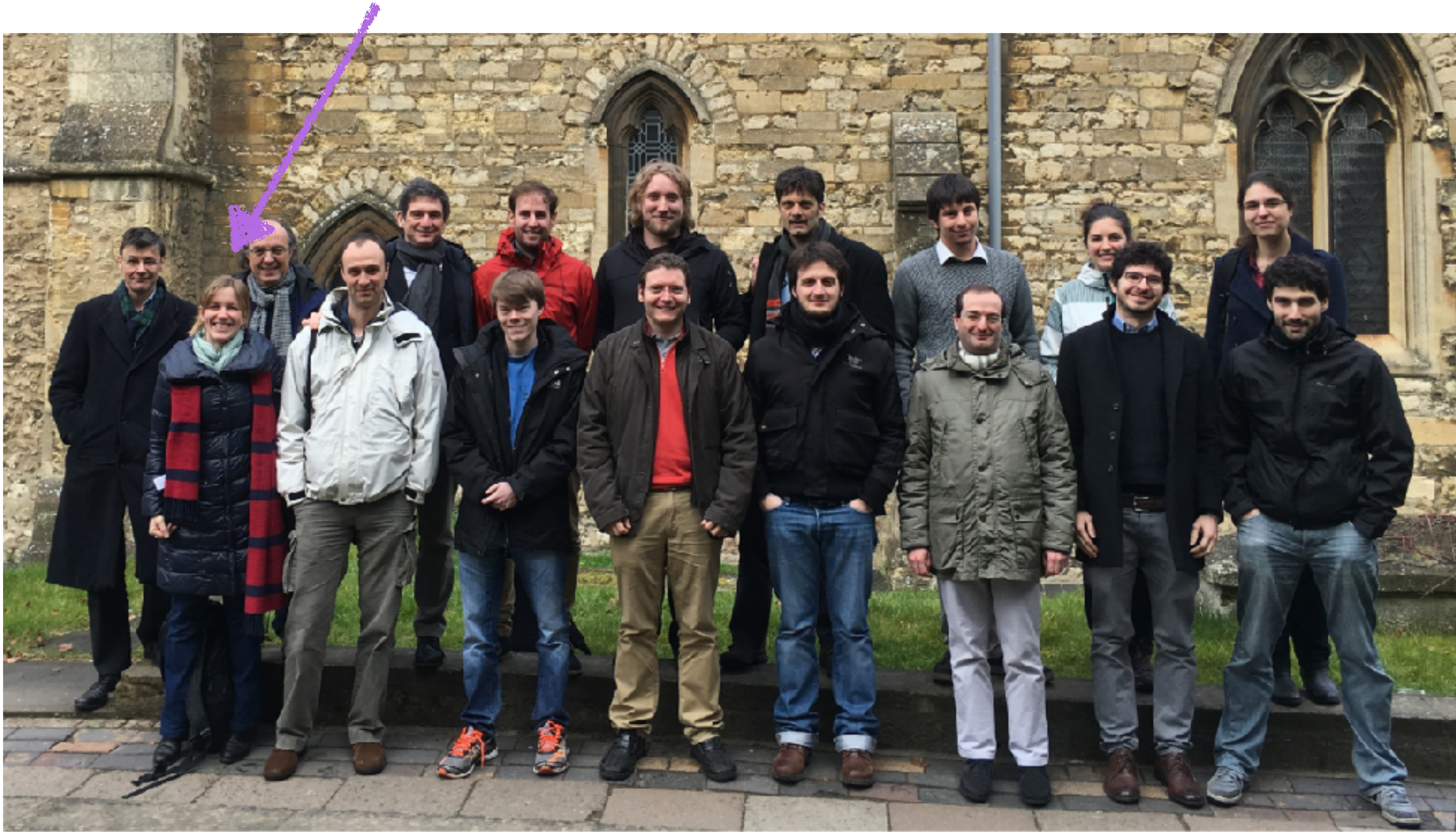
NEW FRONTIERS IN PDF DETERMINATION: NNPDF3.1

Maria Ubiali

Royal Society Dorothy Hodgkin Research Fellow
University of Cambridge



NNPDF

The NNPDF collaboration:

V. Bertone, N. Hartland, J. Rojo (Amsterdam)

J. I. Latorre (Barcelona) MU (Cambridge) S. Carrazza (CERN)

R. D. Ball, L. Del Debbio, P. G. Merrild (Edinburgh)

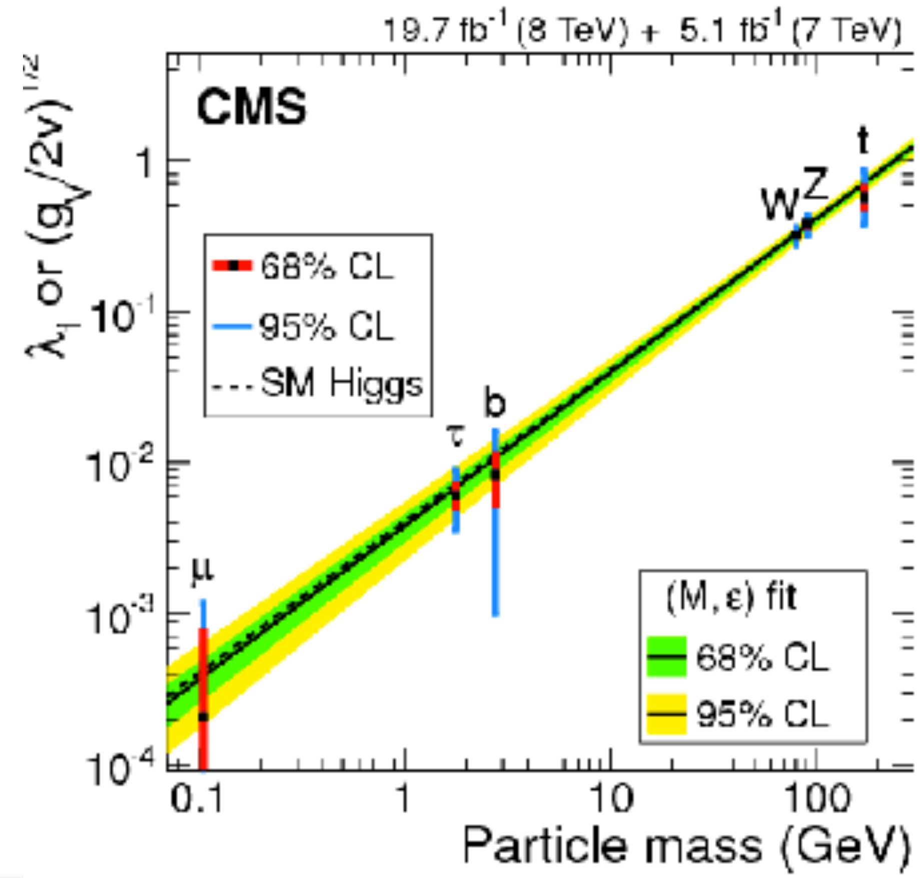
Z. Dim, S. Forte (Milano), A. Guffanti (Torino)

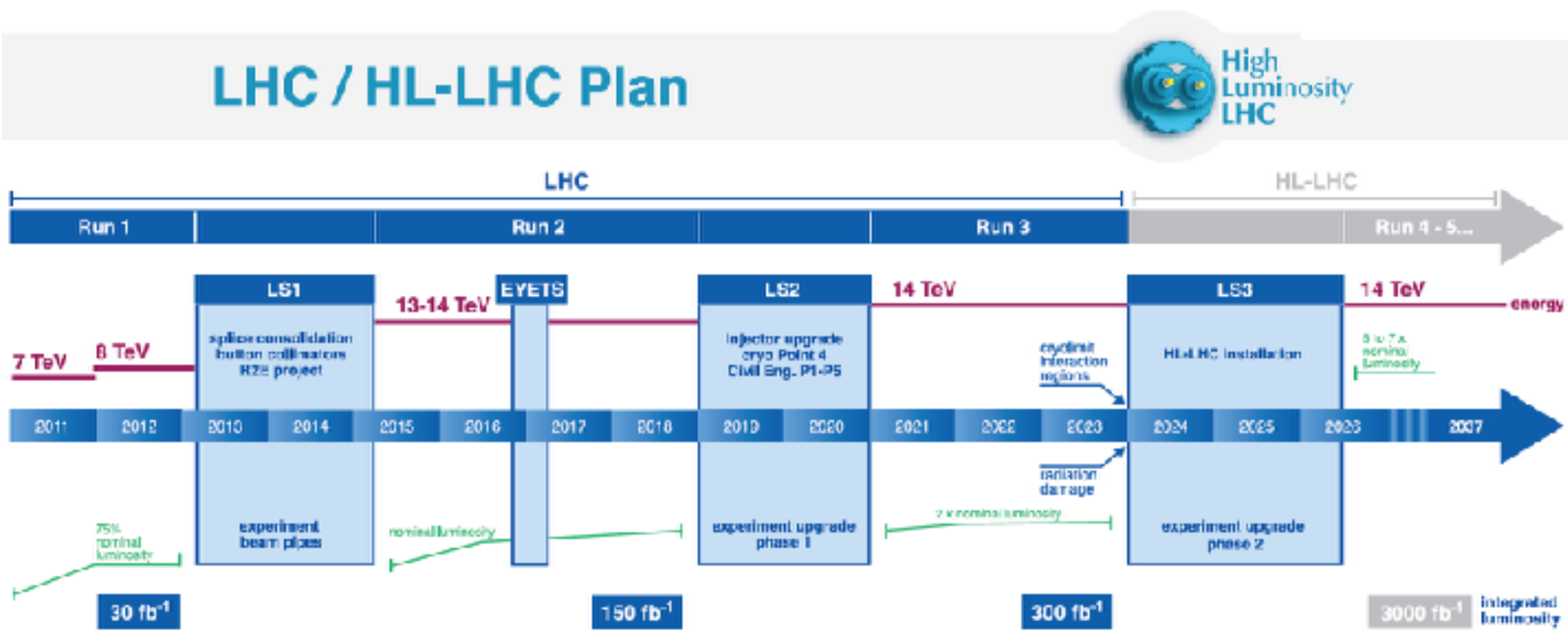
E. Nocera, L. Rottoli (Oxford)

LHC physics at Run II

- Is precision physics possible/necessary at hadron colliders?
  At the LHC a paradigm shift has taken place. Often theoretical predictions have to catch up with accuracy set by experiments

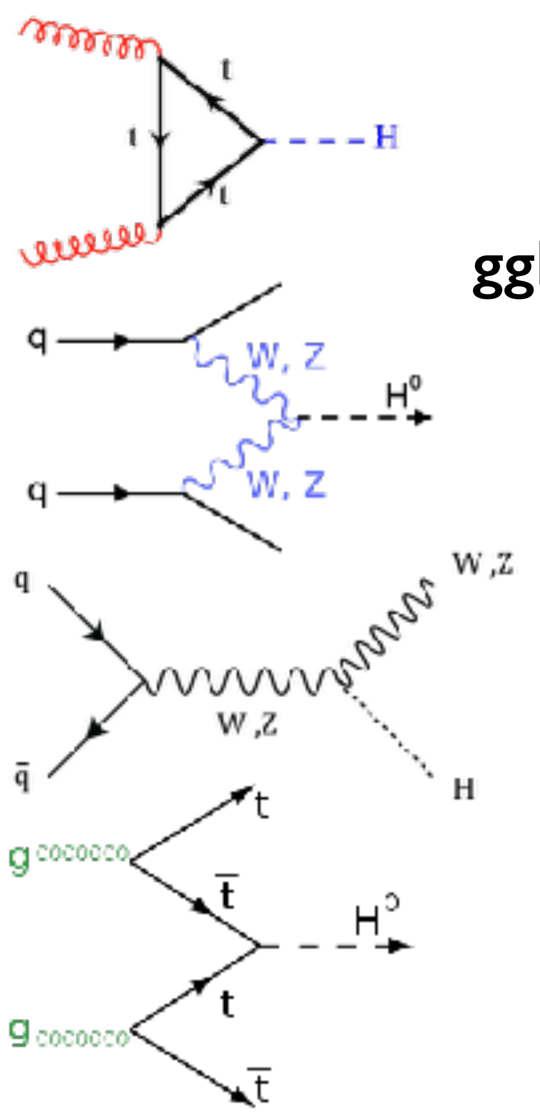
- Precise theoretical predictions are key to indirectly spot new physics signals and/or to characterise any possible “bump”

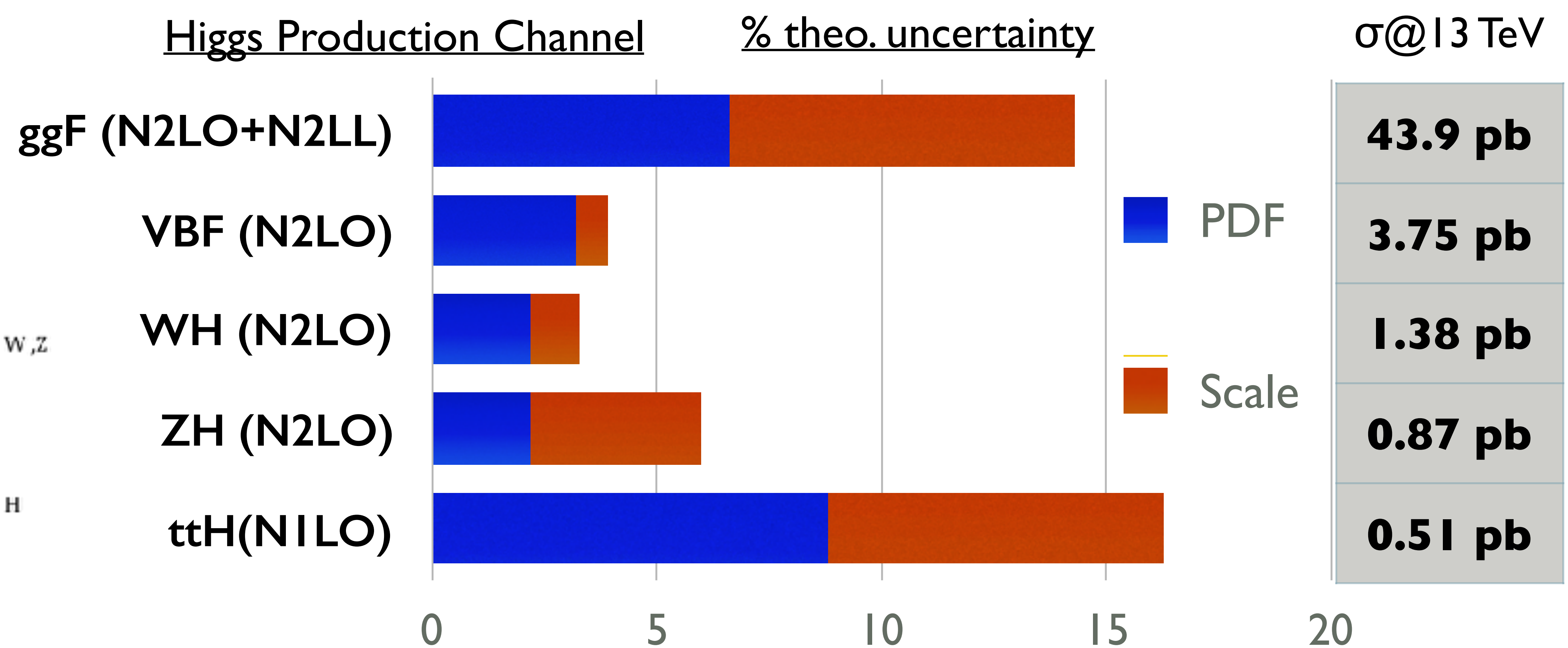




PDFs and LHC interplay

Yellow Report 3 (2013)

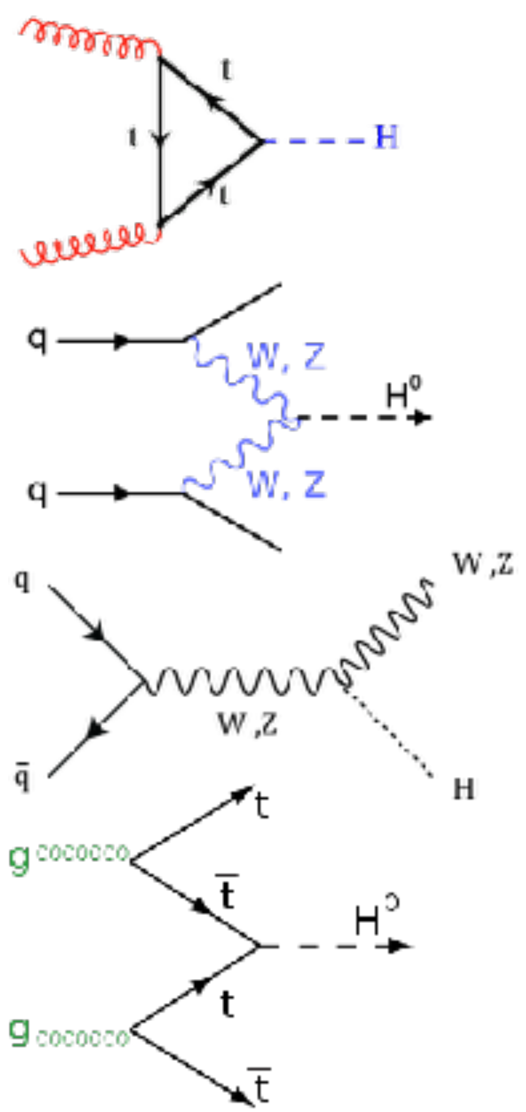


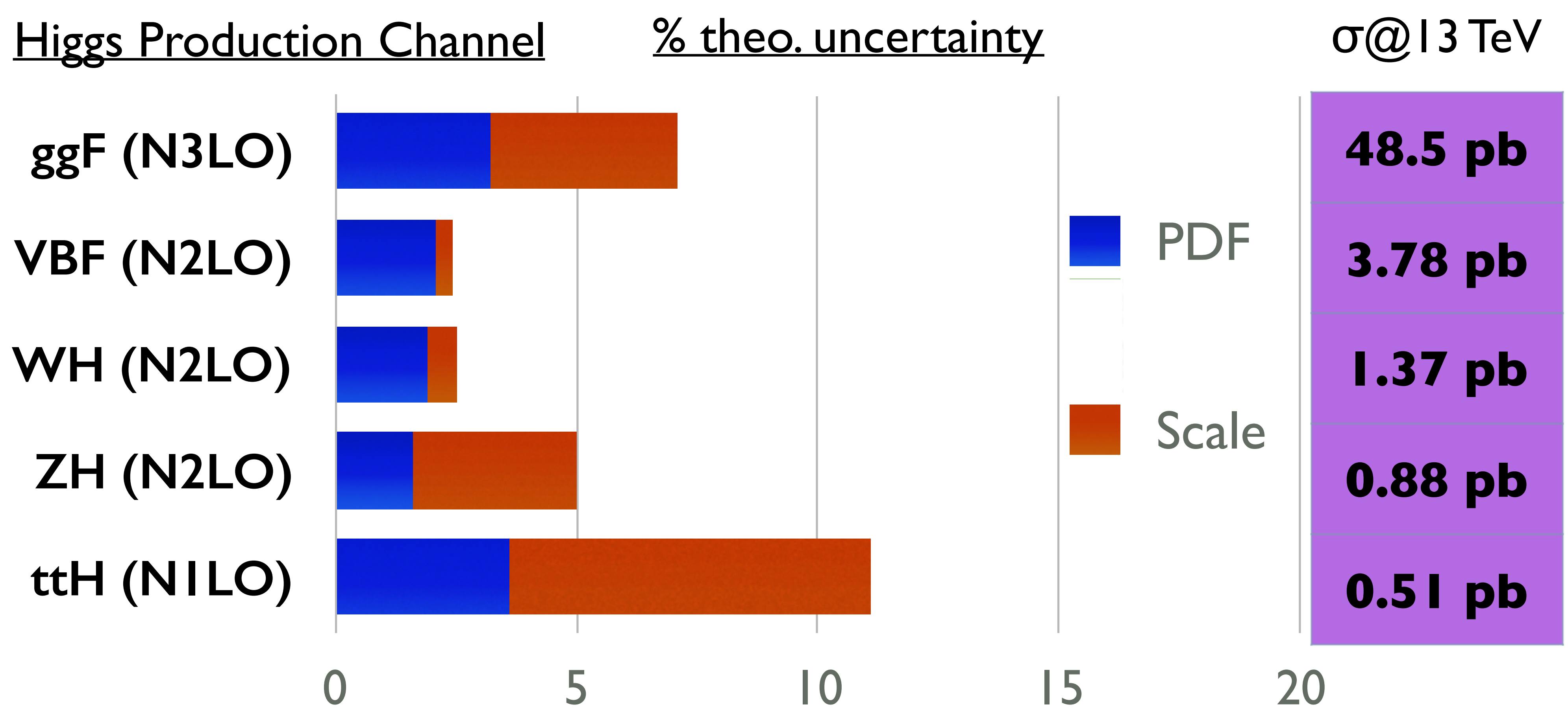


PDF uncertainties limiting factor in the accuracy of theoretical predictions

PDFs and LHC interplay

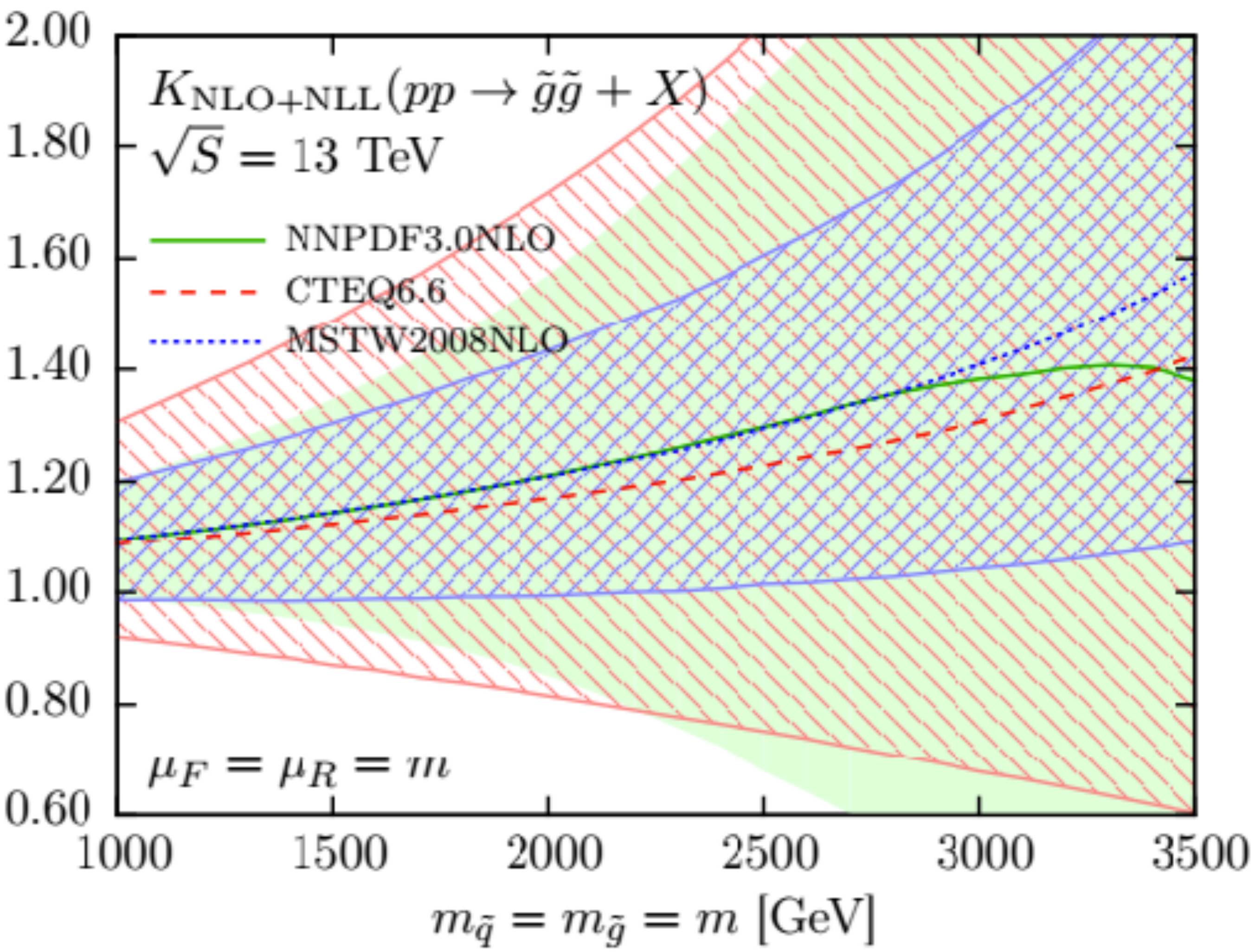
Yellow Report 4 (2016)



