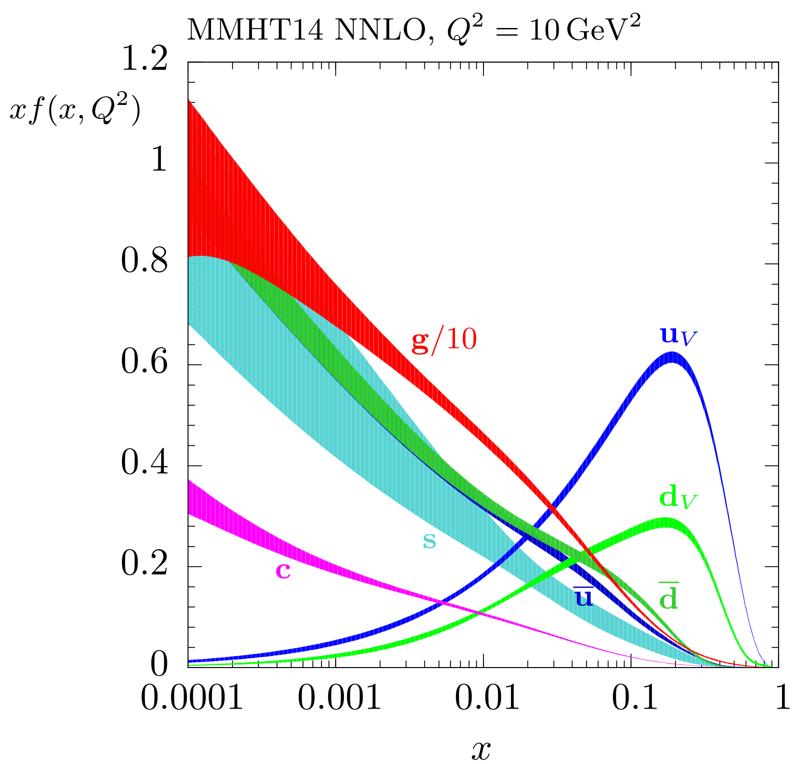
The MMHT View of the Proton

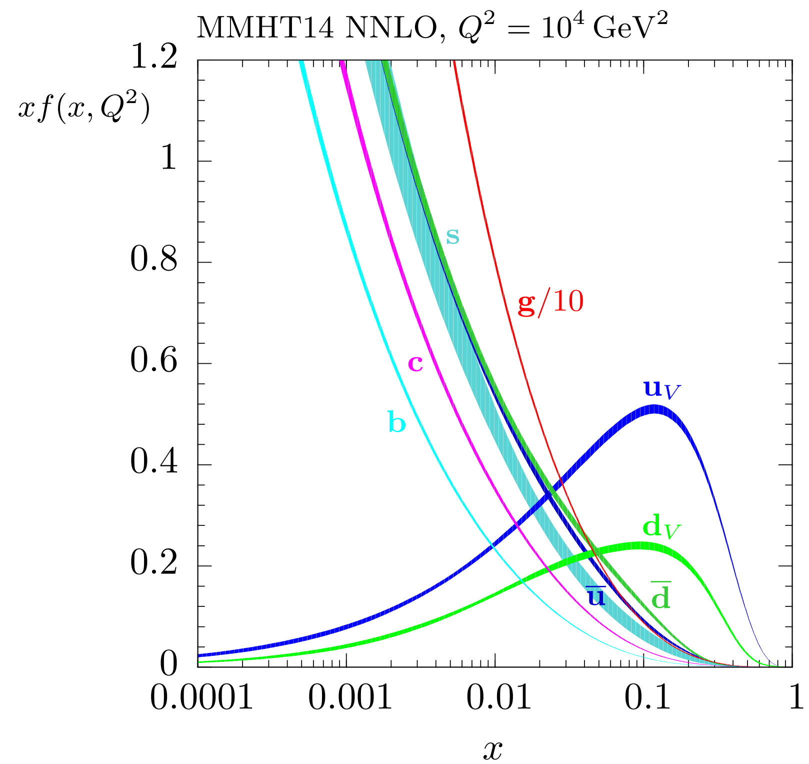
Lucian Harland-Lang, University of Oxford

