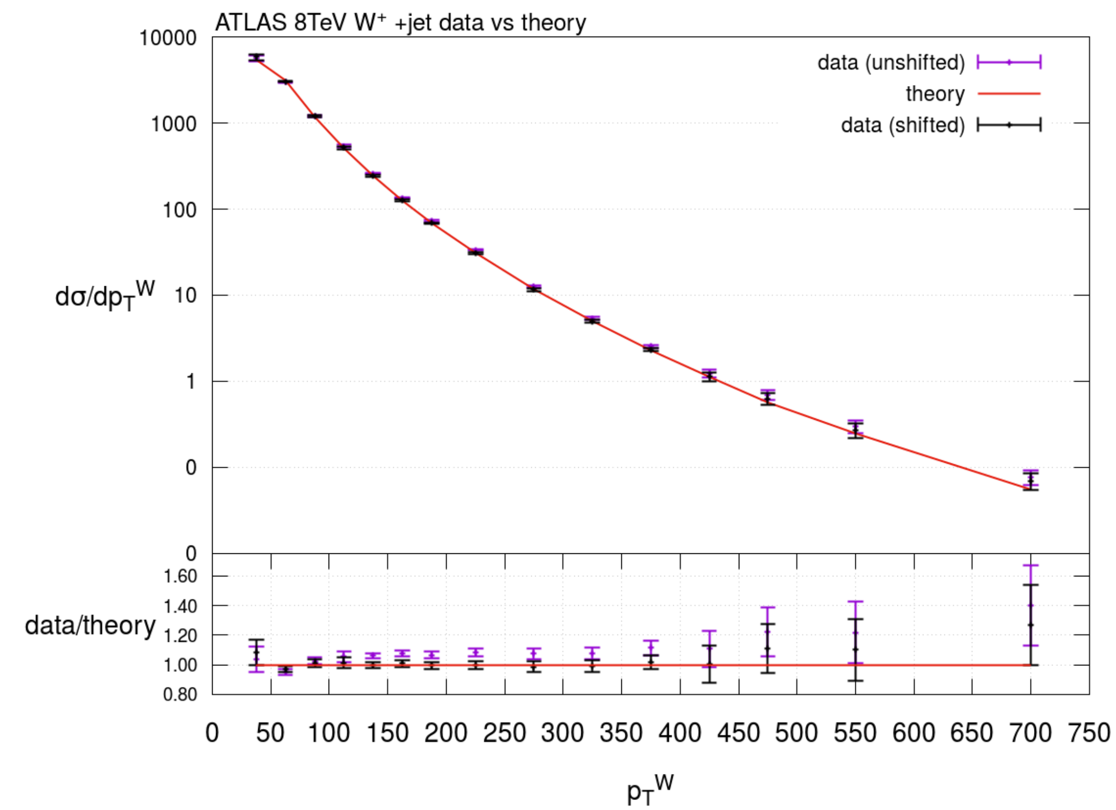


- Not discussed here:

  - ★ Impact of missing higher order corrections. Essential to include these in fit before making firm conclusion about jets vs. dijets and interplay with other data in fit. Work ongoing.

  - ★ Other jet related data. For example, ATLAS 8 TeV W + jets included in MSHT20 fit. Sensitive to high x gluon/quarks. Well fit, with moderate impact.

  - ★ Jets in DIS. No included so far in MSHT20, but to be looked at.



Thank you for listening!

(and over to Tommaso)