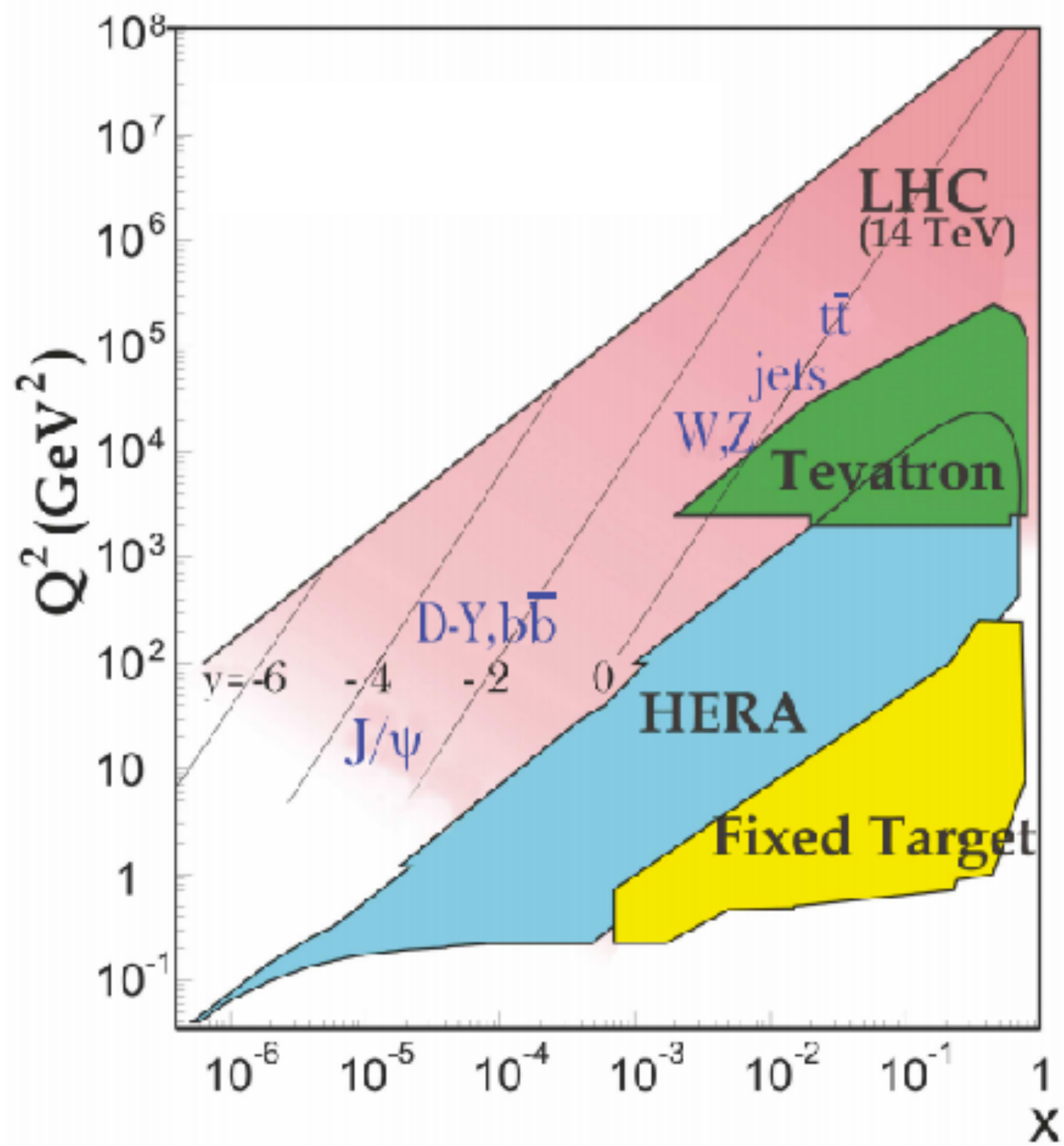


Reduced (still often dominant) PDF uncertainties

PDFs and LHC interplay



Beenakker et al.
EPJC76 (2016)2, 53



PDF uncertainties are a limiting factor in the accuracy of theoretical predictions, both within **SM** and **beyond**

Exploit the power of precise LHC data to reduce PDF uncertainties and "discriminate" among PDF sets

Outline of the talk

- Motivation
- Introduction
  - What PDFs are
  - The NNPDF approach
- The NNPDF3.1 set
  - Fitted versus perturbative charm
  - New constraints from LHC: challenges and opportunities (top, jets, Z pT)
  - Results and phenomenology
- Conclusions and outlook

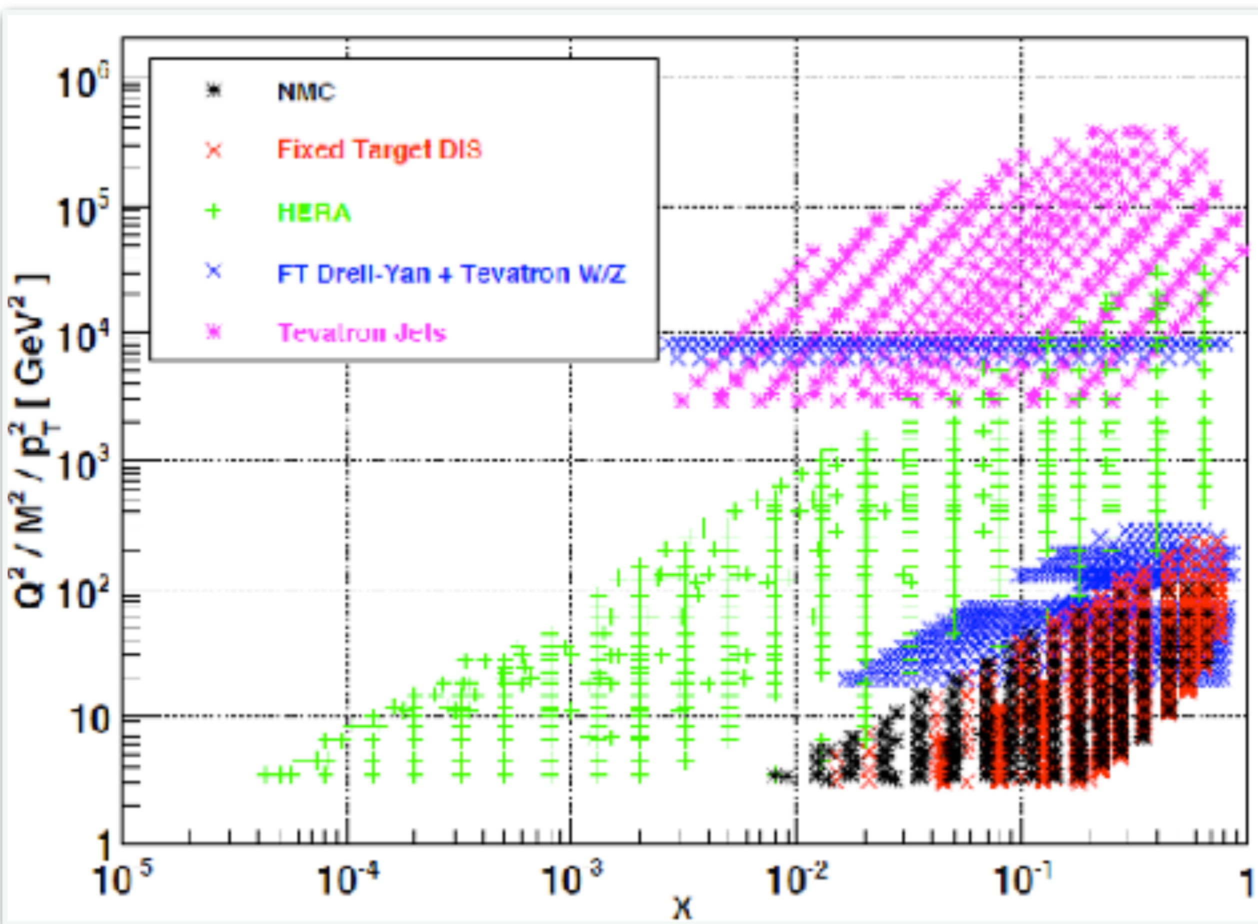
Collinear Factorisation Theorem

$$\frac{d\sigma_H^{pp\to ab}}{dX} = \sum_{i,j=1}^{N_f} f_i(x_1,\mu_F)\, f_j(x_2,\mu_F)\, \frac{d\sigma_H^{ij\to ab}}{dX}(x_1 x_2 S_{\rm had}, \alpha_s(\mu_R), \mu_F) + \mathcal{O}\left(\frac{\Lambda_{\rm QCD}^{2n}}{S_{\rm had}^n}\right)$$

$$\mu^2 \frac{\partial\, f(x,\mu^2)}{\partial\, \mu^2} = \int_z^1 \frac{dz}{z}\, \frac{\alpha_s}{2\pi}\, P(z)\, f\left(\frac{x}{z},\mu^2\right)$$

Dokshitzer, Gribov, Lipatov, Altarelli, Parisi renormalization group equations

Q-dependence: pert. QCD



x-dependence: from data

LO - Dokshitzer; Gribov, Lipatov; Altarelli, Parisi, 1977

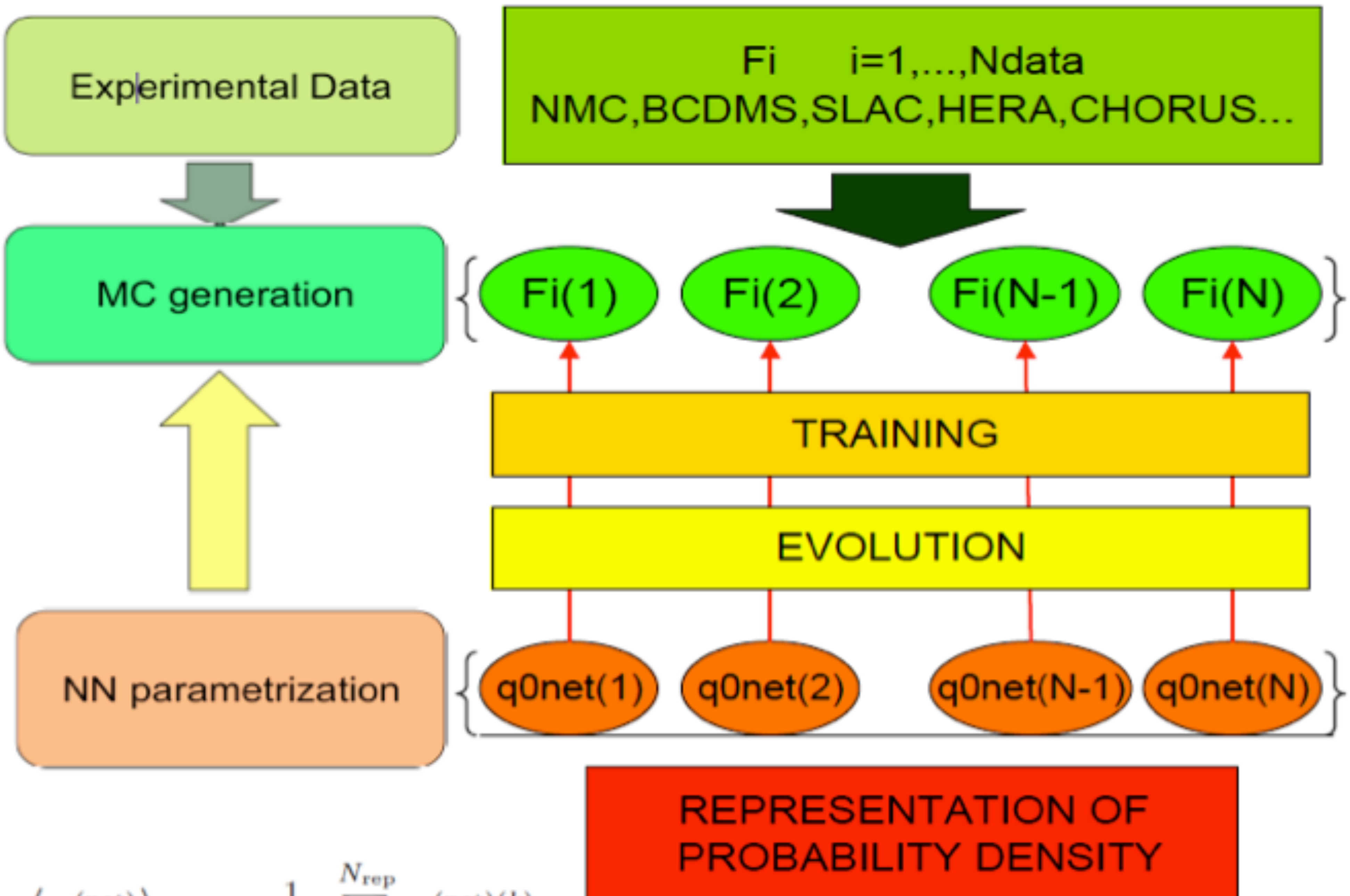
NLO - Floratos,Ross,Sachrajda; Floratos,Lacaze,Kounnas, Gonzalez-Arroyo,Lopez,Yndurain; Curci,Furmanski Petronzio, 1981

NNLO - Moch, Vermaseren, Vogt, 2004

The PDF extraction process

- Choose **experimental data** to fit and include all info on correlations → PDF uncertainty
- **Theory settings**: perturbative order, heavy quark mass scheme, EW corrections, intrinsic heavy quarks, α_s, quark masses value and scheme → *Hidden uncertainty*
- Choose a starting scale Q_0 where pQCD applies
- **Parametrize** independent quarks and gluon distributions at the starting scale → *Parametric versus non-parametric approach*
- Solve **DGLAP equations** from initial scale to scales of experimental data and build up observables
- **Fit** PDFs to data
- Provide **error sets** to compute **PDF uncertainties** → *Hessian versus MC approach*

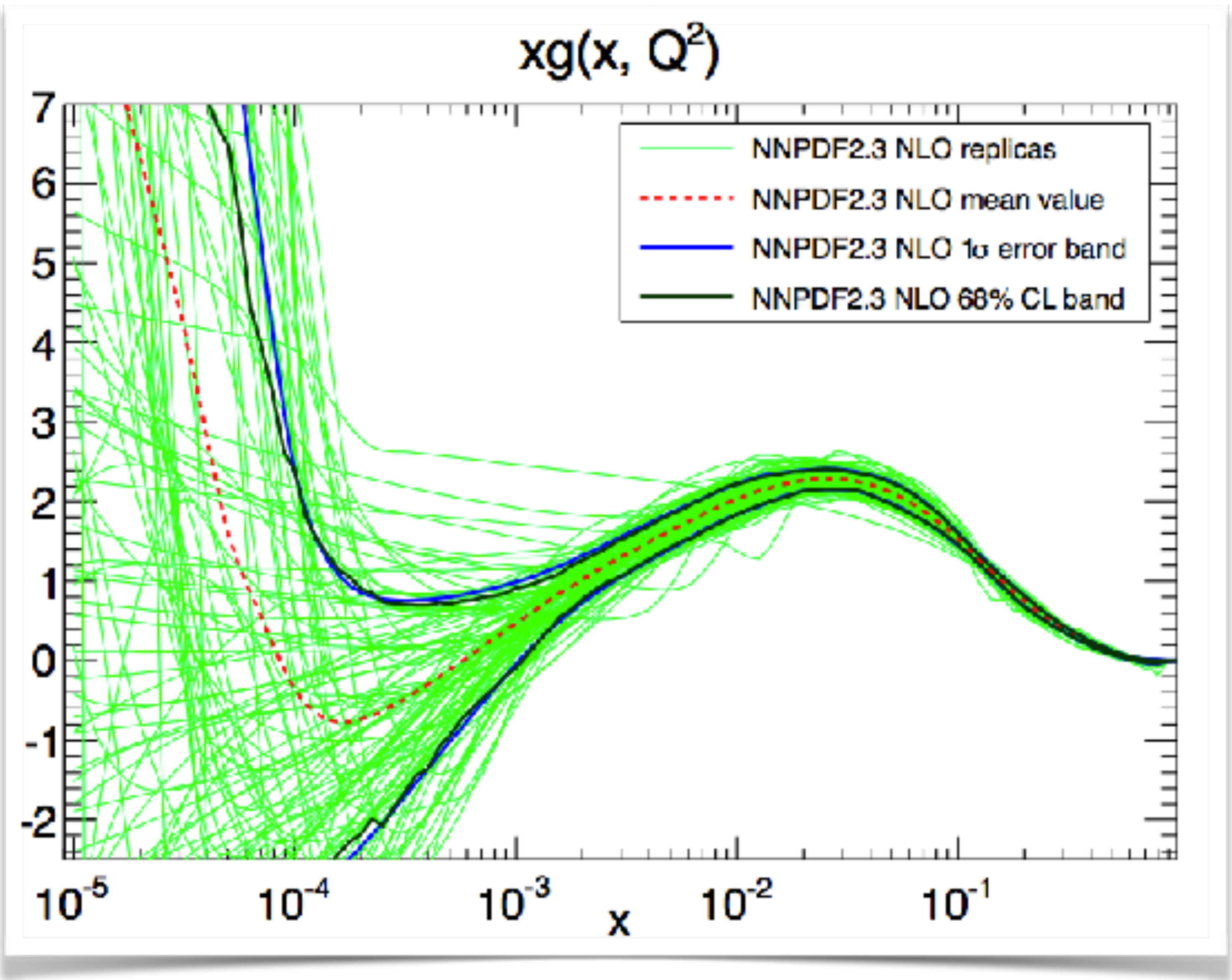
The NNPDF approach



$$\left\langle F_i^{(\mathrm{net})} \right\rangle_{\mathrm{rep}} = \frac{1}{N_{\mathrm{rep}}} \sum_{k=1}^{N_{\mathrm{rep}}} F_i^{(\mathrm{net})(k)}$$

Ball, Del Debbio, Forte, Guffanti, Latorre, Rojo, MU, ArXiv:0808.1231

The NNPDF approach



The N(eural)N(etwork)PDFs:

- Monte Carlo techniques: sampling the probability measure in PDF functional space
- Neural Networks: all independent PDFs are associated to an unbiased and flexible parametrization: O(300) parameters versus O(25) in polynomial parametrization

✓ Precise error estimate not driven by theoretical prejudice
✓ No need to add new parameters when new data are included
✓ Statistical interpretation of uncertainty bands
✓ Possibility to include data via re-weighting: no need to refit

NNPDF3.1

A fast-paced progress ...

2016 First PDF set with fitted charm

2015 First PDF set with threshold resummation

2014 NNPDF3.0: first PDF set validated with closure test

2013 NNPDF2.3QED: first PDF set with fitted photon PDF

NNPDF2.3: first PDF set with LHC data

2012 Reweighting PDFs

2011 NNPDF2.1:Heavy quark mass effects

Determination of α_S from PDF fit

2010 NNPDF2.0:First NNPDF global set

2009 NNPDF1.2: Determination of the proton strangeness: solved NuTeV anomaly

2008 NNPDF1.0: First NNPDF set - only DIS data

… to the future

April 2017 NNPDF3.1

Summer 2017 NNPDF3.1QED (à la LUXqed)

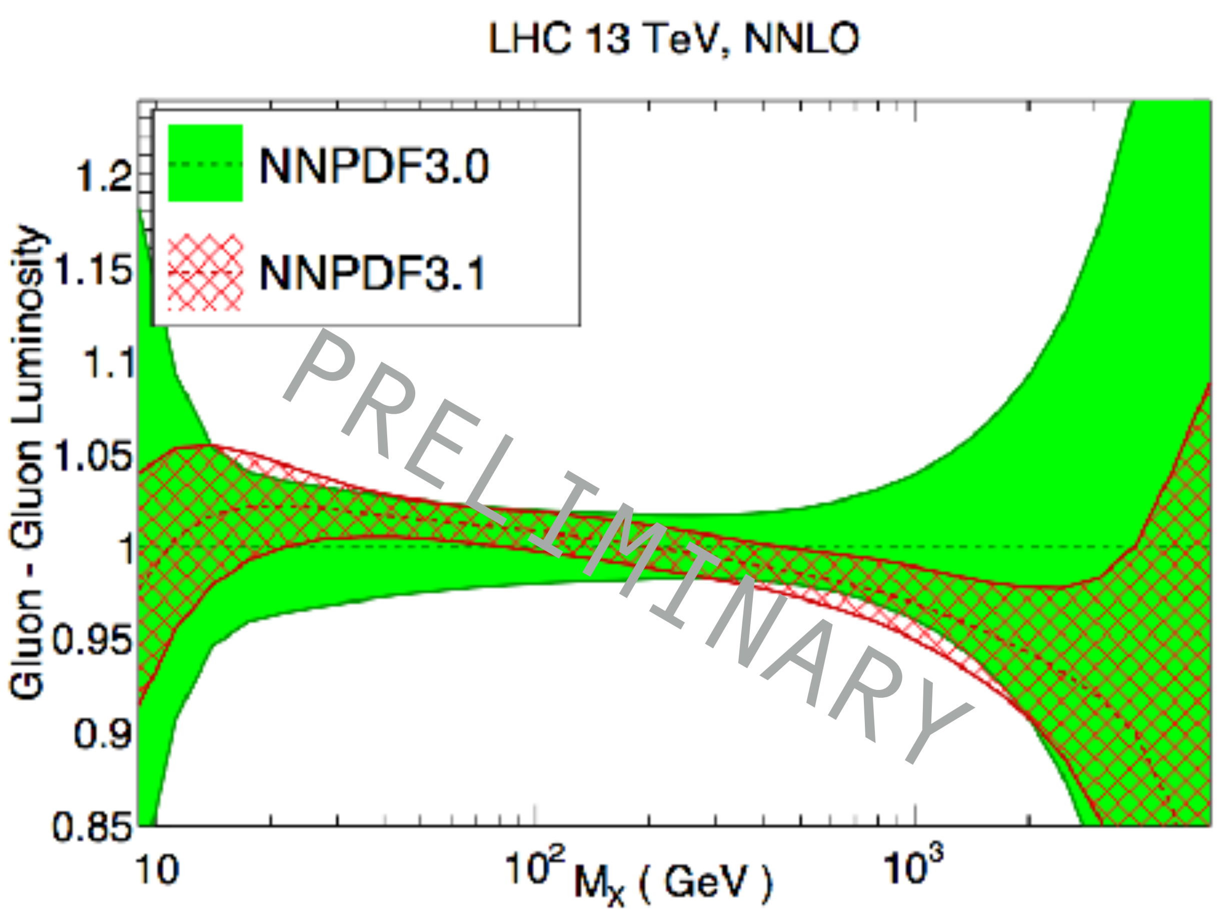
Updated determination of α_S and m_c

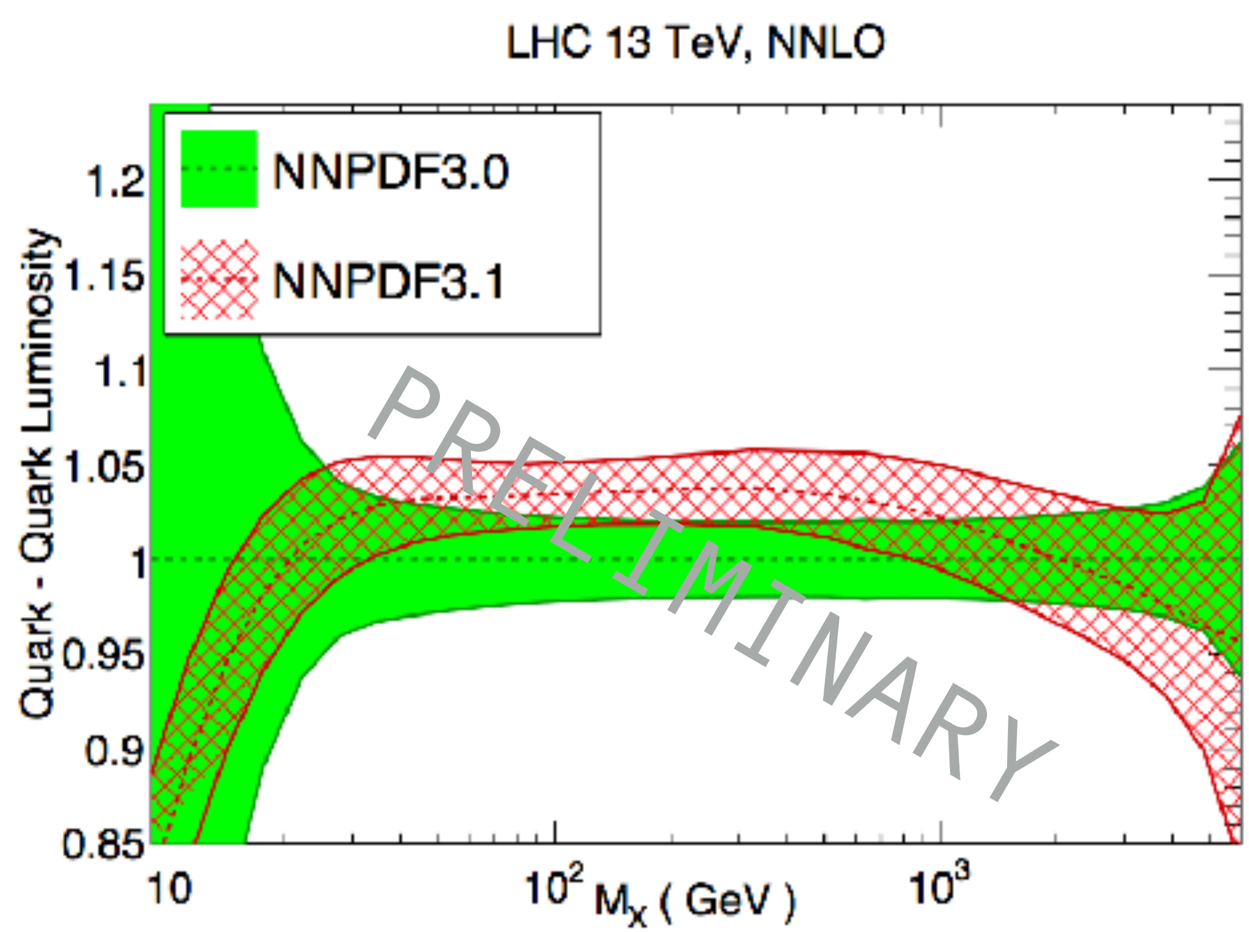
2018 PDF fits with (scale)theory uncertainties

PDF set with small-x resummation

The NNPDF3.1 analysis

- Plethora of new precise measurements and new available precise theoretical calculations call for an updated analysis: top differential distributions, transverse momentum distribution of the Z, combined HERA I-II data, legacy data from Tevatron, etc…
- Main methodological improvement is fitted charm PDFs, which increases stability with respect to choice of charm threshold





NNPDF3.1: fitted charm

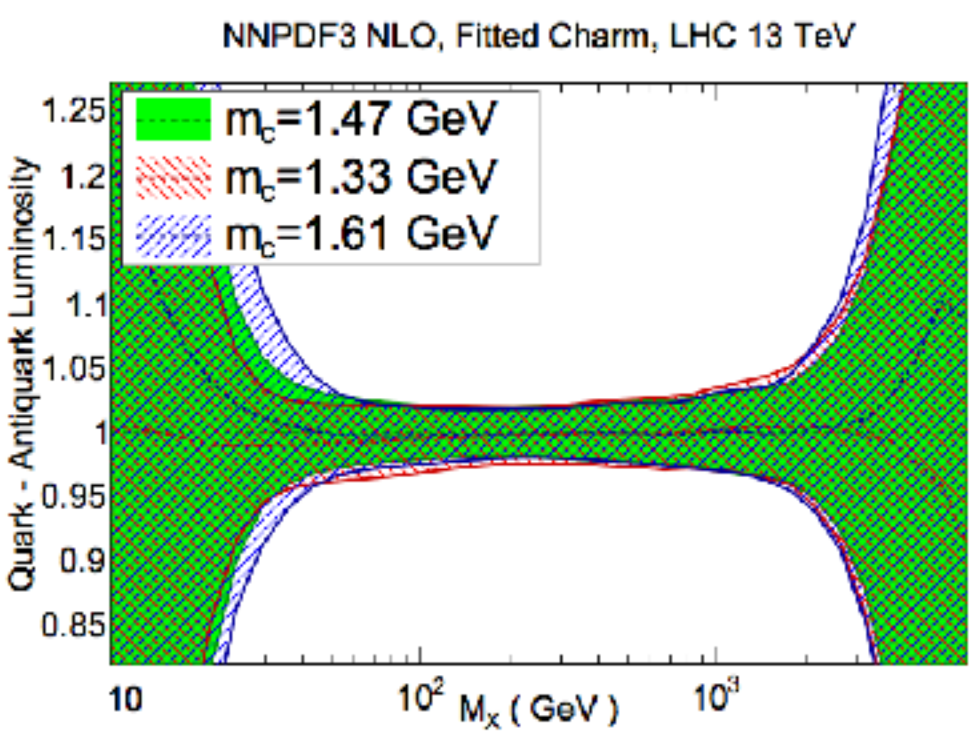
- Most global fits assume scale-independent charm content of the proton vanishes

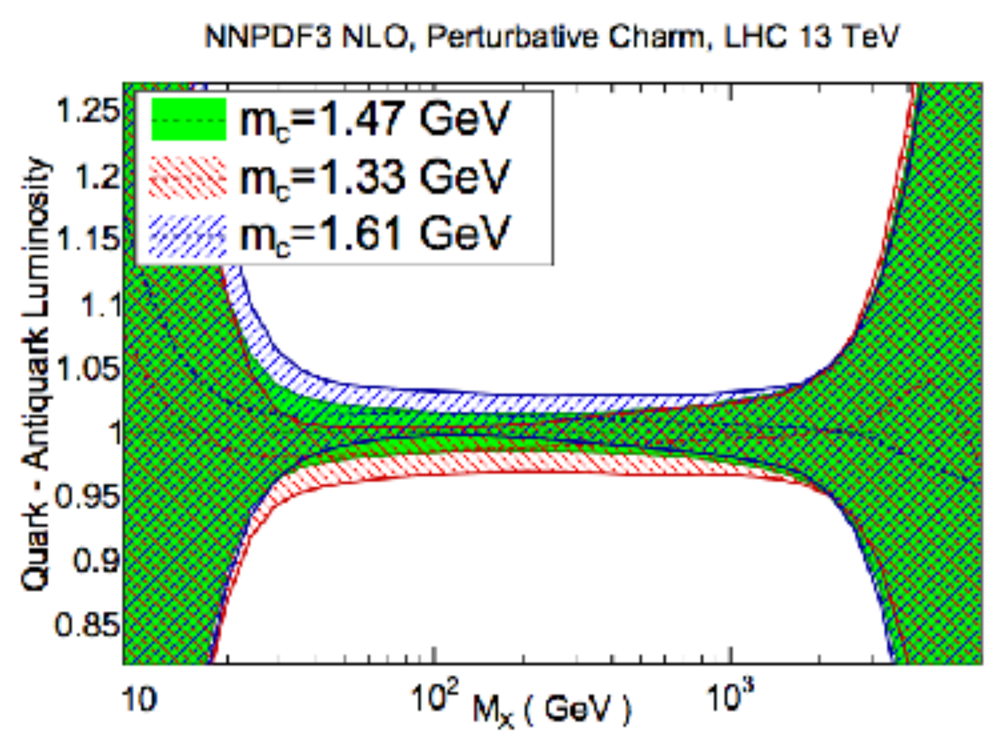
$$\alpha_s^{(4)}(m_h^2) = \alpha_s^{(3)}(m_h^2) + \mathcal{O}(\alpha_s^3)\,,$$
$$f_i^{(4)}(m_h^2) = \sum_j K_{ij}(m_h^2) \otimes f_j^{(3)}(m_h^2)$$

→ modified by non-zero intrinsic charm

- Why fit the intrinsic component of the charm?
  - ☑ Stabilise the dependence on mc
  - ☑ Compare determination with available models

Eur.Phys.J. C76 (2016) no.11, 647



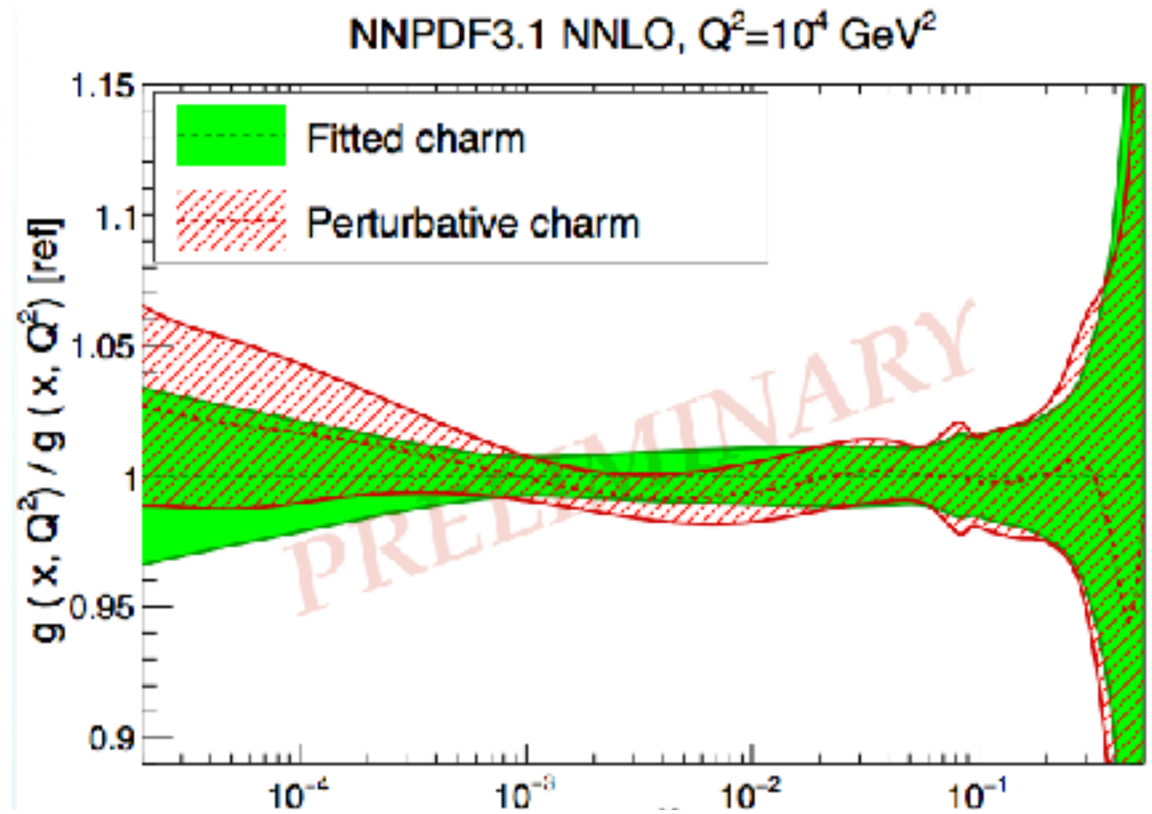


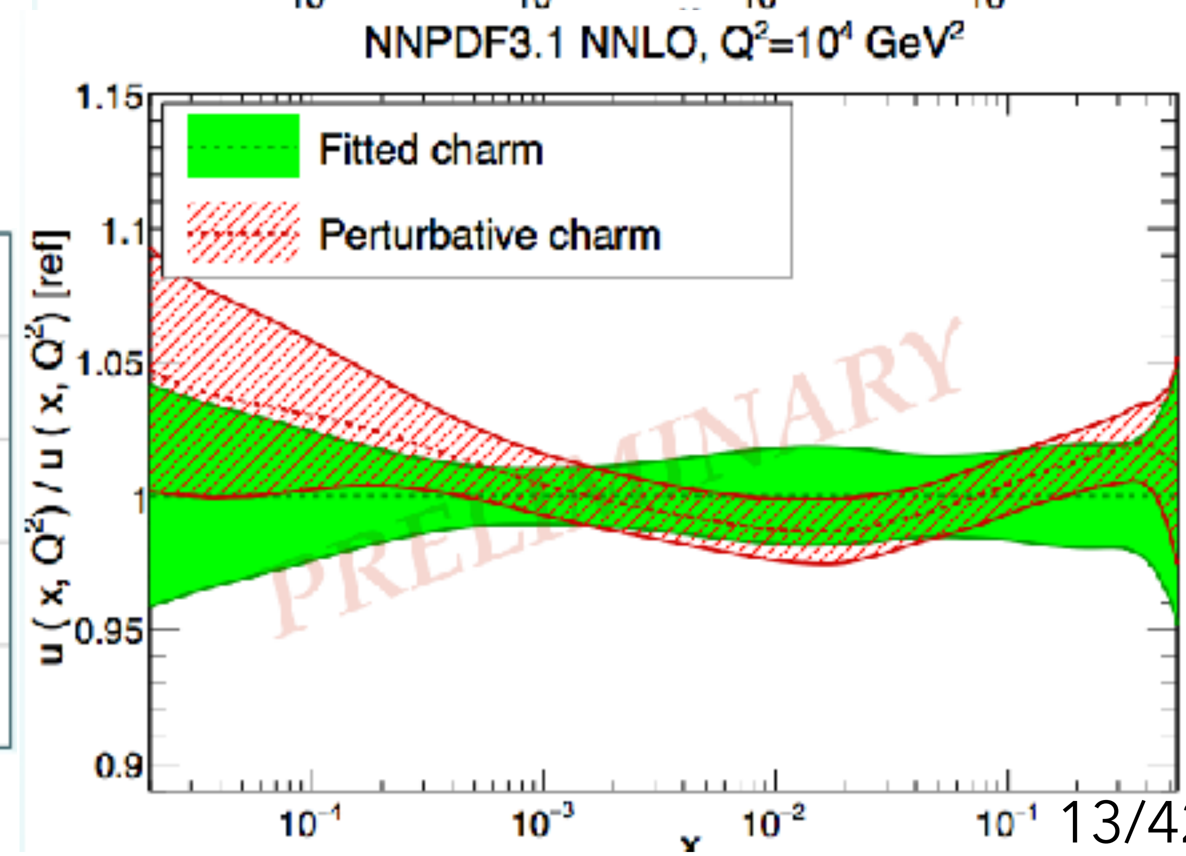
NNPDF3.1: fitted charm

- Differences between NNPDF3.1 NNLO fits with perturbative or fitted charm moderate but non-negligible for precision physics
- Fitted charm slightly better data description
- Both fits will be released

	NNLO FittedCharm	NNLO PertCharm	NLO FittedCharm	NLO PertCharm
HERA	1.16	1.20	1.16	1.16
ATLAS	1.13	1.19	1.45	1.50
CMS	1.04	1.06	1.20	1.20
LHCb	1.46	1.46	1.94	1.93

PRELIMINARY





NNPDF3.1: new data

NNPDF3.0 + **NNPDF3.1**

Combined HERA inclusive data	q and g at small/med x
ATLAS jets 2.76 TeV and 7 TeV **+ 2011 data 7 TeV**	gluon large x
ATLAS high-mass DY at 7 TeV **+ low mass**	**q/q~ separation**
ATLAS W pT data at 7 TeV	g and q at moderate x
ATLAS & CMS differential Z pT data at 7 & 8 TeV	**g and q at moderate x**
CMS (Y,M) double diff distributions 7 TeV **+ 8 TeV**	flavour separation
CMS jets at 7 TeV **+ 2.76 and 8 TeV jet data**	**gluon large x**
CMS muon charge asymmetry at 7 TeV **+ 8 TeV**	quark separation
CMS W+c at 7 TeV	strangeness
LHCb Z rapidity distribution at 7 TeV **+ 8 TeV (full data)**	small/large x quarks
ATLAS+CMS tt total xsec at 7/8 TeV	gluon large x
ATLAS+CMS tt differential xsec at 7/8 TeV	gluon large x
D0 legacy W asymmetry data	q/q~ separation

NNPDF3.1: data implementation

- PDF evolution and DIS structure functions up to NNLO computed with **APFEL** in **FONLL** scheme

- Hadronic data computed using **APPLgrid/fastNLO** interfaced to MCFM/aMC@NLO/NLOjet++ & bin-by-bin NNLO/NLO C factors for each process

- **APFELgrid** used to combined PDF evolution and interpolated coefficient functions

$$\sigma = \sum_{i,j}^{n_f} \sum_{\alpha,\beta}^{n_x} W_{ij\alpha\beta}\, f_i(x_\alpha, Q_0^2) f_j(x_\beta, Q_0^2)$$

APFEL

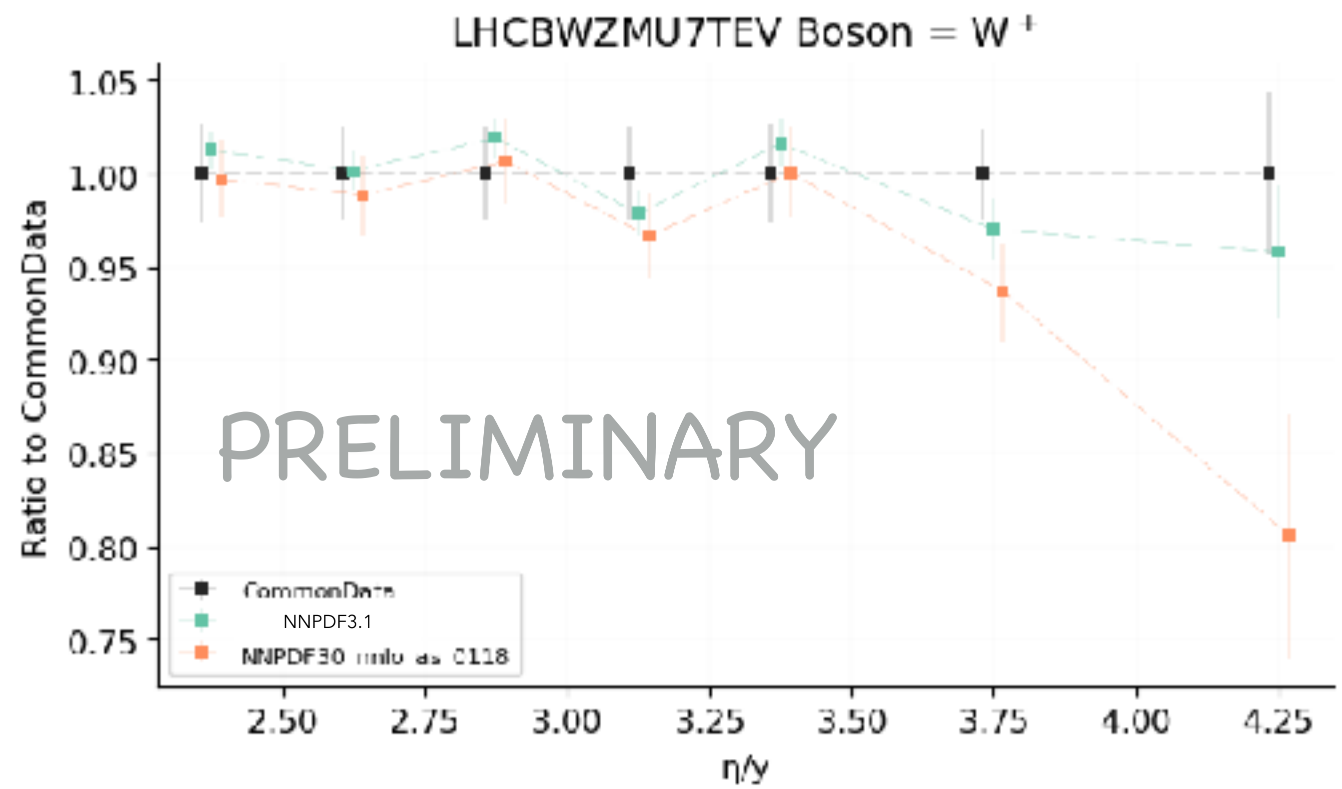
the APPLgrid project

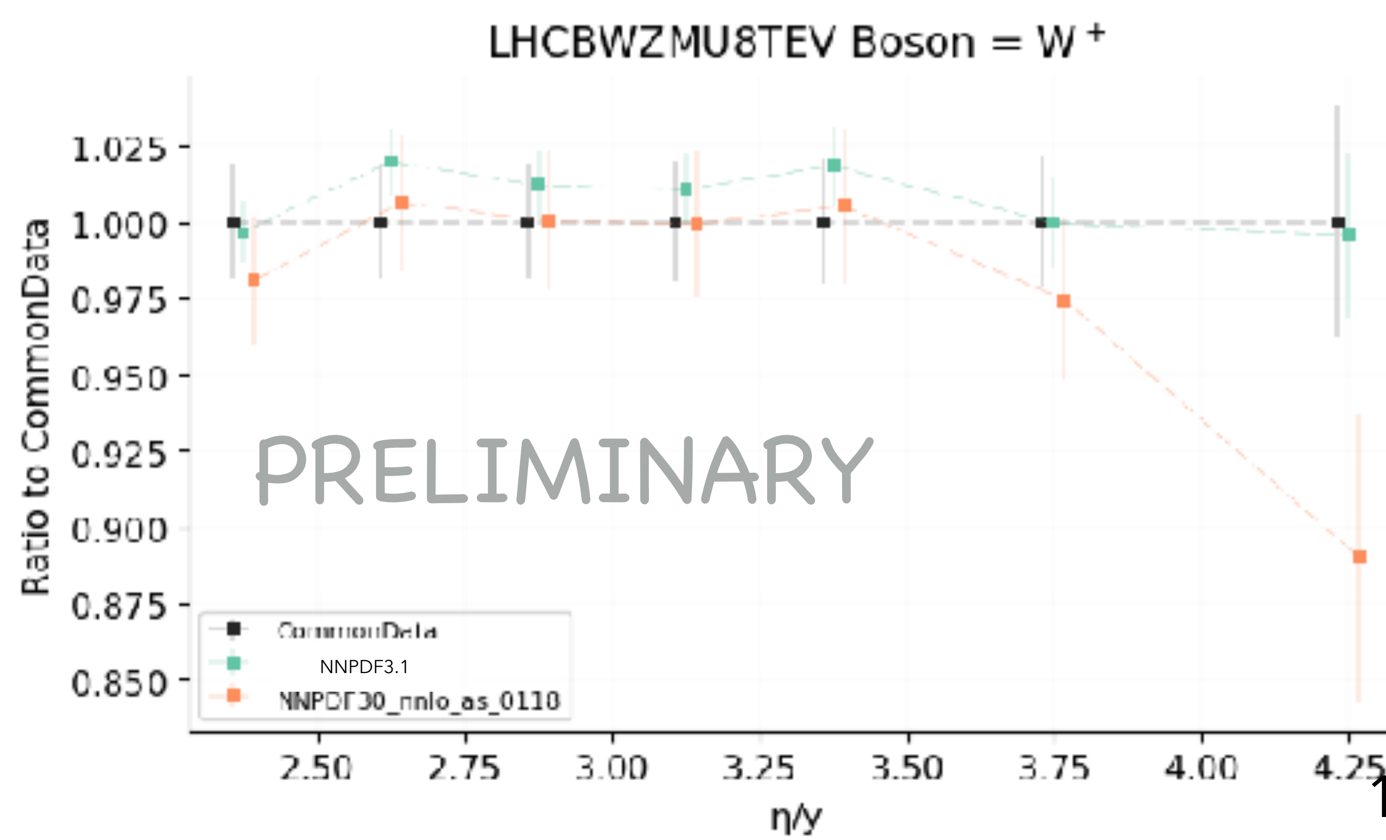
Observable	APPLGRID	**APFELcomb**
W^+ production	1.03 ms	0.41 ms (2.5x)
Inclusive jet production	2.45 ms	20.1 μs (120x)

APPLgrid, Carli et al EPJC66 (2010) 503-524 & FASTNLO, Kluge et al
APFELgrid, Bertone et al 1605.02070
aMCfast, Berton et al JHEP 1408 (2014) 166
MCgrid, Del Debbio et al Comput.Phys.Commun. 185 (2014) 2115-2126

NNPDF3.1: LHCb 7 and 8 TeV data

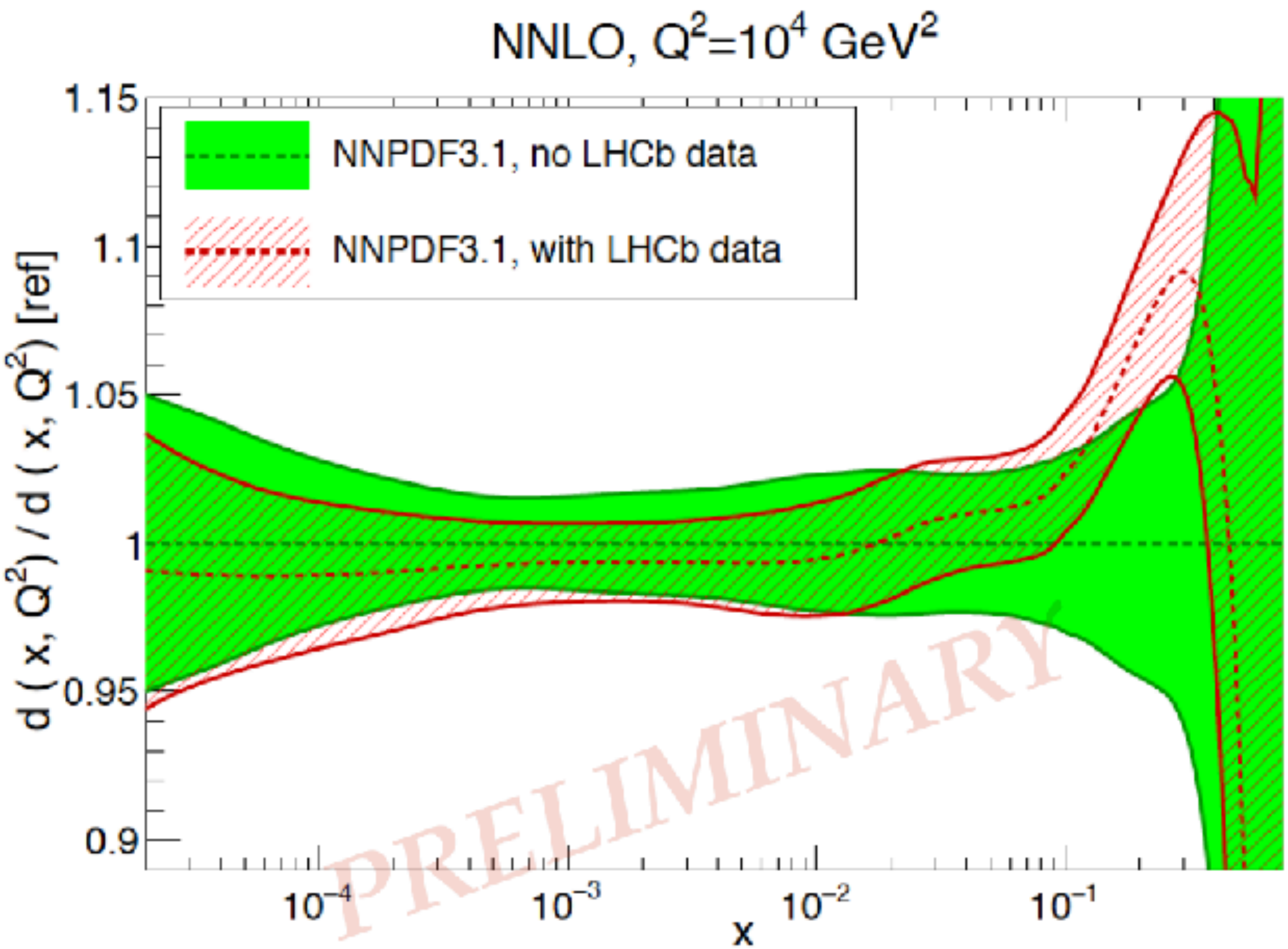
- LHCb published complete 7 TeV and 8 TeV Z and W measurements in electron and muon channels in the forward region
- Forward W/Z production data improve flavour-separation especially at large-x
- Good theoretical description and sizeable impact





NNPDF3.1: LHCb 7 and 8 TeV data

- LHCb published complete 7 TeV and 8 TeV Z and W measurements in electron and muon channels in the forward region
- Forward W/Z production data improve flavour-separation especially at large-x
- Good theoretical description and sizeable impact



NNPDF3.1: new NNLO calculations

- NNLO calculations are essential to reduce theoretical uncertainties in PDF analyses
- Stunning progress has been made on some key processes for PDF determination
- Not all of them yet fully exploited (jets and direct photon production)

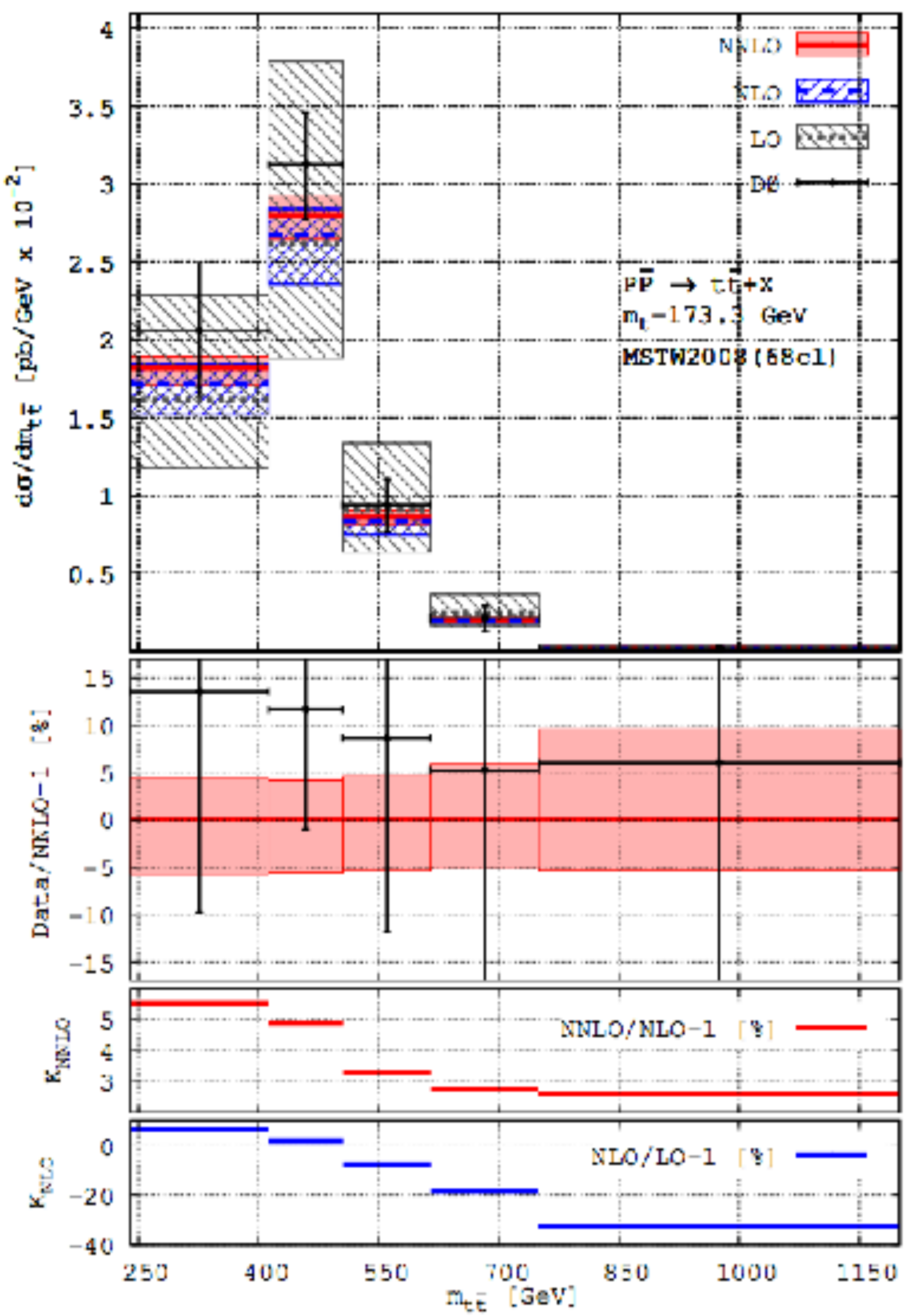
✓ NNLO top pair production (total and differential)
  Czakon, Fiedler, Mitov [PRL 116(2016) 082003]
  Czakon, Mitov [JHEP 1301(2015)]

✓ W/Z+j and W/Z transverse momentum distributions
  Gehrmann-De Ridder et al [1605.04295]
  Boughezal, Liu, Petriello [1602.08140]
  Boughezal, Liu, Petriello [1602.06965]
  Boughezal et al [PRL 116(2016) 152001 & 062002]
  Gehrmann-De Ridder et al [1507.02850]

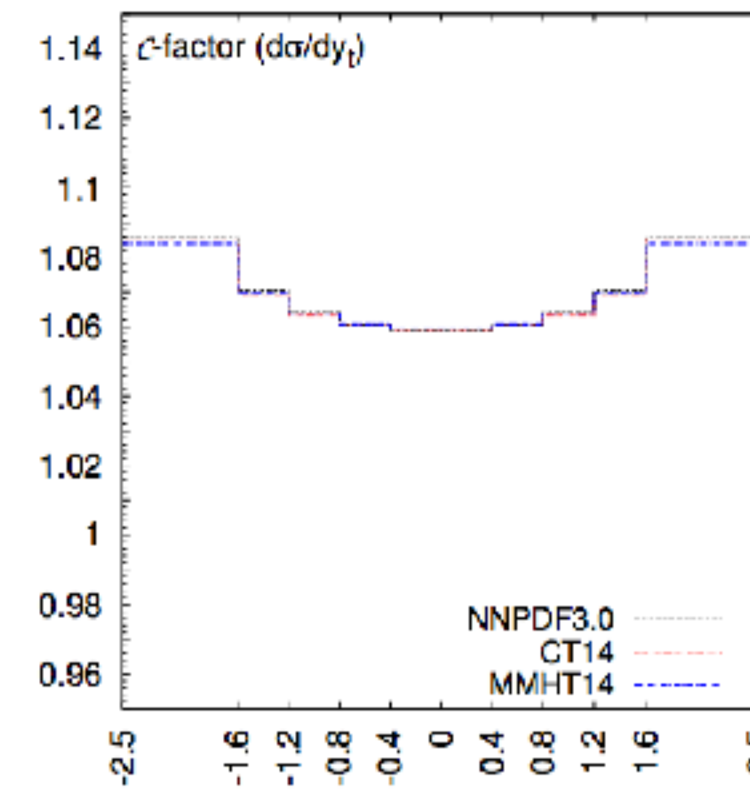
✓ Inclusive jet cross section
  Currie et al [JHEP 1401 (2014) 110]
  Gehrmann-De Ridder et al [PRL 110 (2016) 162003]

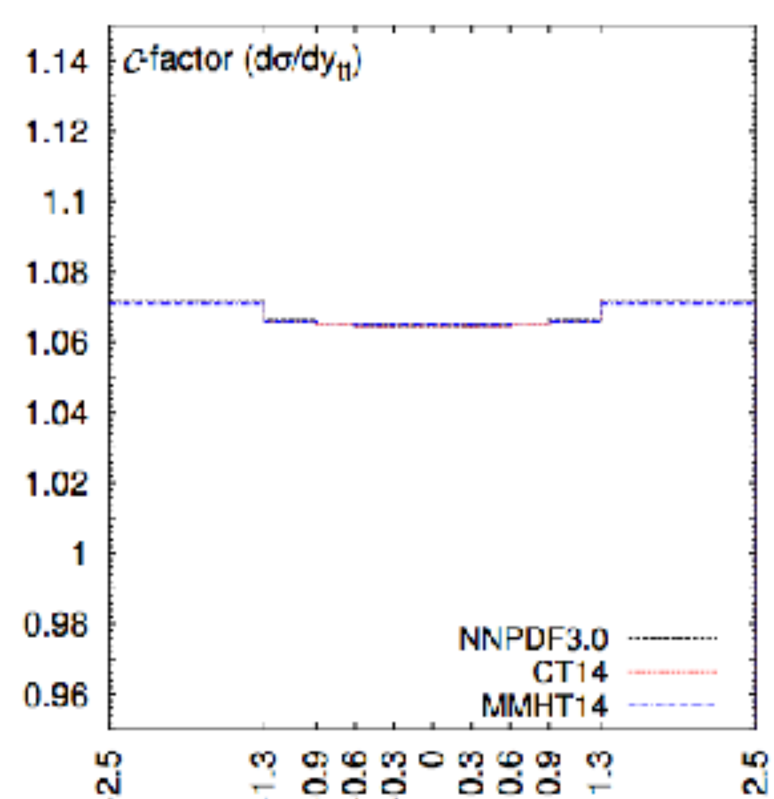
✓ Direct photon production
  Campbell, Ellis, Williams [1612.04333]

NNPDF3.1: top differential distributions



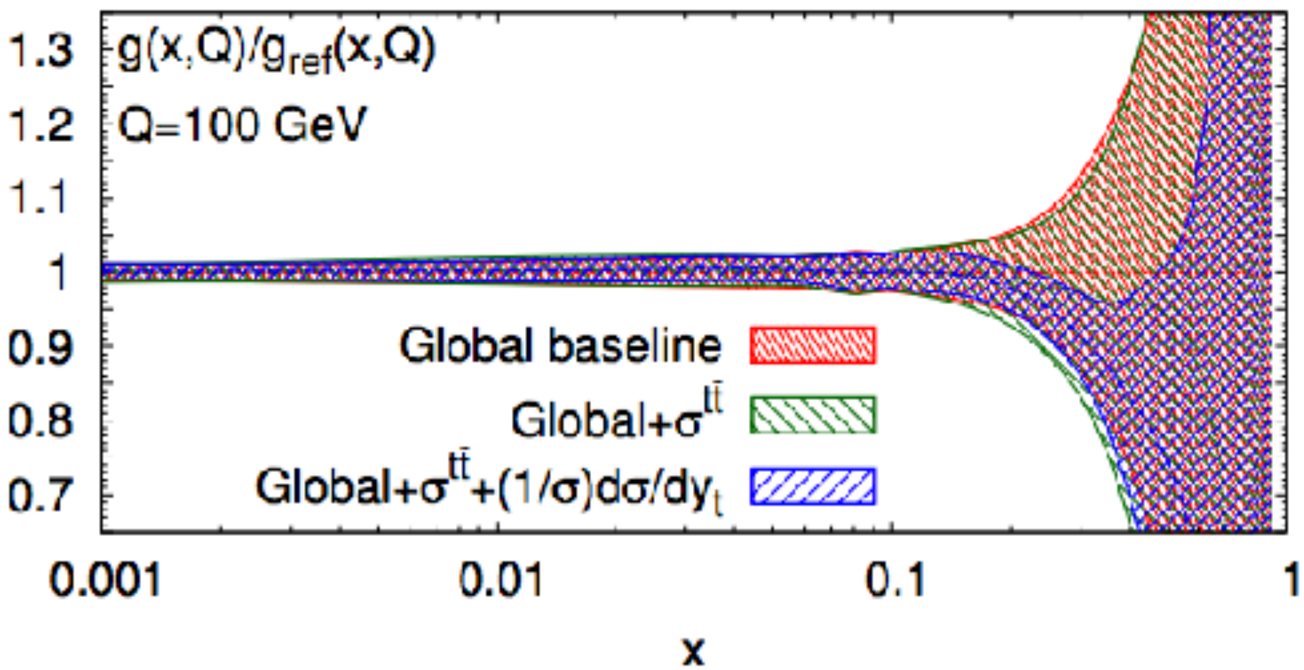
$$\frac{\sigma_{\mathrm{NNLO}} \times \mathscr{L}_{\mathrm{NNLO}}}{\sigma_{\mathrm{NLO}} \times \mathscr{L}_{\mathrm{NNLO}}}$$

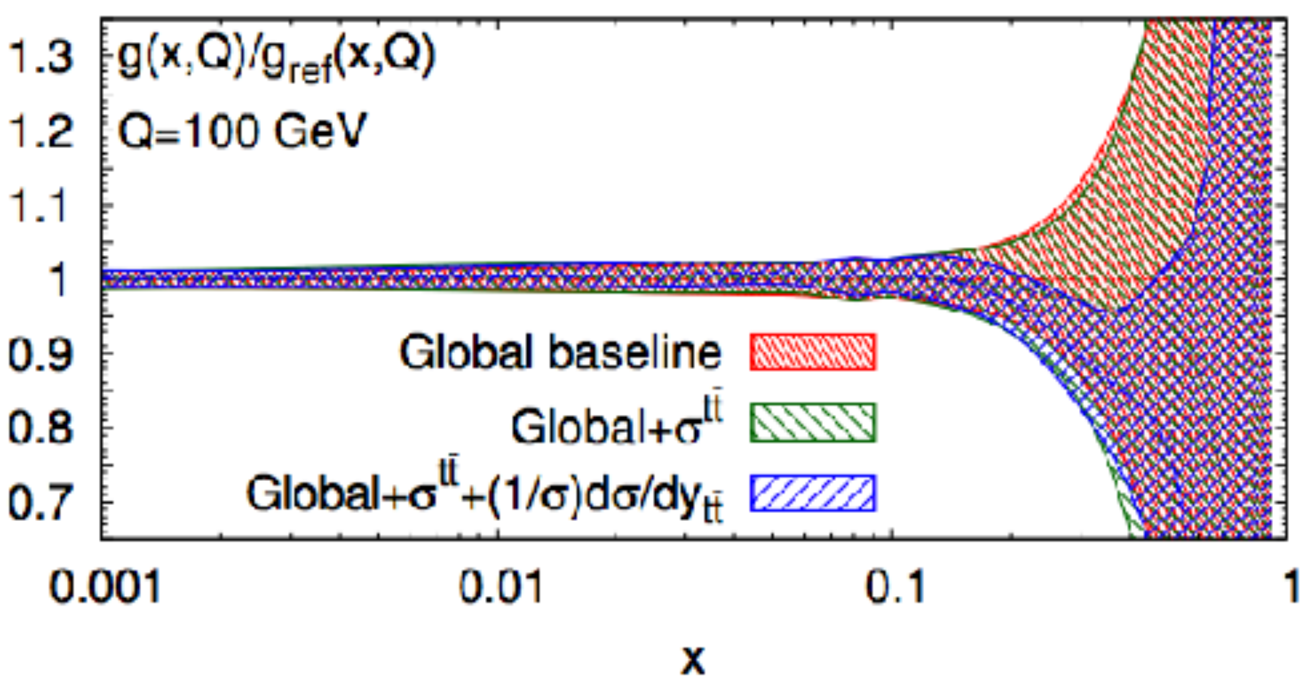


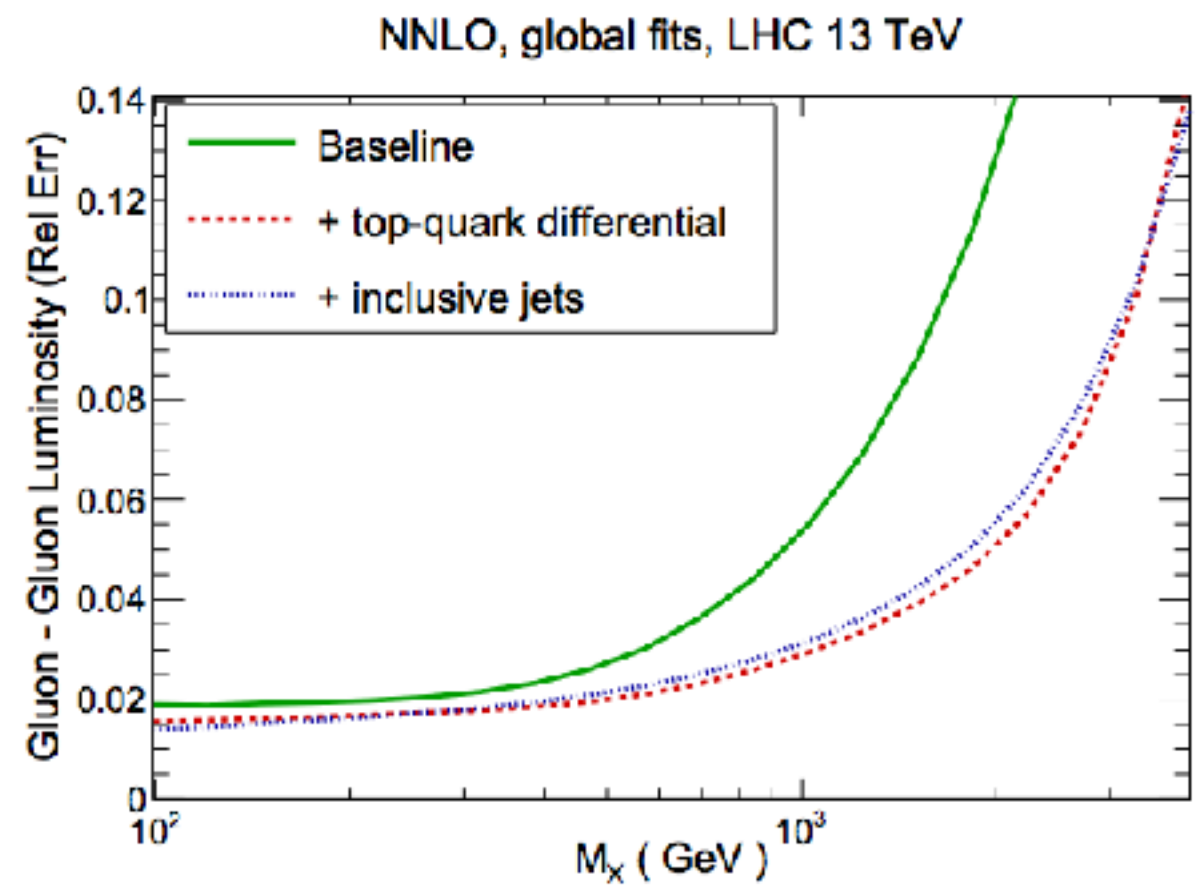


Czakon, Fiedler, Mitov [PRL 116(2016) 082003]

NNPDF3.1: top differential distributions



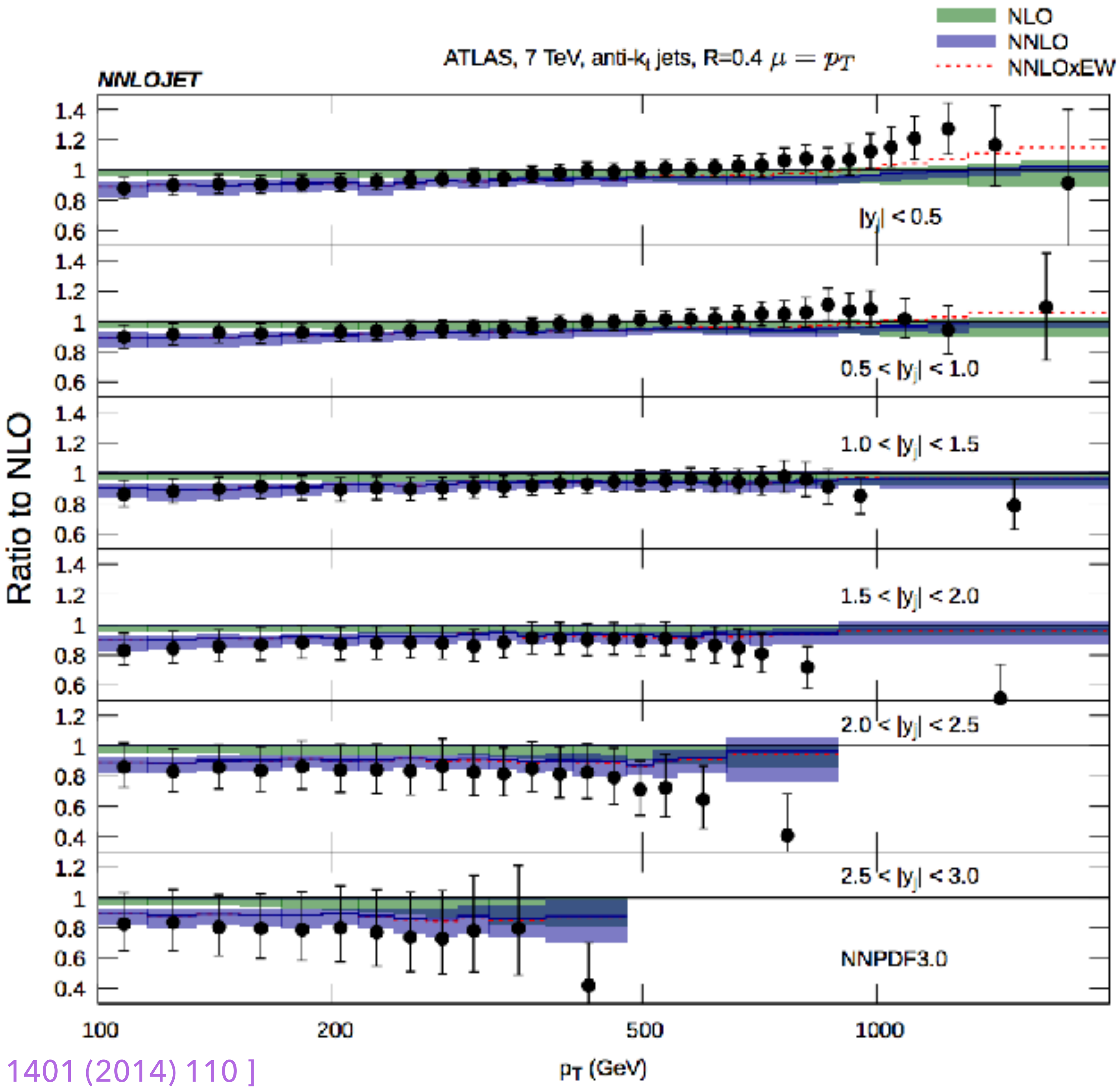




- Most constraining is inclusion of y_t list from ATLAS and y_{tt} from CMS jointly with total xsec
- Competitive reduction of gluon uncertainty with jets measurement
- Slight tension between ATLAS and CMS in NNPDF3.1 ($\chi^2_{ATLAS} \sim 1.6$, $\chi^2_{CMS} \sim 0.9$)

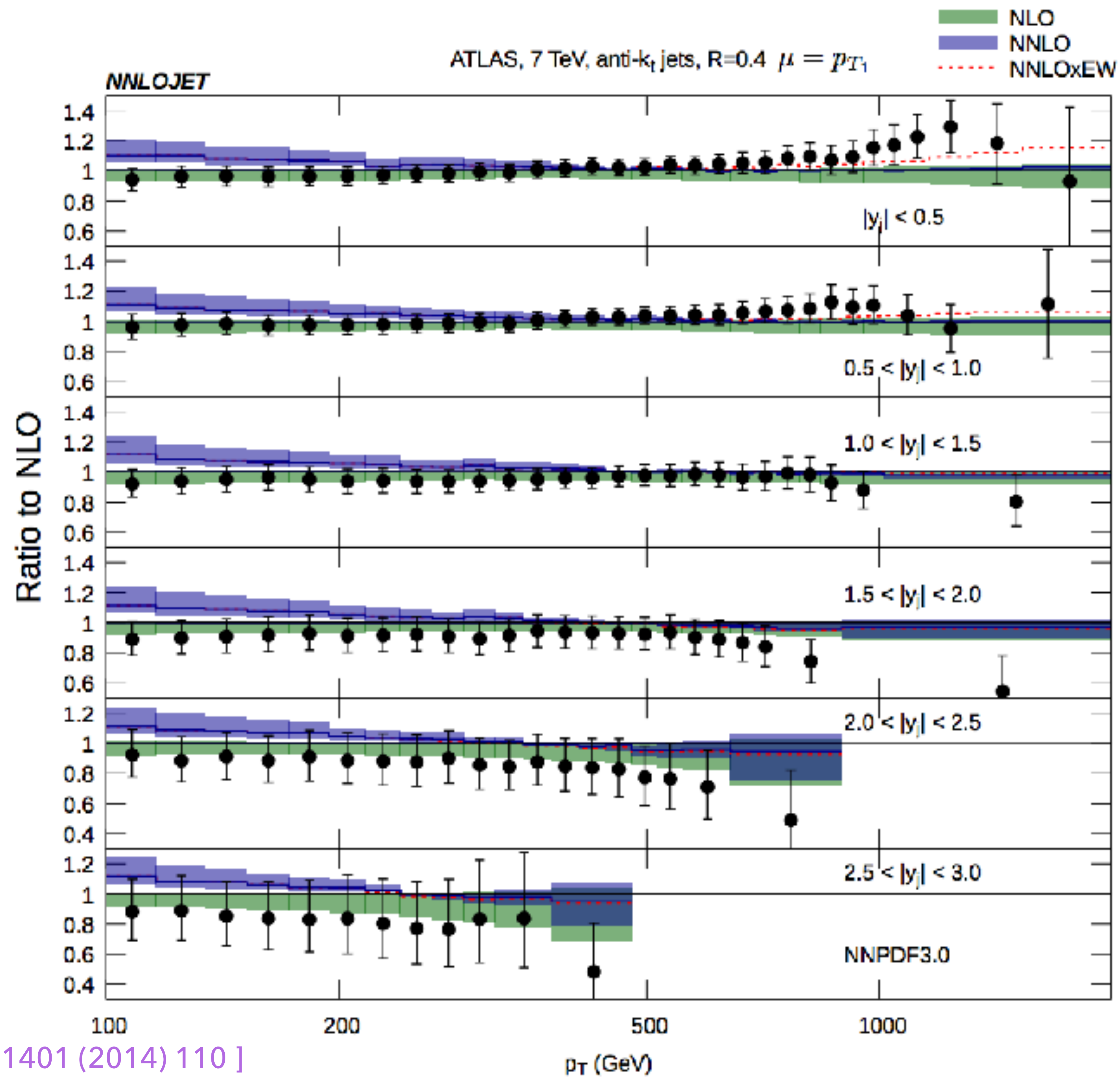
Czakon, Hartland, Mitov, Nocera and Rojo, arXiv: 1611.08609

NNPDF3.1: inclusive-jet data



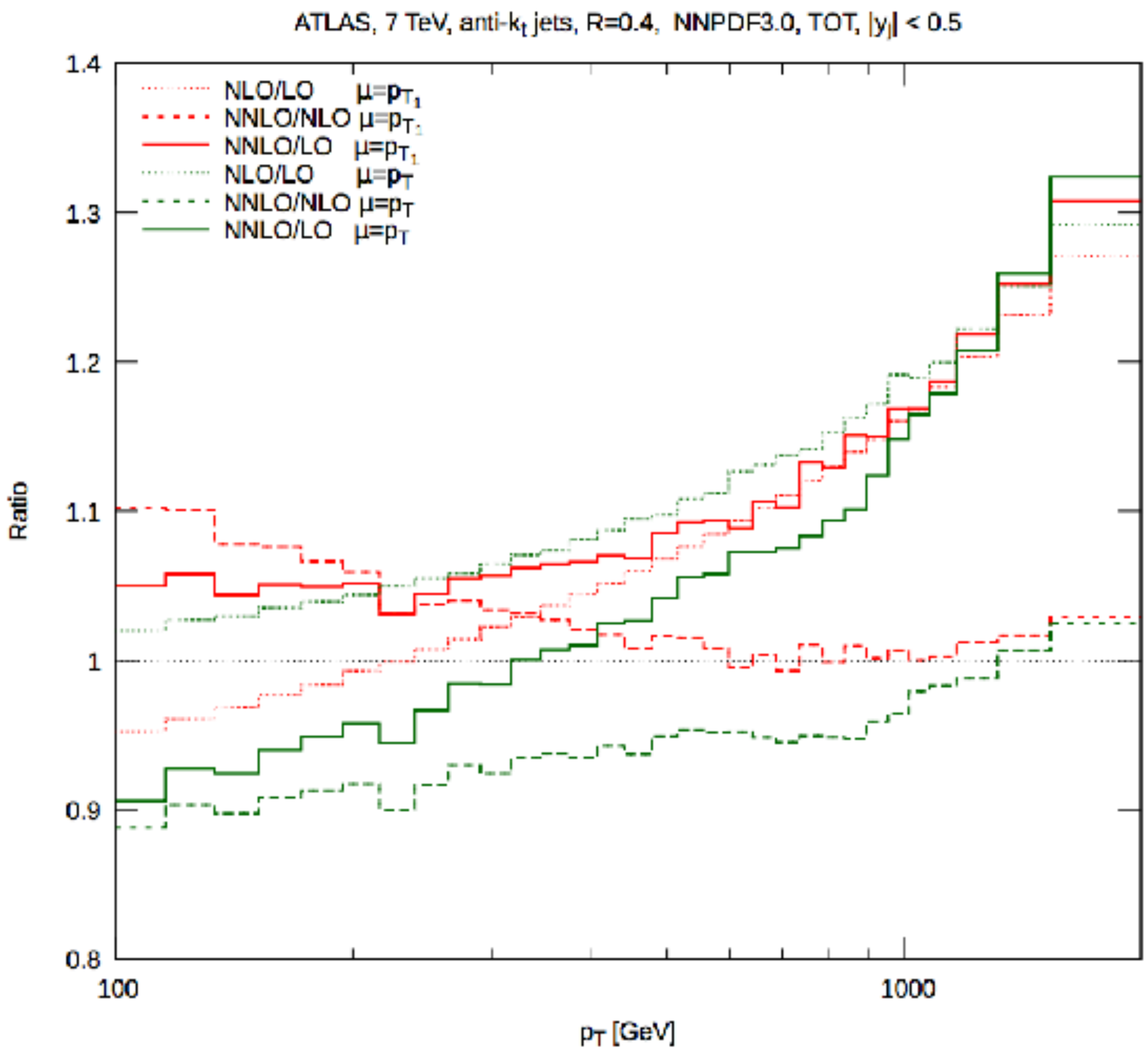
Currie et al [JHEP 1401 (2014) 110]

NNPDF3.1: inclusive-jet data



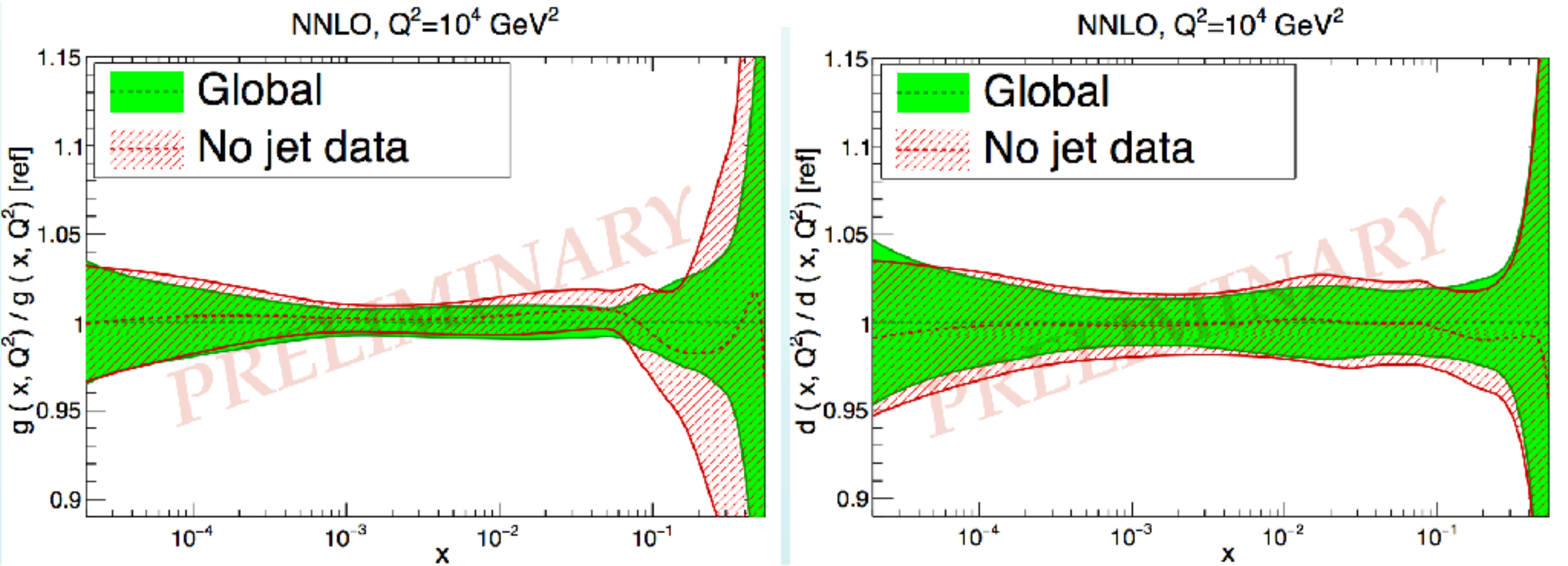
Currie et al [JHEP 1401 (2014) 110]

NNPDF3.1: inclusive-jet data



- NNLO corrections known for all partonic channels (leading colour contribution only)
- Different scales predict opposite behaviour of the K-factor
- NNLO/NLO K-factors available only for ATLAS 7 TeV data
- In NNPDF3.1 use NLO matrix elements for jets computed with individual jet pT as central scale and NLO scale uncertainty added as additional uncorrelated uncertainty

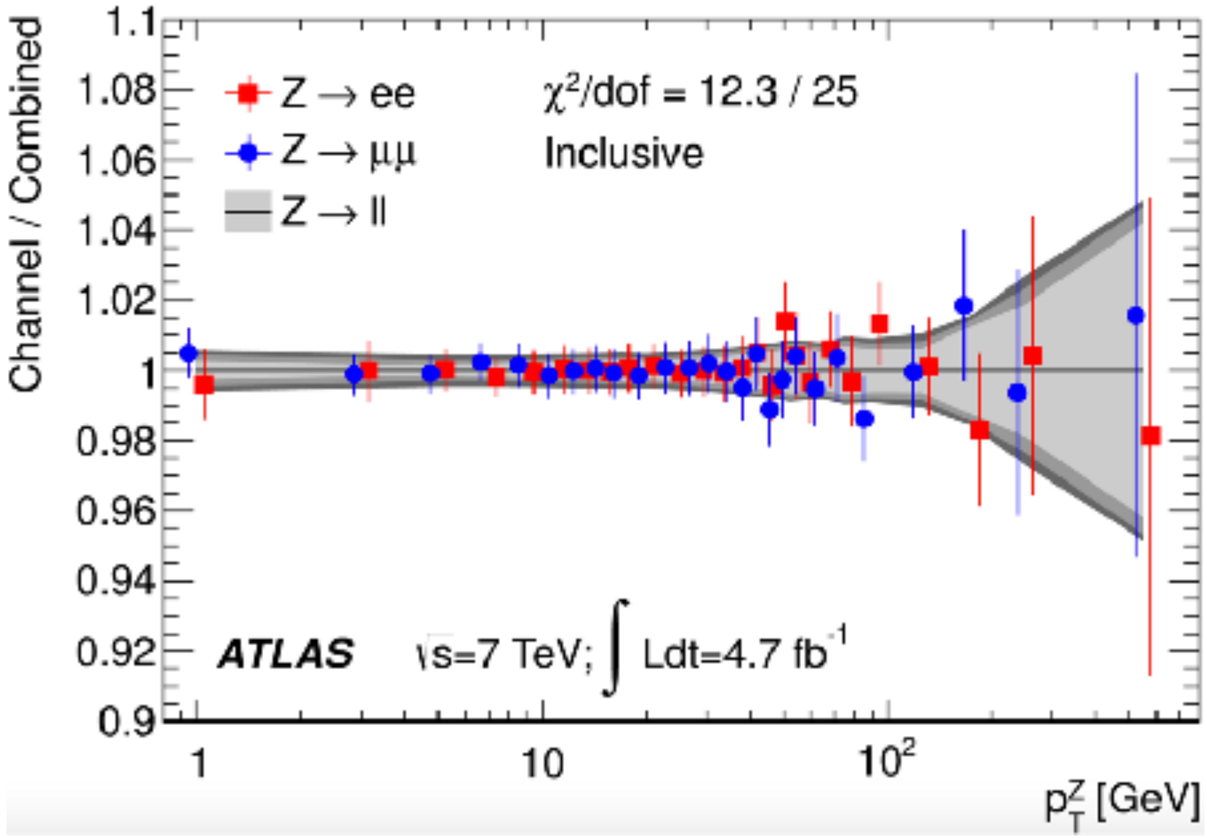
NNPDF3.1: inclusive-jet data

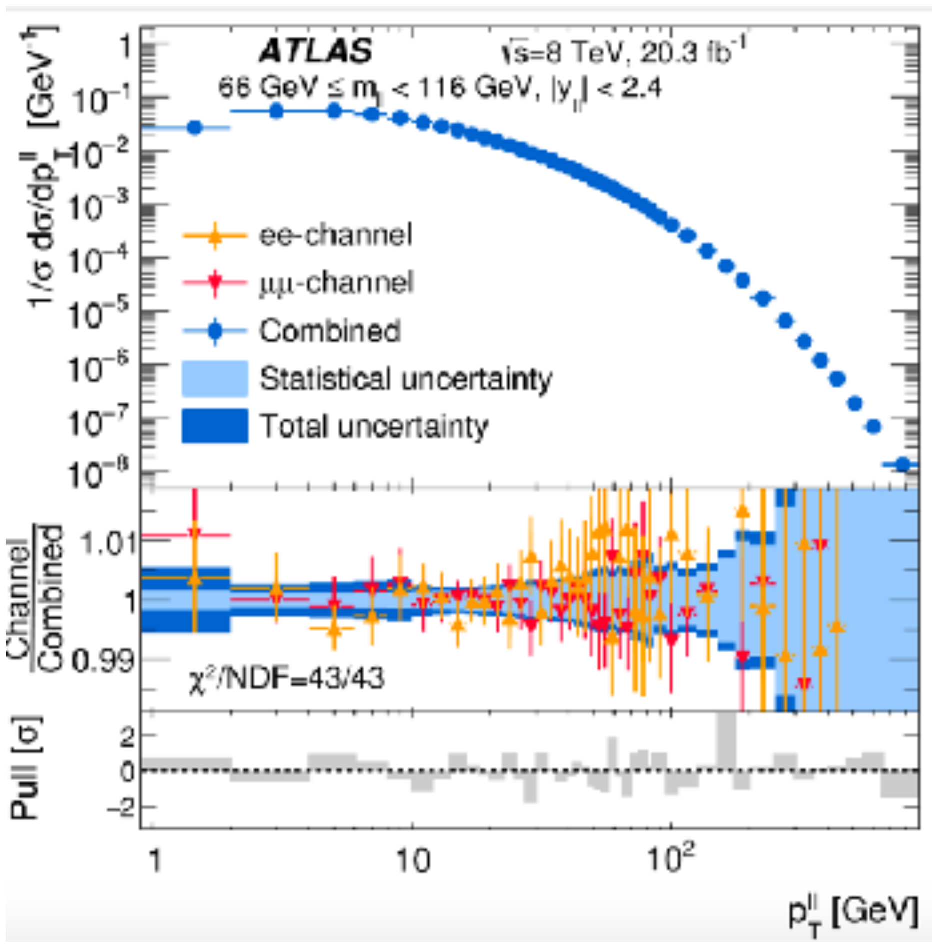


- In NNPDF3.1 included only central rapidity bin with good fit quality $\chi^2_{NLO} = 1.06$, $\chi^2_{NNLO} = 1.12$
- When all rapidity bins are included and full bin-by-bin correlation kept into account then description of the data becomes very bad
- Given that NLO scale uncertainty contains the NNLO - NLO shift, the issue is most likely related to experimental correlations

NNPDF3.1: Z p_T distributions

- Experimental precision < 1% up to pT~200 GeV
- Interesting case-study to probe current theory-experiment frontier





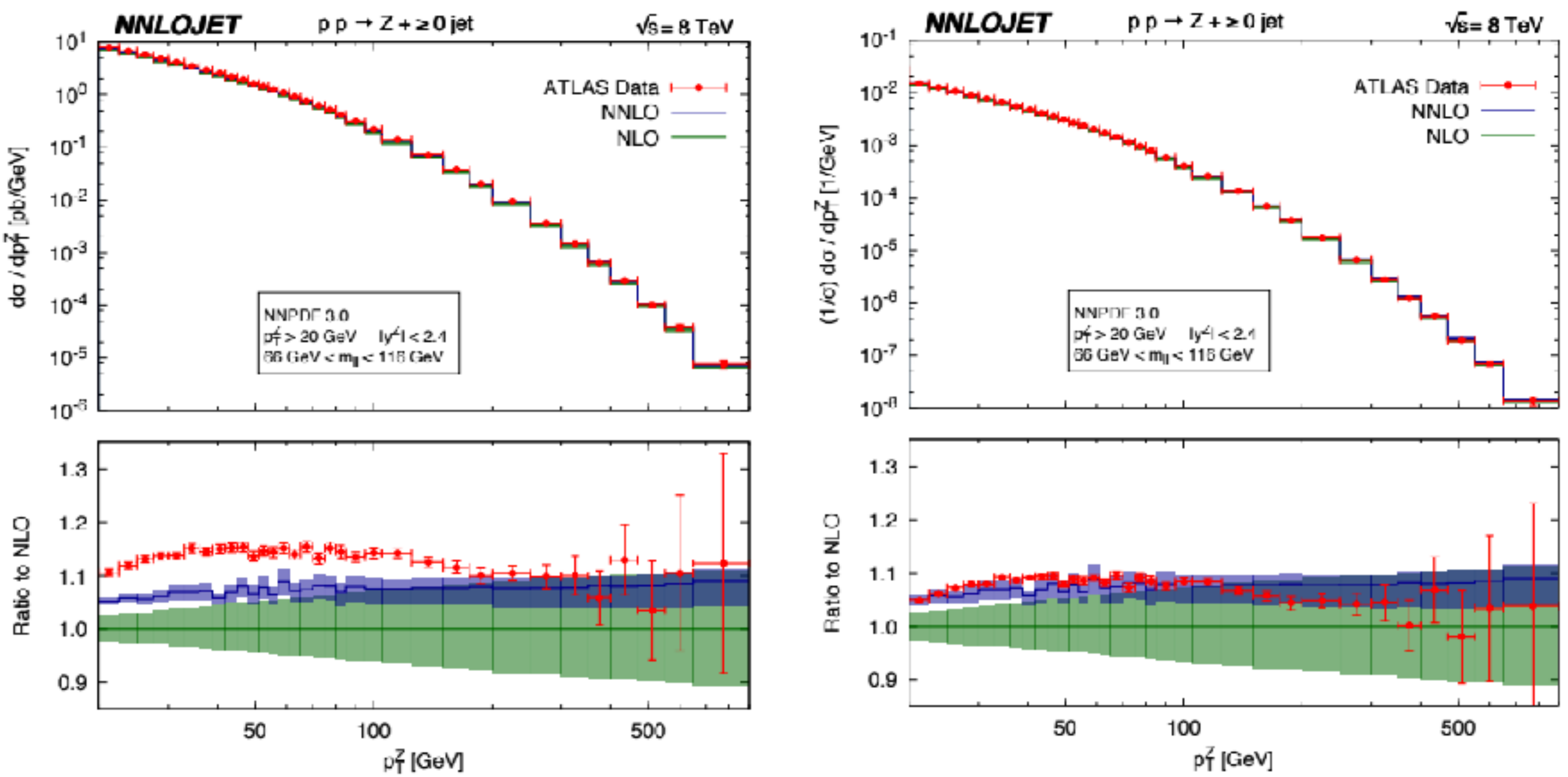
☑ ATLAS Z pT @LHC7, normalised distributions, 3 rapidity bins (0.0 < Y < 1.0, 1.0 < Y <2.0 , 2.0<Y<2.5)
~**50** data in perturbative region pT > 30 GeV

☑ ATLAS Z pT @LHC8, normalised/unnormalised distributions, 6 rapidity bins in Z peak + low/high M
~**150** data in perturbative region pT > 30 GeV

☑ CMS Z pT @LHC8, normalised/unnormalised distributions, 5 rapidity bins in Z peak
~**50** data in perturbative region pT > 30 GeV

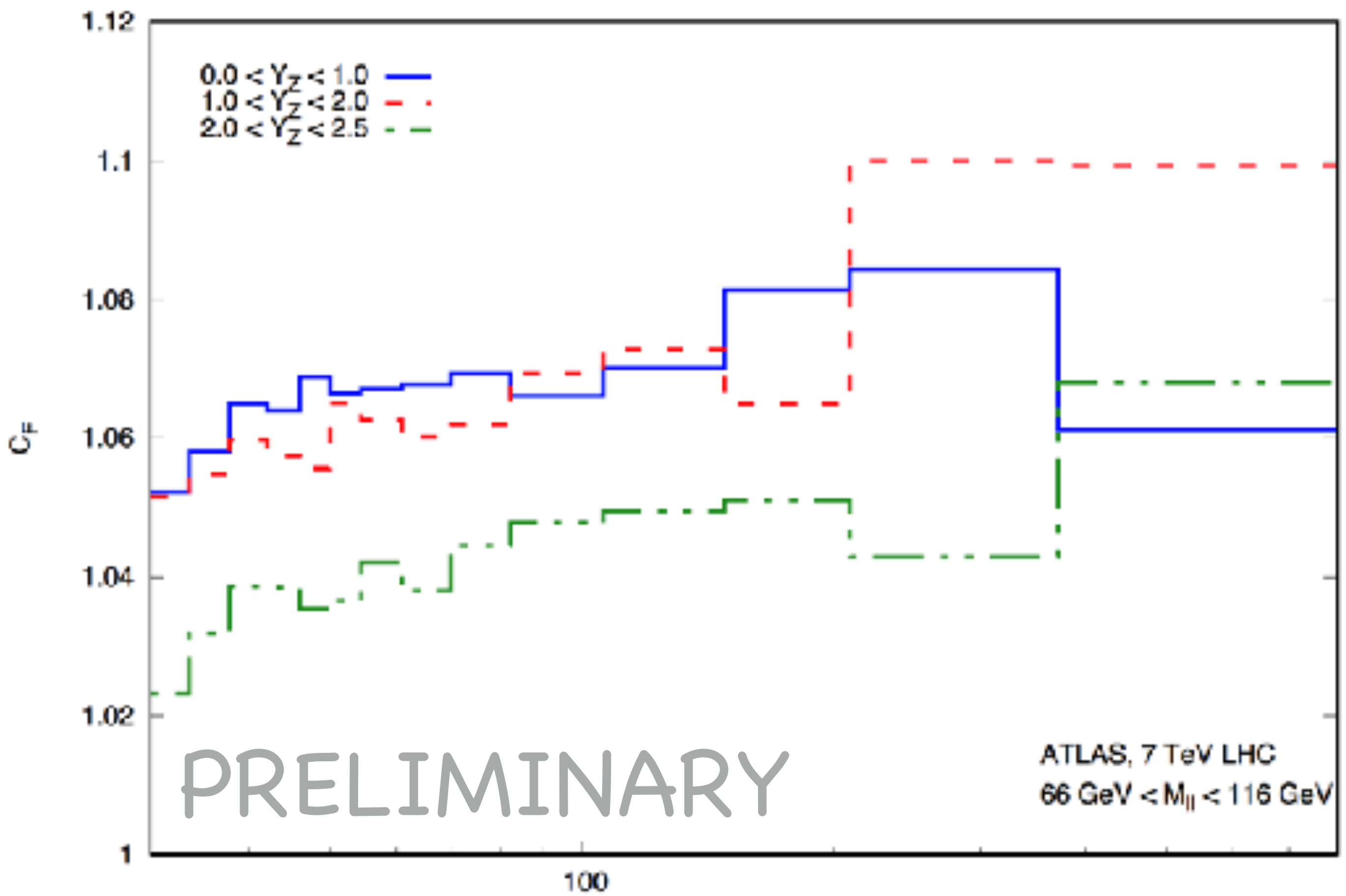
NNPDF3.1: Z p_T distributions

- Experimental precision < 1% up to pT~200 GeV
- Interesting case-study to probe current theory-experiment frontier



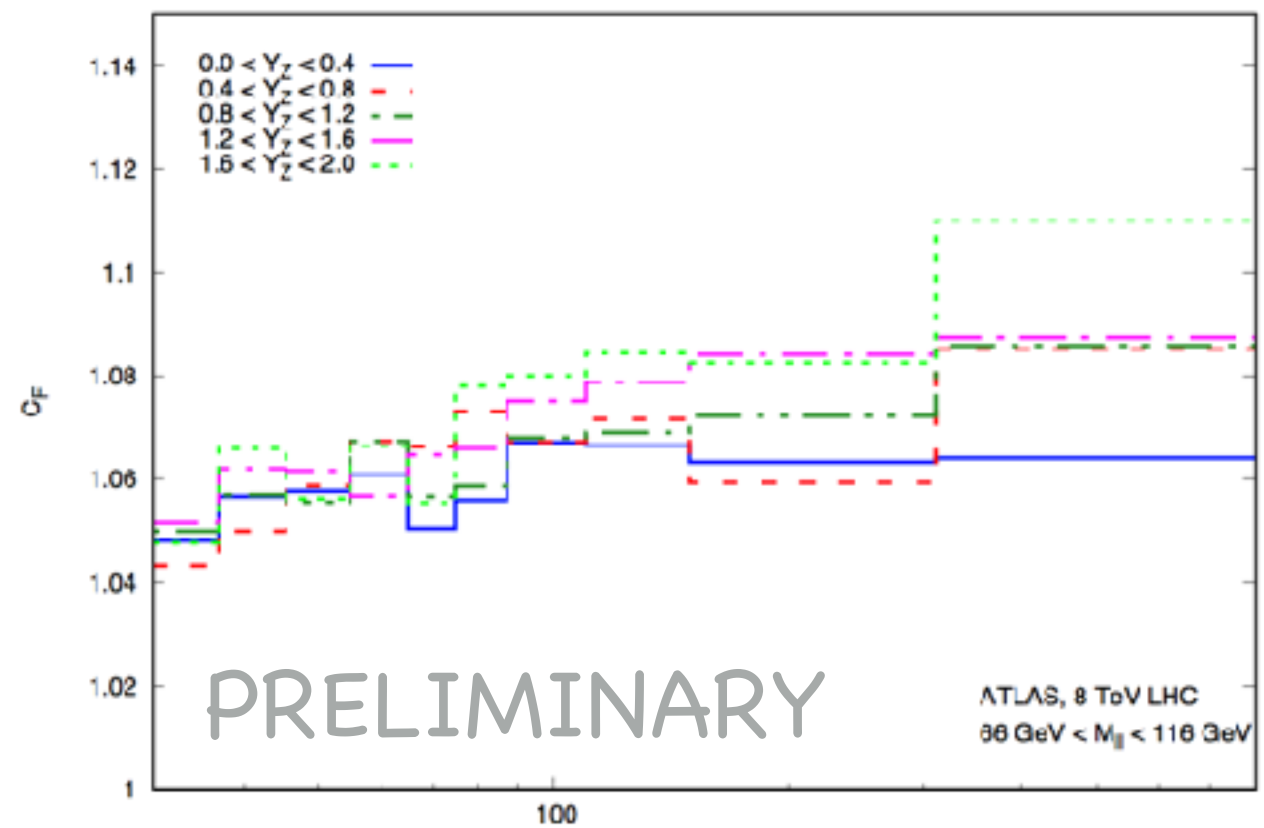
Gehrmann-De Ridder et al [1605.04295]

NNPDF3.1: Z p_T distributions



- NNLO calculation performed using N-jettiness subtraction scheme, by using recent calculation of Z+j at NNLO [Boughezal et al, PRL 116 (2016)] and relaxing cut on final state jet

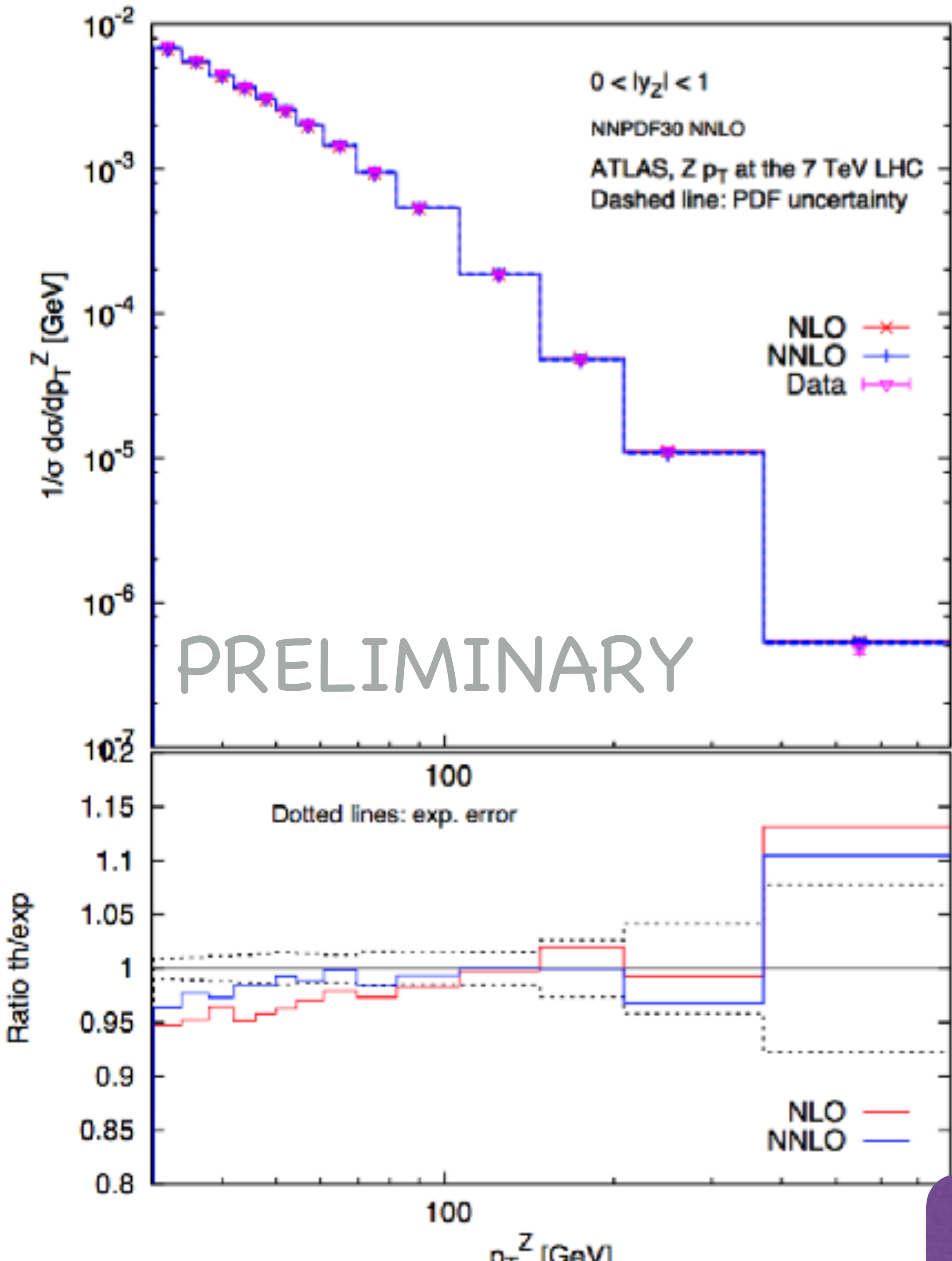
$$\mu_R = \mu_F = \sqrt{(p_T^Z)^2 + M_{ll}^2}$$



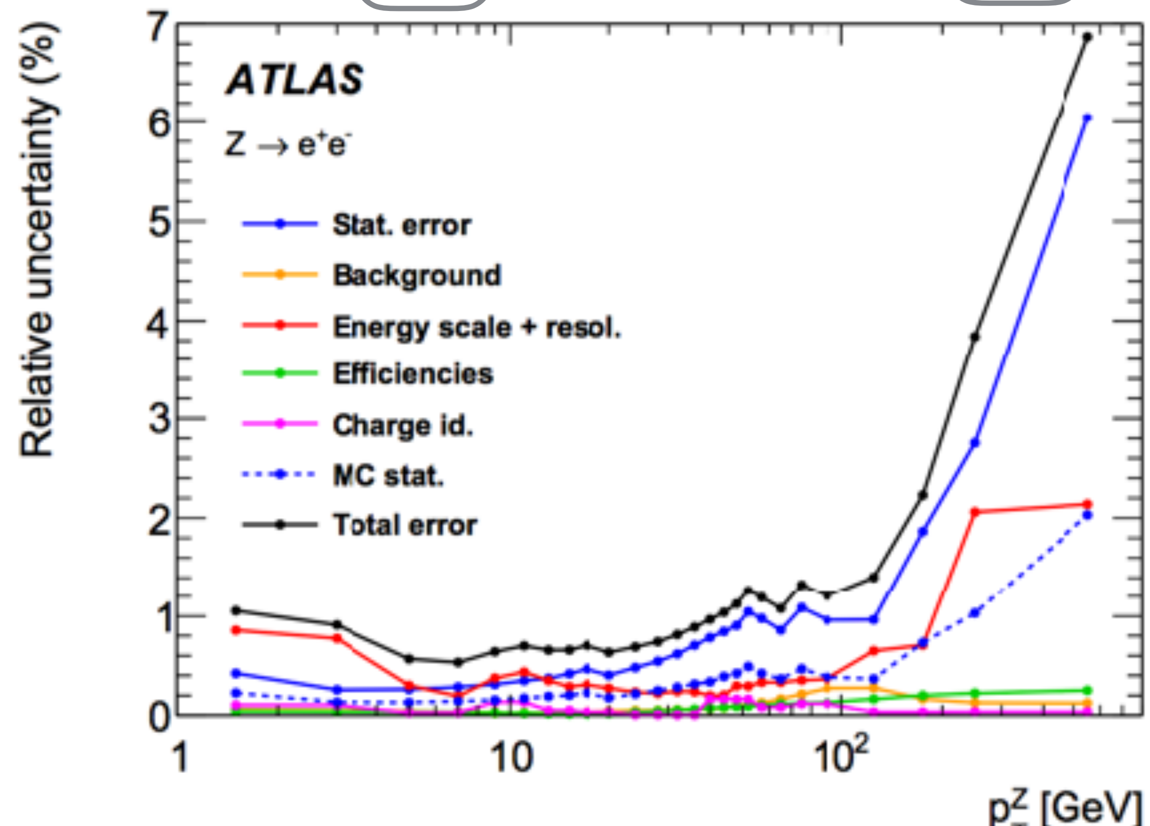
- NNLO/NLO K-factors 5% - 10% increase with p_T
- EW corrections only relevant for the highest pT bins in the Z-mass peak and for high-mass ATLAS measurement

Boughezal, Guffanti, Petriello, MU - in progress

NNPDF3.1: Z p_T distributions

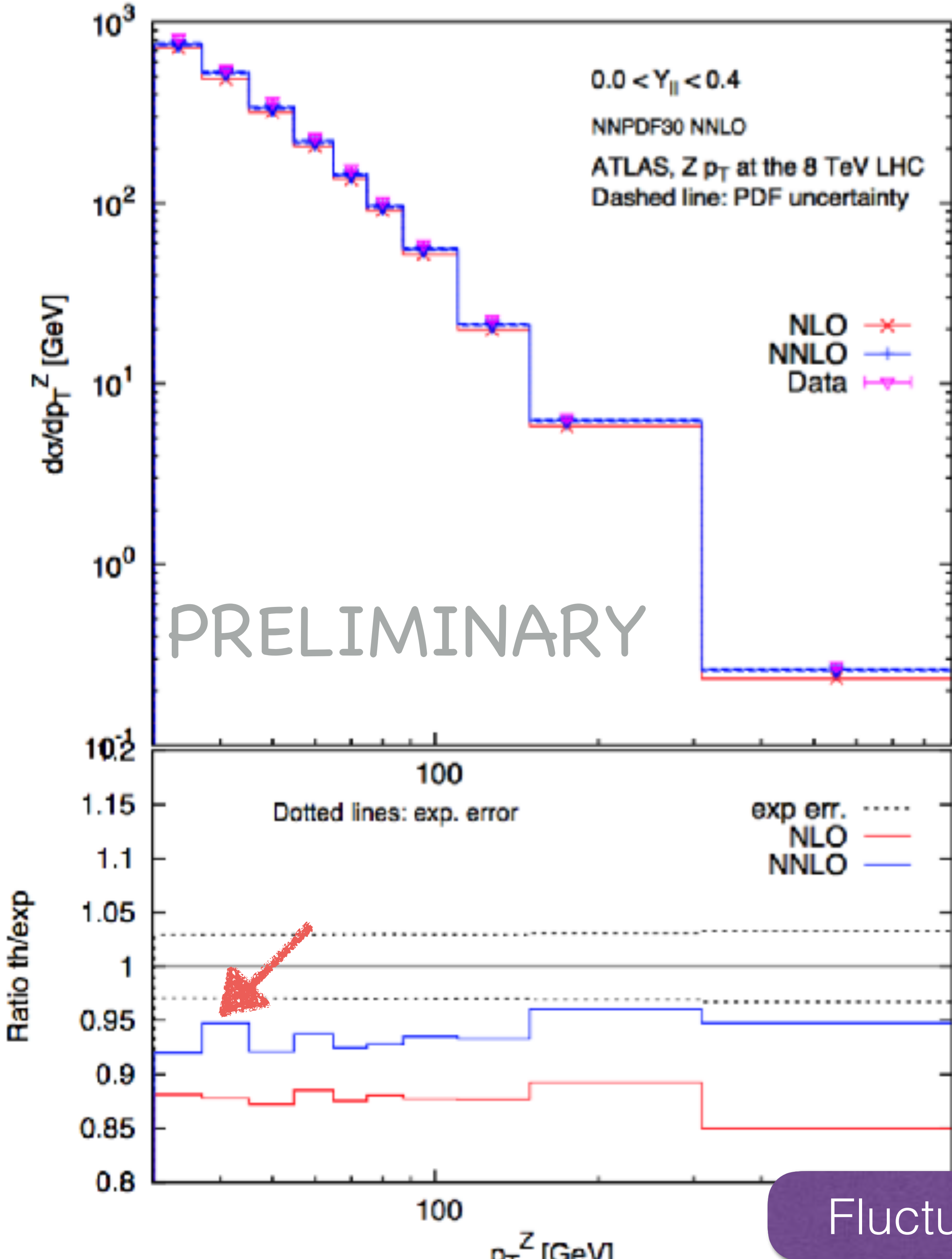


	NLO	NNLO	
	$\chi^2_{\rm stat} = 10.$	$\chi^2_{\rm stat} = 1.3$	/d.o.f.
	$\chi^2_{\rm uncor} = 7.3$	$\chi^2_{\rm uncor} = 0.9$	
	$\chi^2_{\rm diag} = 6.0$	$\chi^2_{\rm diag} = 0.7$	
	$\chi^2_{\rm full} = 16.$	$\chi^2_{\rm full} = 1.9$	



NNLO correction crucial to get the right shape!

NNPDF3.1: Z p_T distributions



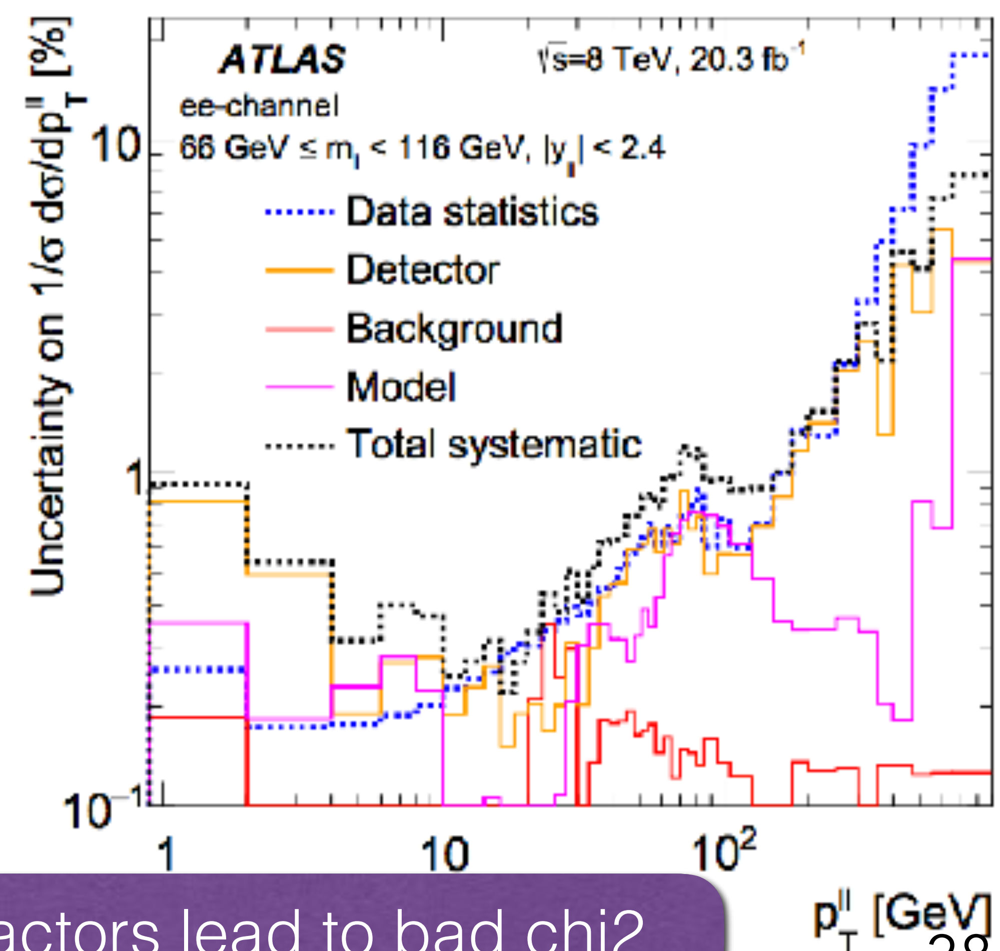
NLO

$$\chi^2_{\rm stat} = 140$$
$$\chi^2_{\rm uncor} = 115$$
$$\chi^2_{\rm diag} = 1.70$$
$$\chi^2_{\rm full} = 2.80$$

NNLO

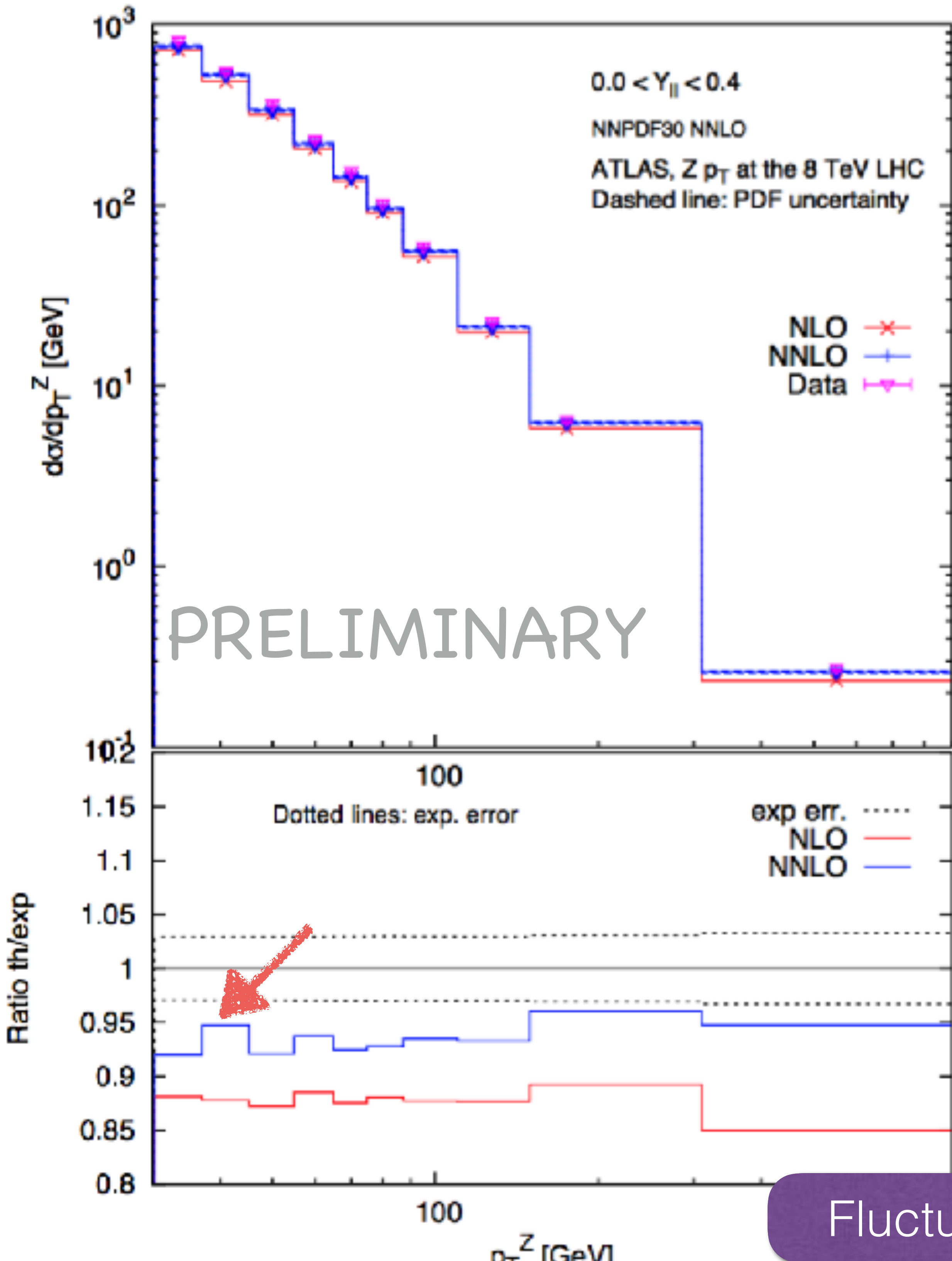
$$\chi^2_{\rm stat} = 51.$$
$$\chi^2_{\rm uncor} = 42.$$
$$\chi^2_{\rm diag} = 0.62$$
$$\chi^2_{\rm full} = 5.61$$

/d.o.f.



Fluctuations in K-factors lead to bad chi2

NNPDF3.1: Z p_T distributions



NLO

$$\chi^2_{\rm stat} = 140$$
$$\chi^2_{\rm uncor} = 115$$
$$\chi^2_{\rm diag} = 1.70$$
$$\chi^2_{\rm full} = 2.80$$

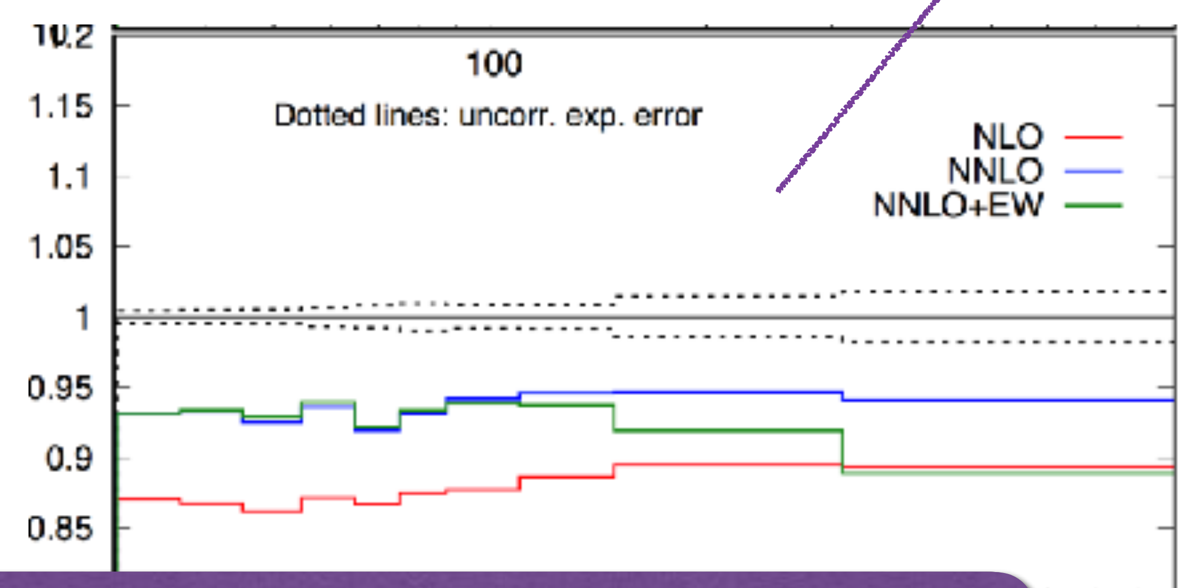
NNLO

$$\chi^2_{\rm stat} = 51.$$
$$\chi^2_{\rm uncor} = 42.$$
$$\chi^2_{\rm diag} = 0.62$$
$$\chi^2_{\rm full} = 5.61$$

/d.o.f.

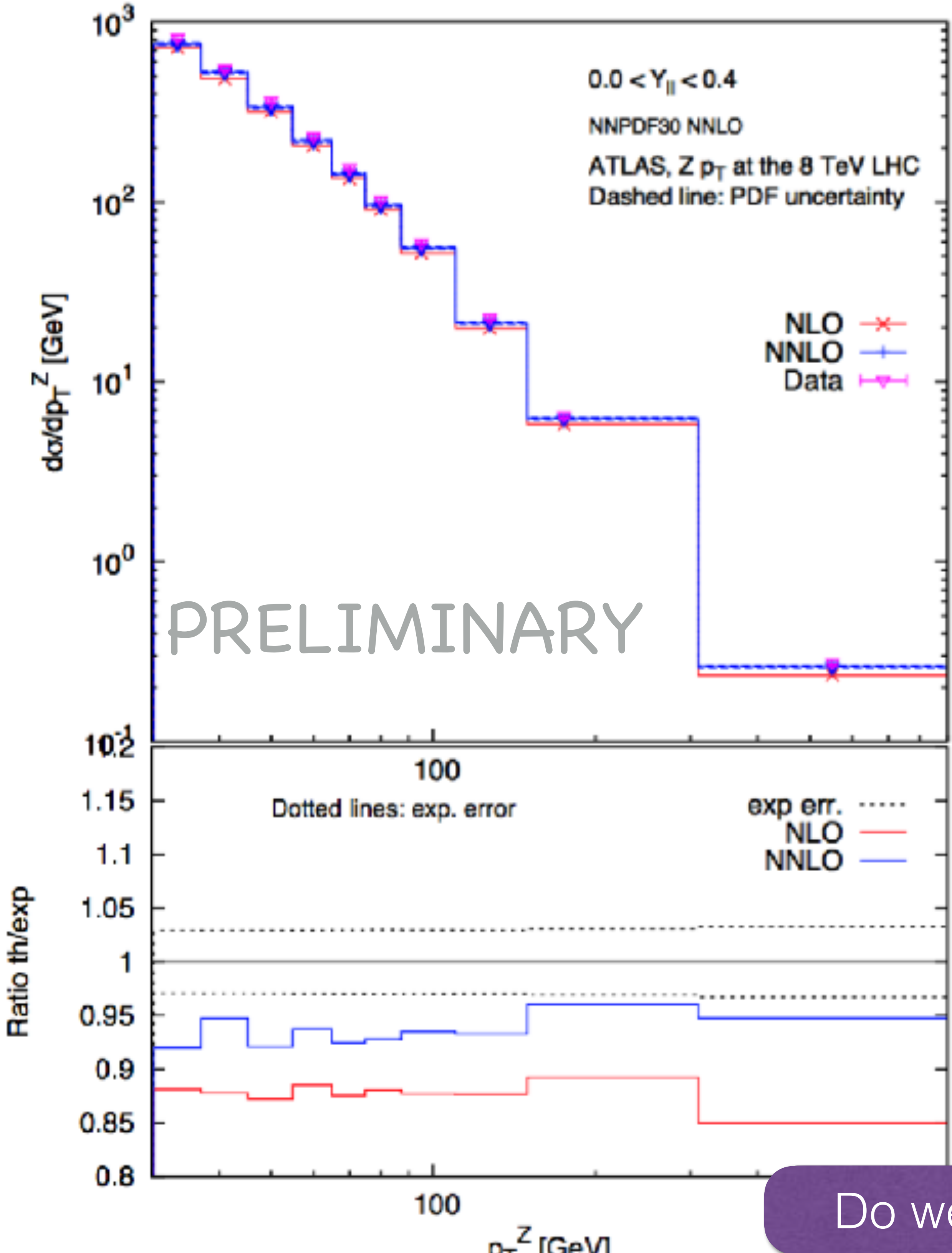
2.0

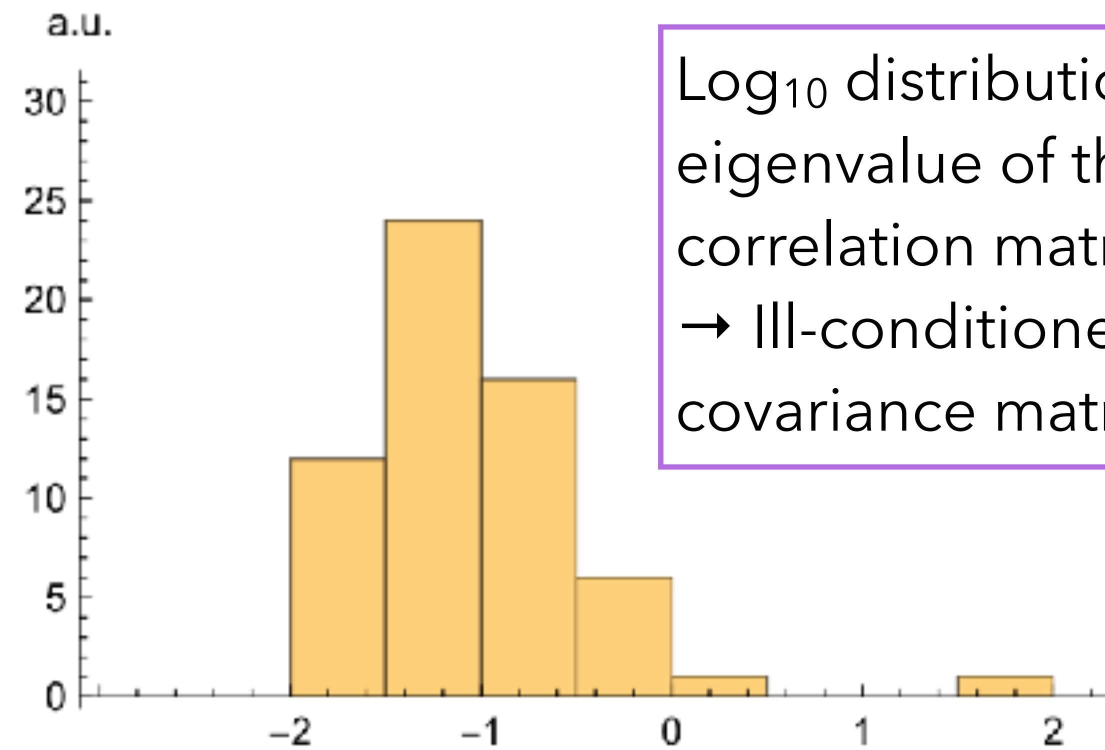
10x more statistics



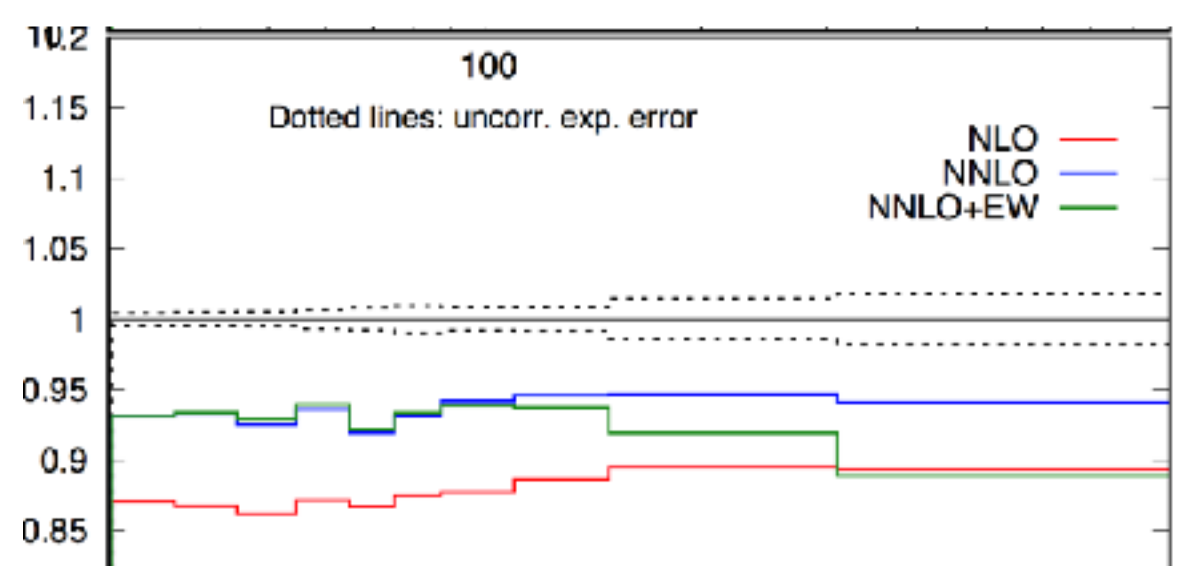
Fluctuations in K-factors lead to bad chi2

NNPDF3.1: Z p_T distributions





Log$_{10}$ distribution of eigenvalue of the correlation matrix → Ill-conditioned covariance matrix



Do we need uncertainty on covariance matrix?

NNPDF3.1: Z p_T distributions

- Uncorrelated uncertainties are very small, at the level of few per-mille
- This requires the shape of theory predictions to be correct to the same accuracy, which can be challenging for CPU-intensive NNLO calculations
- We tackle this by including the MC stat integration error from the theory prediction as an additional uncorrelated systematic error in the chi2
- This also implies that even very small variations of the correlation model (which ultimately determines what is correlated and what uncorrelated) can lead to very large variations of the chi2 for same input theory

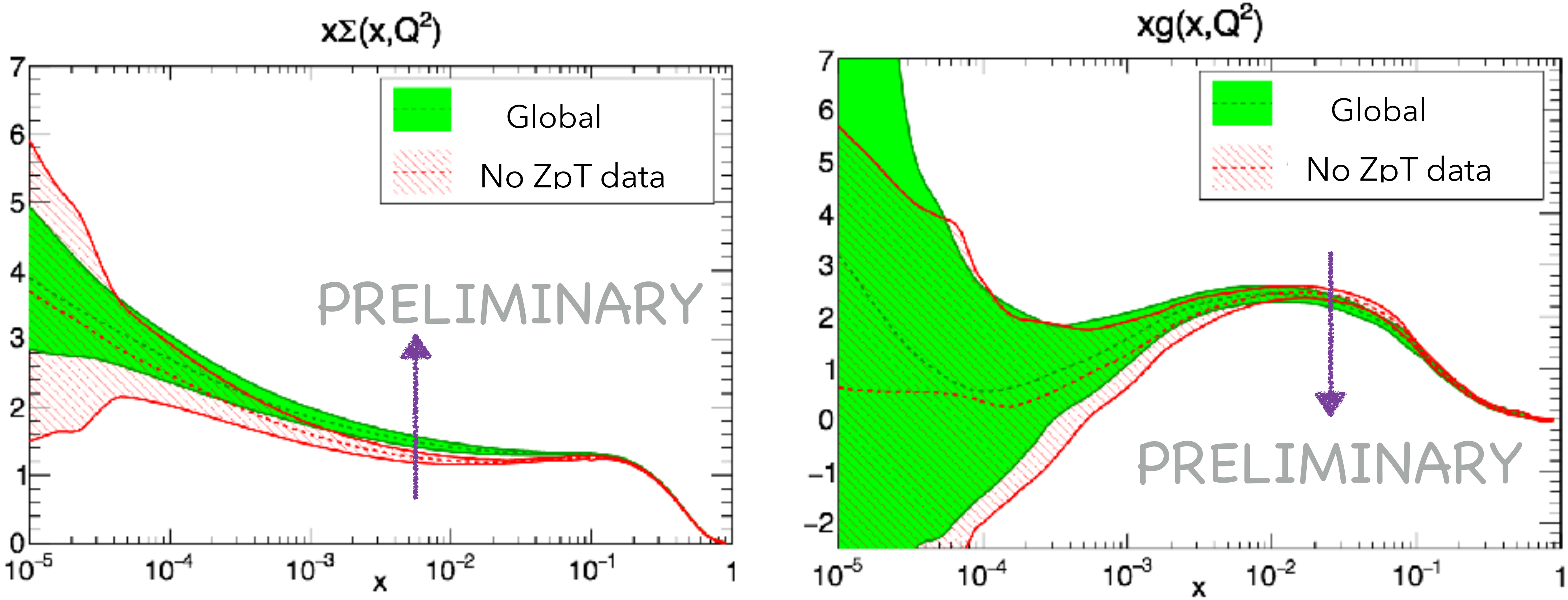
1% additional uncertainty

PRELIMINARY

$\chi^2_{\mathrm{ATLAS7tev}}$	$\chi^2_{\mathrm{ATLAS8tev,mdist}}$	$\chi^2_{\mathrm{ATLAS8tev,ydist}}$	$\chi^2_{\mathrm{CMS8tev}}$
(6.66)	(0.94)	(1.54)	(1.32)
3.31	(0.98)	(1.81)	(1.35)
(7.40)	0.93	1.40	(1.38)
(6.46)	(0.94)	(1.56)	1.30
3.80	0.98	1.69	1.34
(7.09)	0.93	1.46	1.33

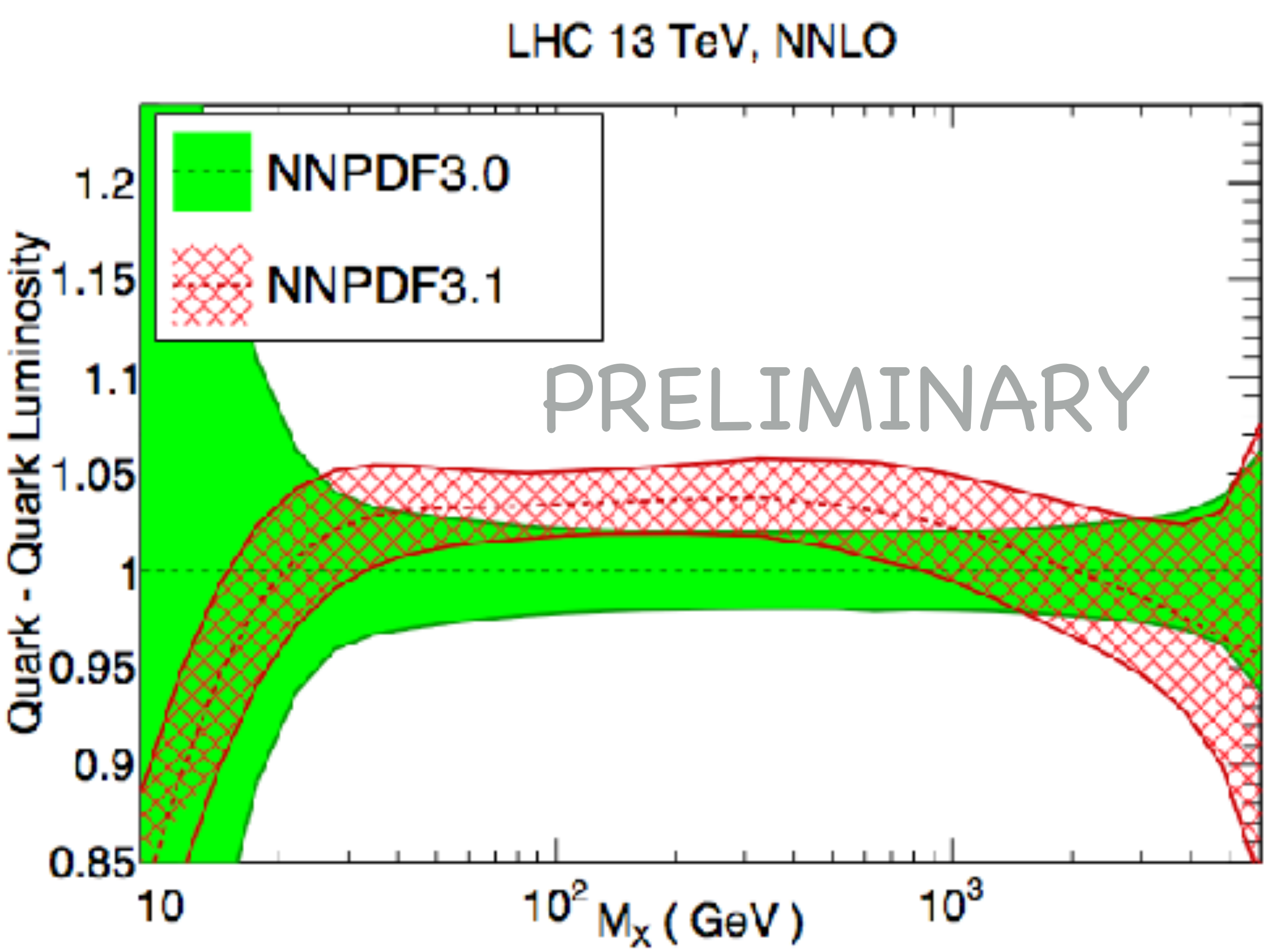
- Investigating on origin of the large chi2 for ATLAS 7 TeV (tension with global fit)

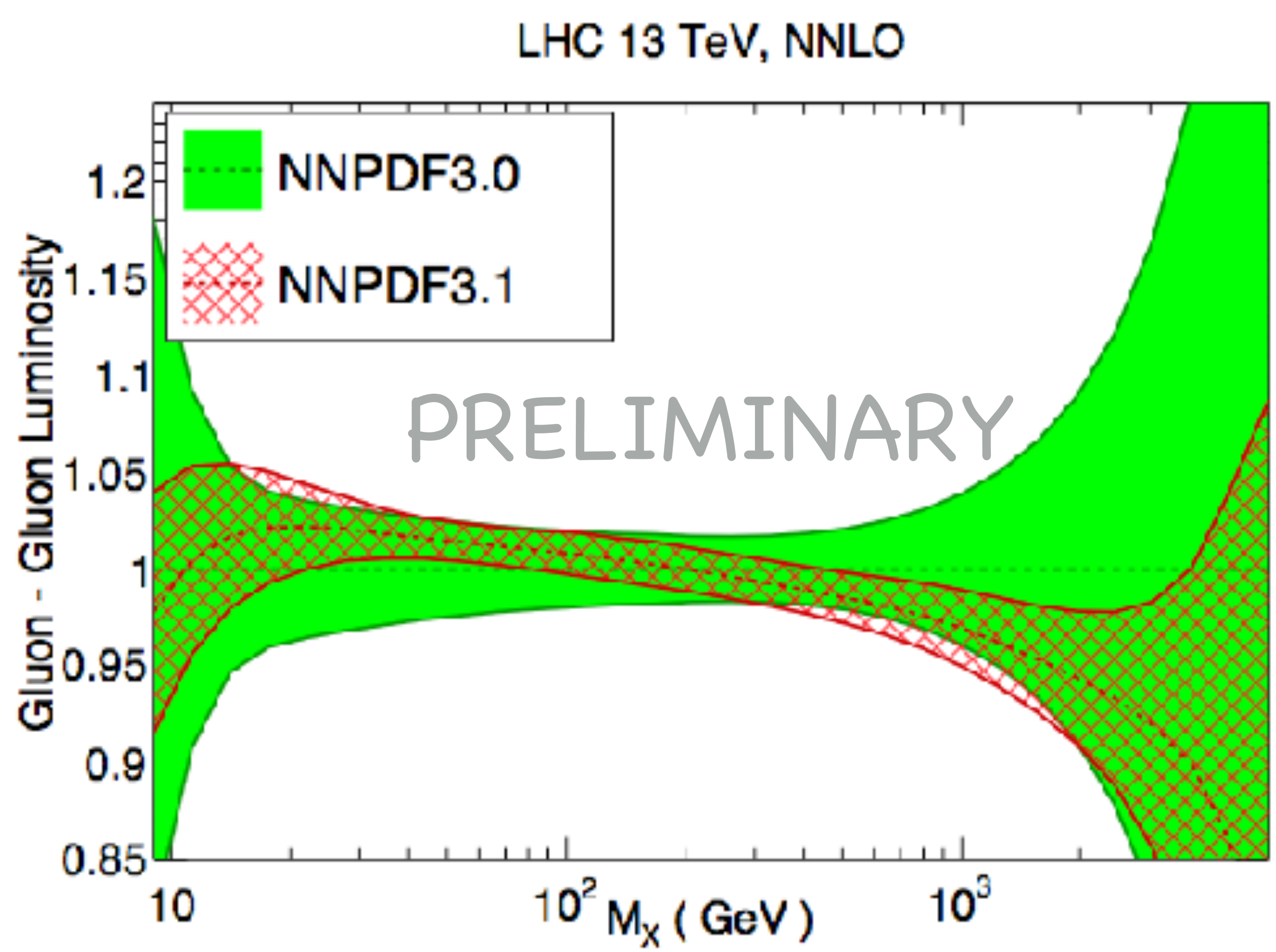
NNPDF3.1: Z p_T distributions



- Impact of Z pT distributions is quite strong, they increase the singlet and decrease the gluon in regions in which we expect them to be correlated with measurement
- Incompatibility between ATLAS 7 TeV data and global fit and ATLAS 7 TeV data and ATLAS 8 TeV data under investigation

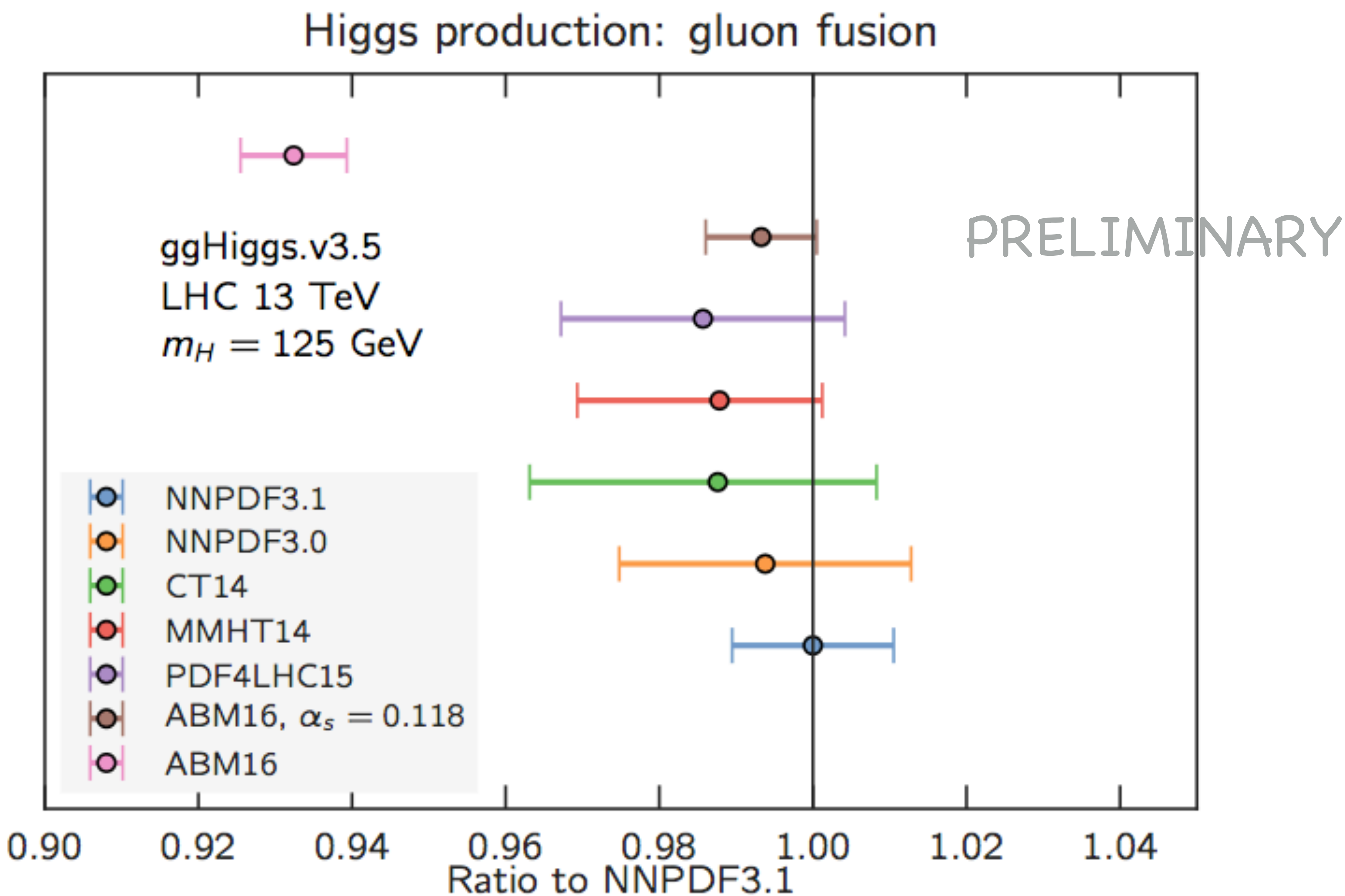
The NNPDF3.1 set



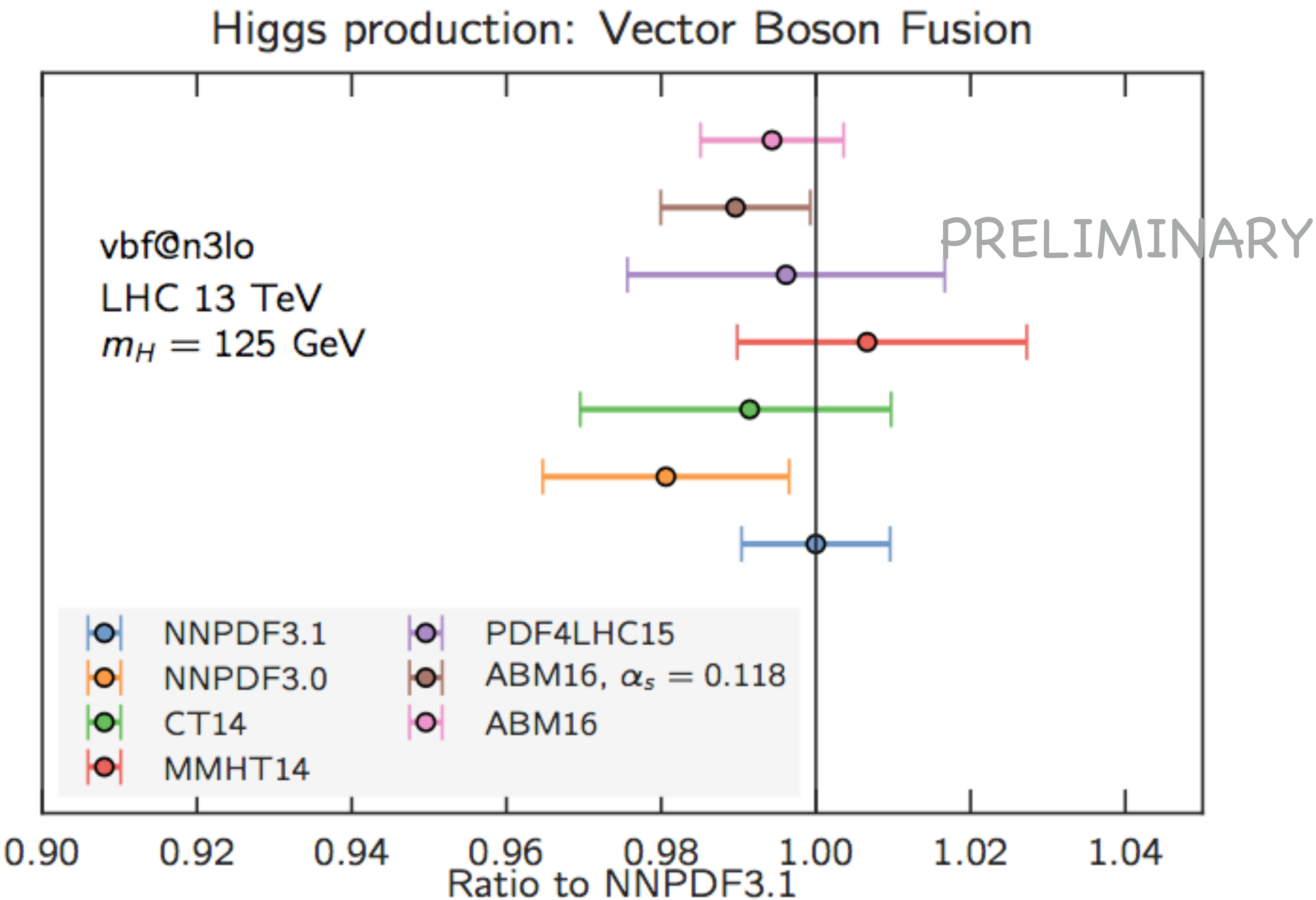


- Changes in gluon mostly due to new data, mostly reducing gluon uncertainty (top dist, jet dist, Z pT dist)
- Still under investigation, but jets, top and Z pT (8 TeV) seem to point in the same direction, no tension
- Changes in quarks due partially to new data (LHCb, Tevatron, CMS) and partially to fitted charm

Phenomenology

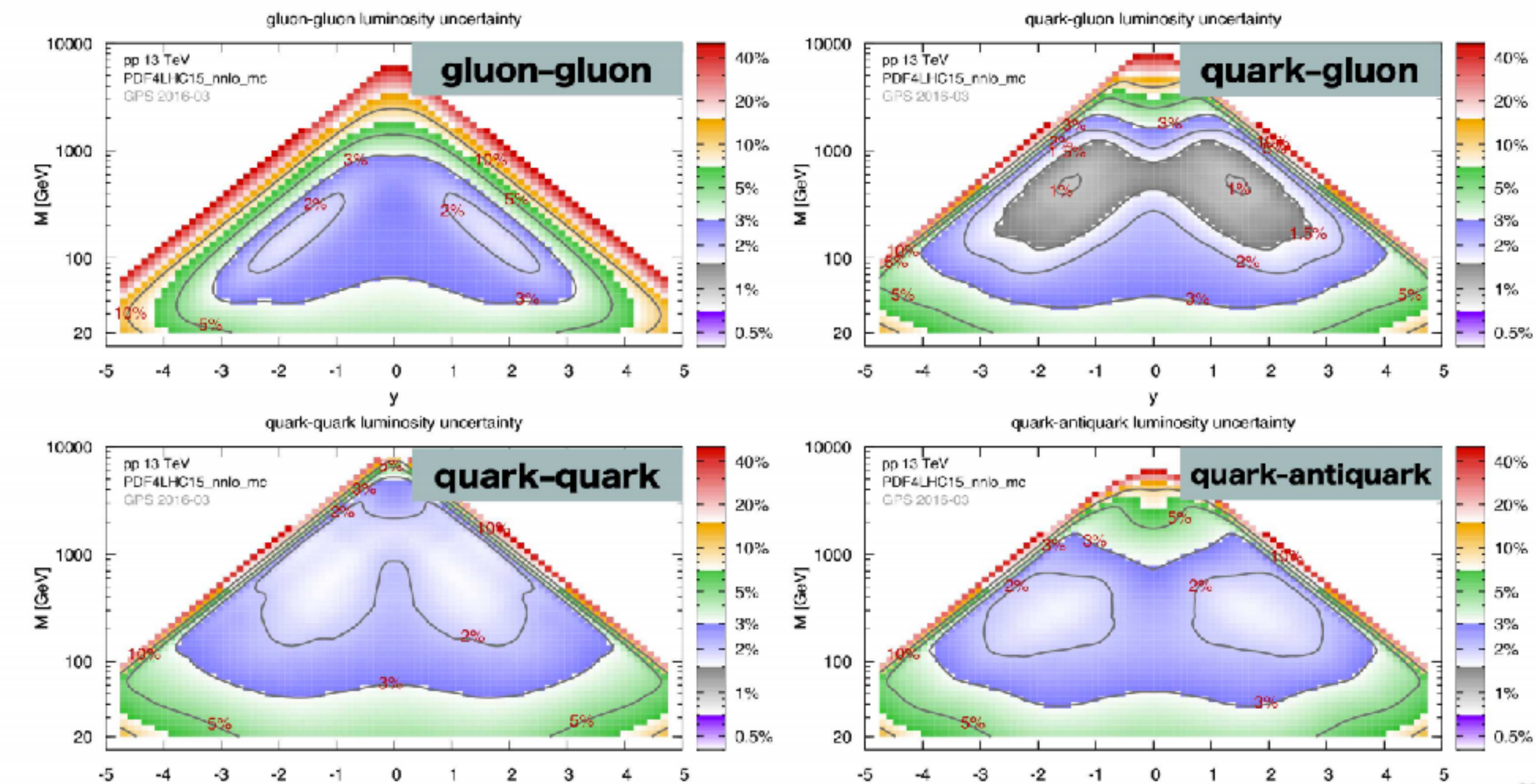


Phenomenology



New frontiers

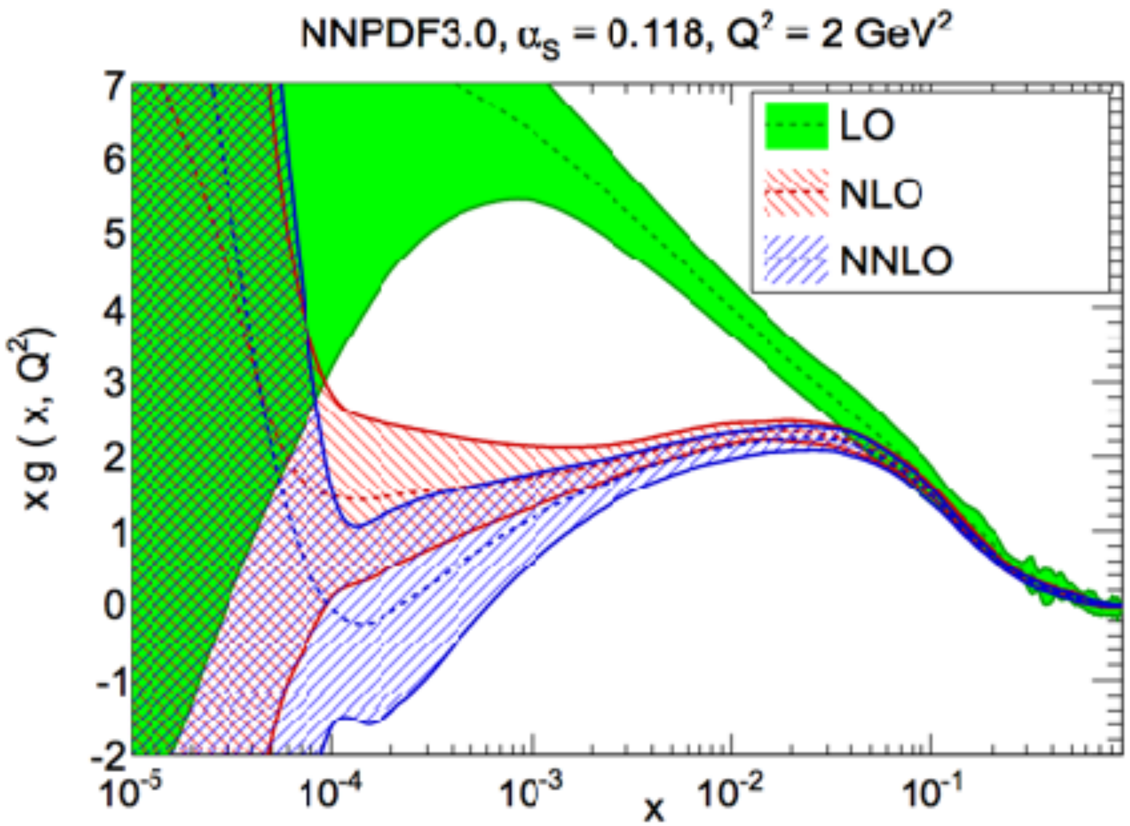
PDF uncertainties

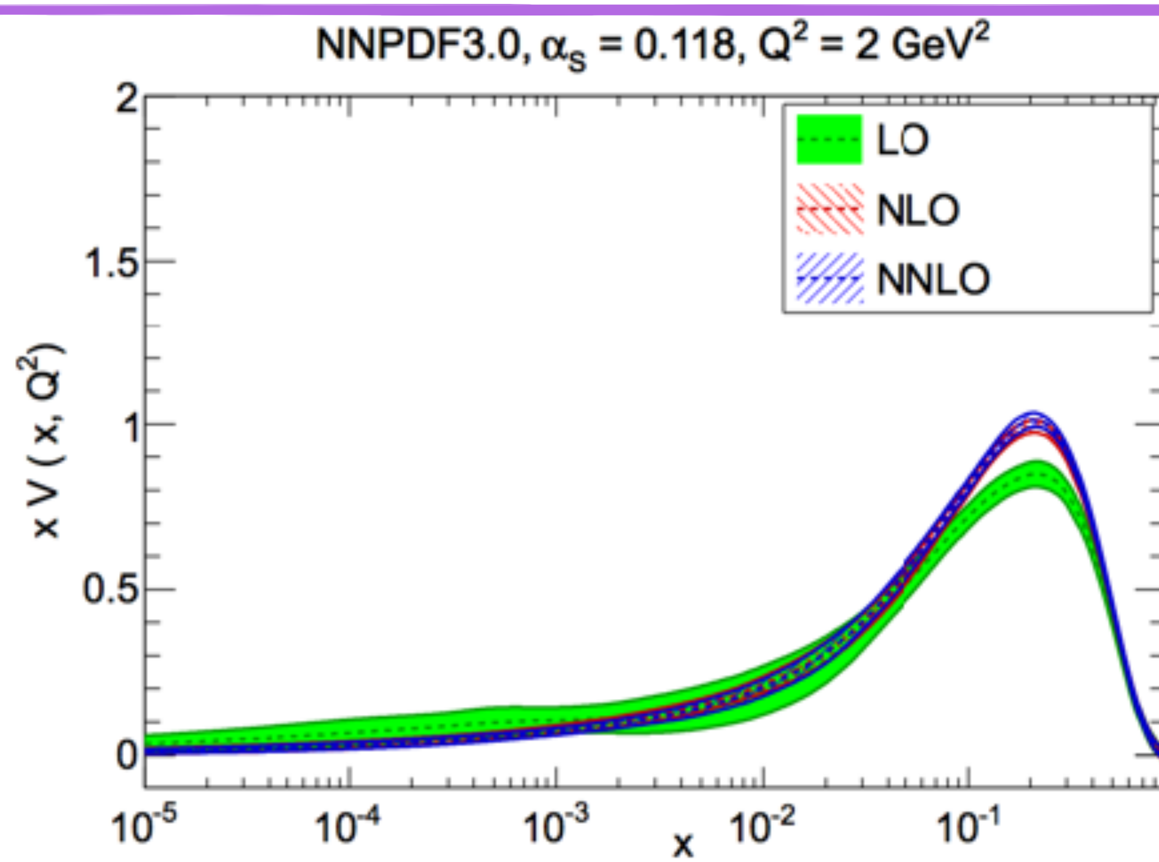


G. Salam, LHCP

Do we trust 1% accuracy in parton luminosities?

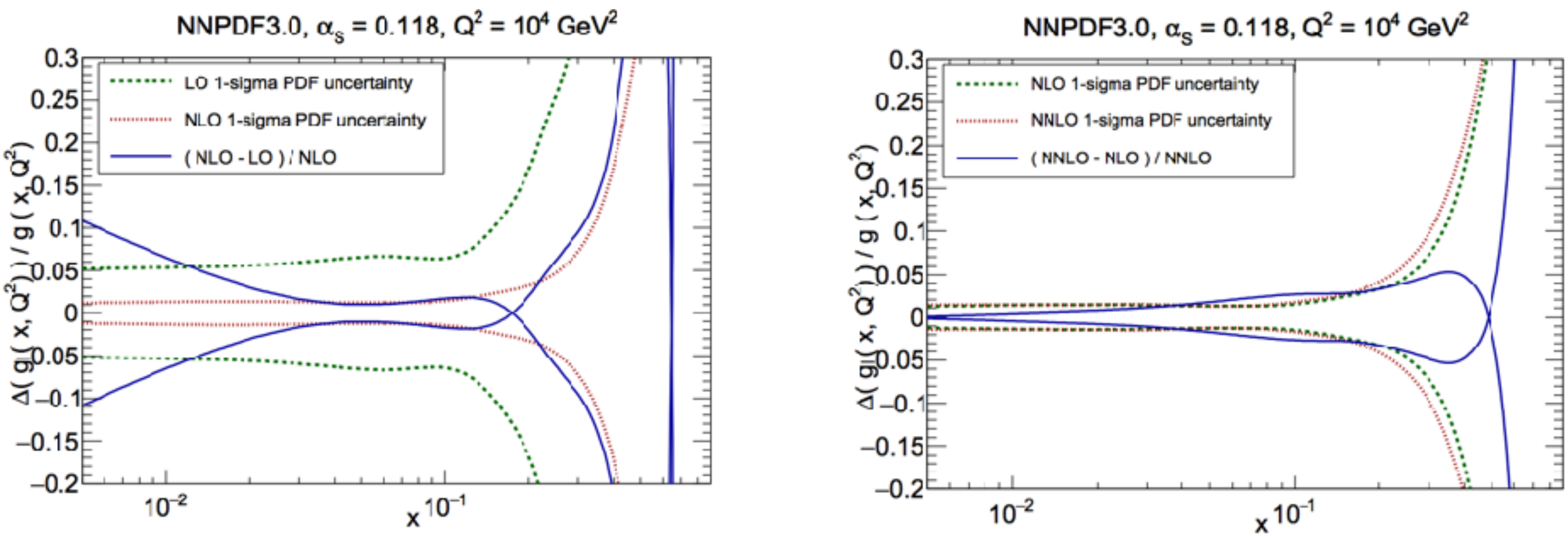
Theory uncertainties





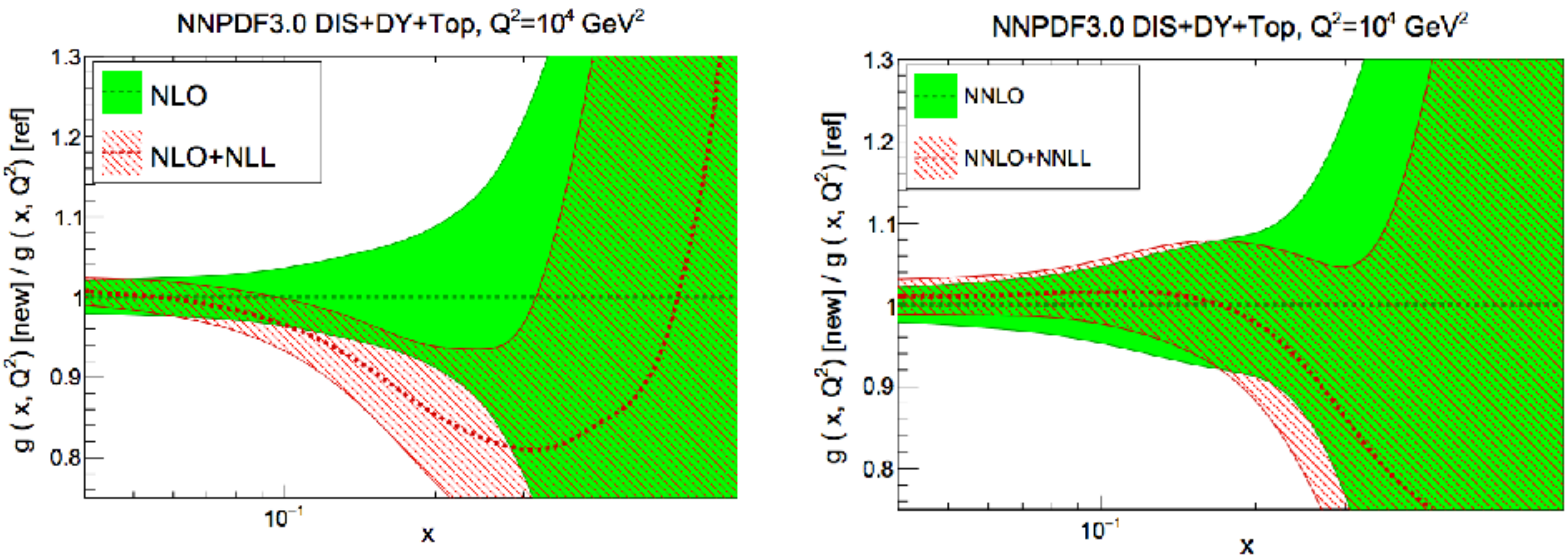
- PDF fits performed with given fixed perturbative order, value of α_S and heavy quark masses (estimated by combining PDF sets determined with different values
- PDF uncertainties only reflect lack of information from data given the theory
- Changes in theory may cause shifts outside the error band, can we estimate that?
- LO fits are merely qualitative, NLO quantitative and NNLO precise, but how much?

Theory uncertainties



- If we knew the next order we could compute the shift: at NLO theory uncertainty is comparable to the experimental one
- NNLO used to be considered subdominant, but now?
- Cacciari Houdeau method **[JHEP 1109 (2011) 039]** look at the behaviour of perturbative expansion promising
- What about NNNLO PDFs? Main bottleneck is missing anomalous dimensions
- Currently testing scale variations in NNLO fits

Beyond fixed-order accuracy

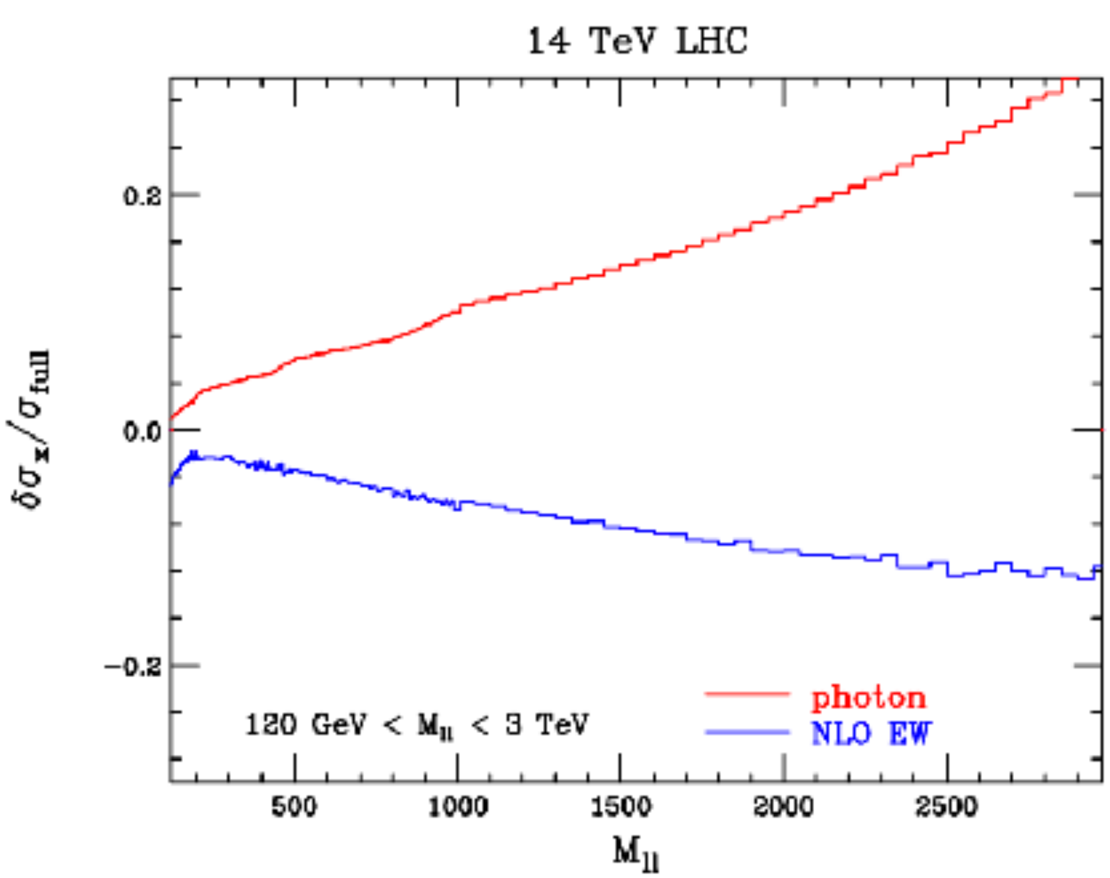


Bonvini et al, JHEP 1509 (2015) 191

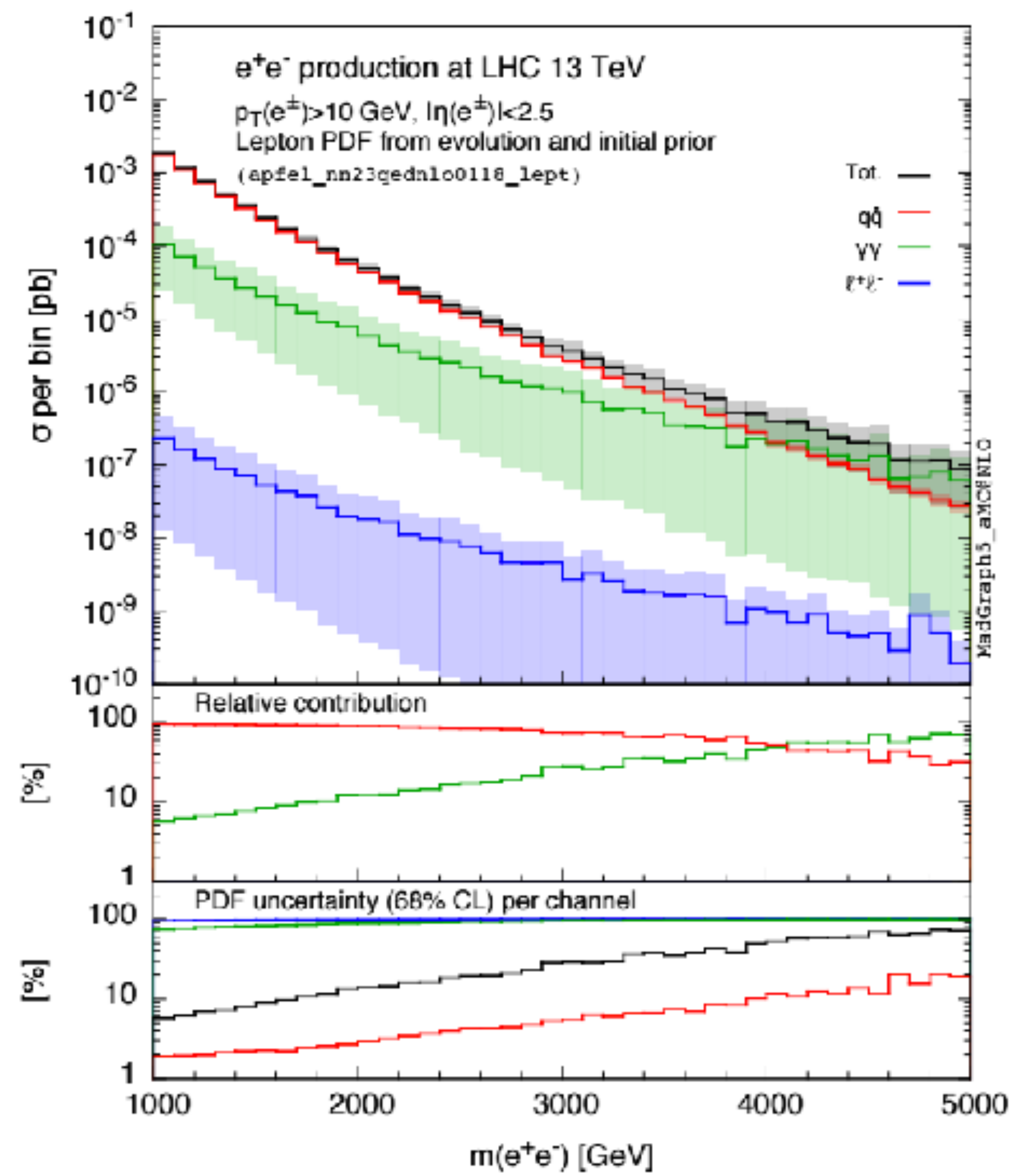
- Threshold-resummed PDFs made recently available [Bonvini et al, JHEP 1509 (2015) 191]
- Gluon suppressed as compared to fixed-order PDFs mostly due to enhancement of NLO+NLL xsecs used in the fit of DIS structure functions and DY distributions
- This suppression partially or totally compensates enhancements in partonic cross sections. Phenomenologically relevant for new physics processes [Beenakker et al. EPJC76 (2016)2, 53]
- Work in progress on small-x, pT resummation, PS resummation

EW corrections

- EW corrections become relevant at the current precision level as are sizeable at large invariant mass
- Full inclusion of EW corrections requires initial γ PDF, which we thought induced large uncertainty



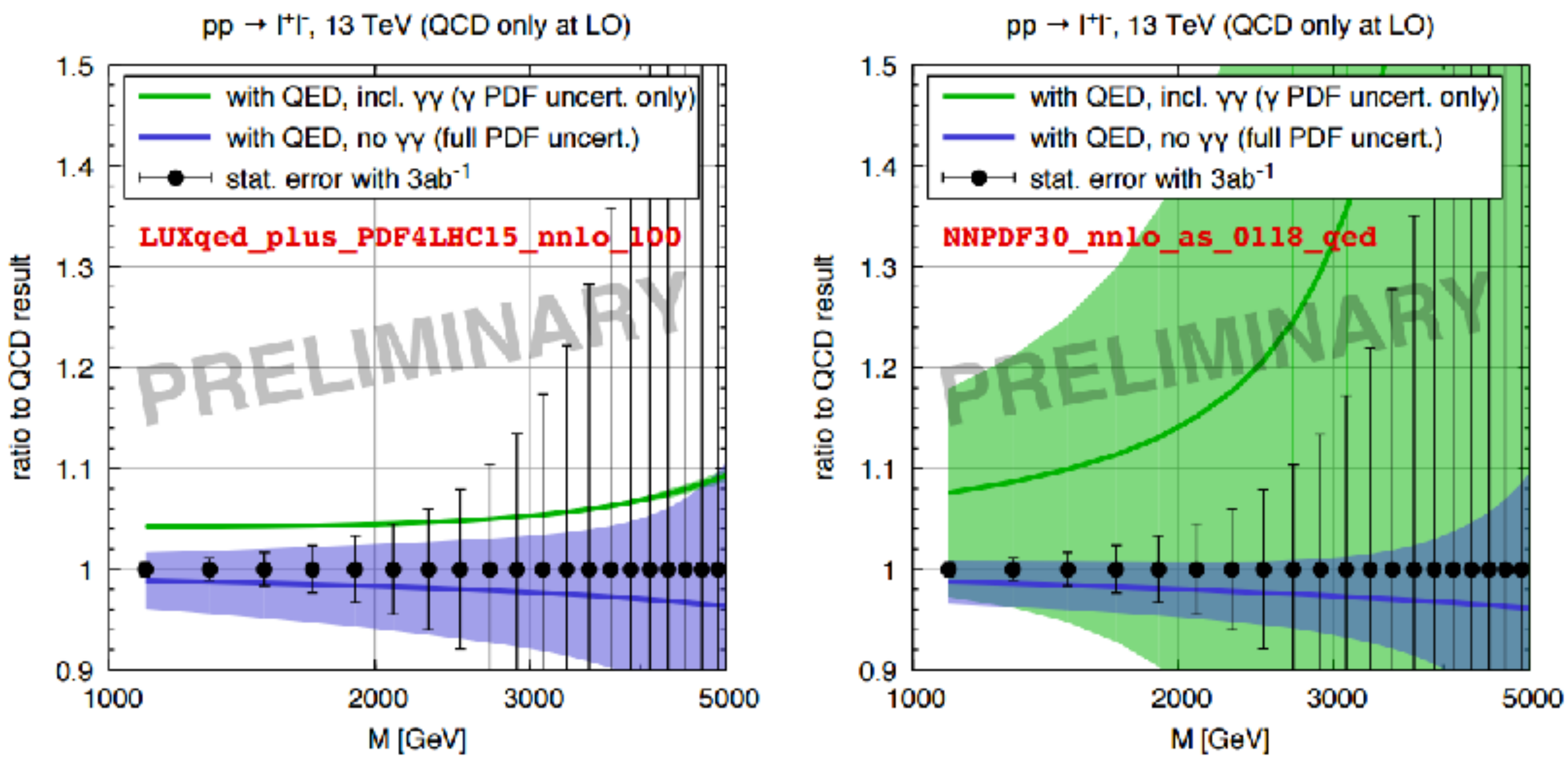
Boughezal et al [Phys.Rev. D89 (2014)3, 034030]



Bertone et al [JHEP 1511 (2015) 194]

Photon PDF

- Data-driven NNPDF approach inducing a large uncertainty on photon PDF
- Breakthrough: LUX PDF [Manohar, Nason, Salam, Zanderighi,1607.04266]
- Take a BSM interaction, compute the cross section with the Master Formula or with the Parton Model formula. Extract photon PDF by identifying the two cross sections.
- Theory constraint reduces uncertainty by a huge factor
- NNPDF3.1QED: include LUX constraint in a PDF fit (as for momentum sum rules)



Conclusions

- Parton Distribution Functions essential ingredient for LHC phenomenology
- Accurate PDFs are required for precision SM measurements
- NNPDF3.1 includes many new precise data from HERA combination to Tevatron legacy data to new LHC data (some never fitted before such as Z pT and top differential distributions)
- Good stability with respect to 3.0, reduced uncertainty in the gluon and better quark-flavour separation
- Precision of the data and correlation-dominated uncertainties very challenging for PDF fitters: is an additional uncorrelated uncertainty the way forward?
- Fitted charm improves the quality of the fit, both perturbative charm set and fitter charm sets will be released

Outlook

Exploit precise **LHC data** to reduce PDF uncertainties

Experimental **correlations** bound to be dominant errors

EXPER.

Reduce **theoretical uncertainty** in PDF fits: resummation, EW effects, HQ masses, intrinsic HQ, parton shower

Introduce a way to measure residual **theoretical uncertainty** in PDF fits

THEORY

The higher the energy regime, the more theory boundaries are probed
The smaller the experimental uncertainty, the more crucial is theory uncertainty