Birmingham University Particle Physics Seminar
25 April 2018

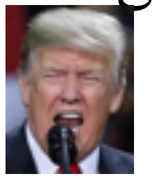
Outline

- Introduction - what are **PDFs** and why are they important?
- Role of precision **LHC** data in PDF fits. Two examples:
    - **Vector boson** production and the proton strangeness.
    - **Jet** production at NNLO.
- **New calculations** - the photon PDF.
- **Ongoing work** - MMHT18 and Ultimate PDFs.





Introductory Remarks

- The extraction of Parton Distribution Functions (PDFs) is a huge subject - could spend entire seminar (many slides, many slides ) on one specific sub-topic.
- Here I will give an overview and pick out a few interesting developments and questions/issues relevant to the high precision LHC. For more details/ broader overview:

The Structure of the Proton in the LHC Precision Era

Jun Gao[a], Lucian Harland-Lang[b], Juan Rojo[c,d]

[a]*Institute of Nuclear and Particle Physics,*
Shanghai Key Laboratory for Particle Physics and Cosmology,
School of Physics and Astronomy, Shanghai Jiao Tong University, Shanghai, China
[b]*Department of Physics and Astronomy, University College London, WC1E 6BT, United Kingdom*
[c]*Department of Physics and Astronomy, VU University, De Boelelaan 1081, 1081HV Amsterdam, The Netherlands*
[d]*Nikhef, Science Park 105, NL-1098 XG Amsterdam, The Netherlands*

Abstract

We review recent progress in the determination of the parton distribution functions (PDFs) of the proton,

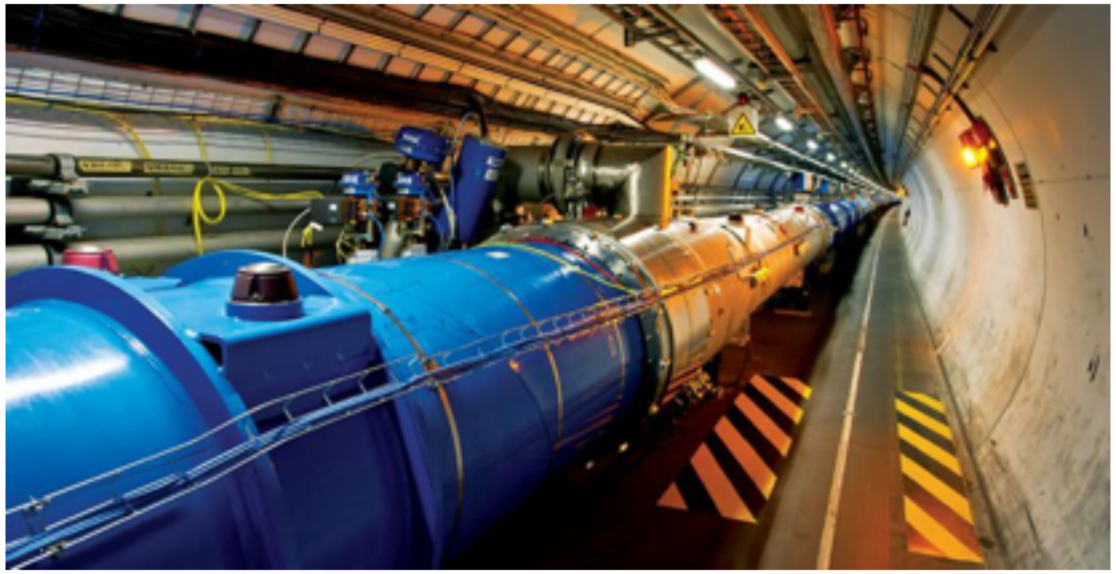
arXiv:1709.04922 (Published in Physics Reports).

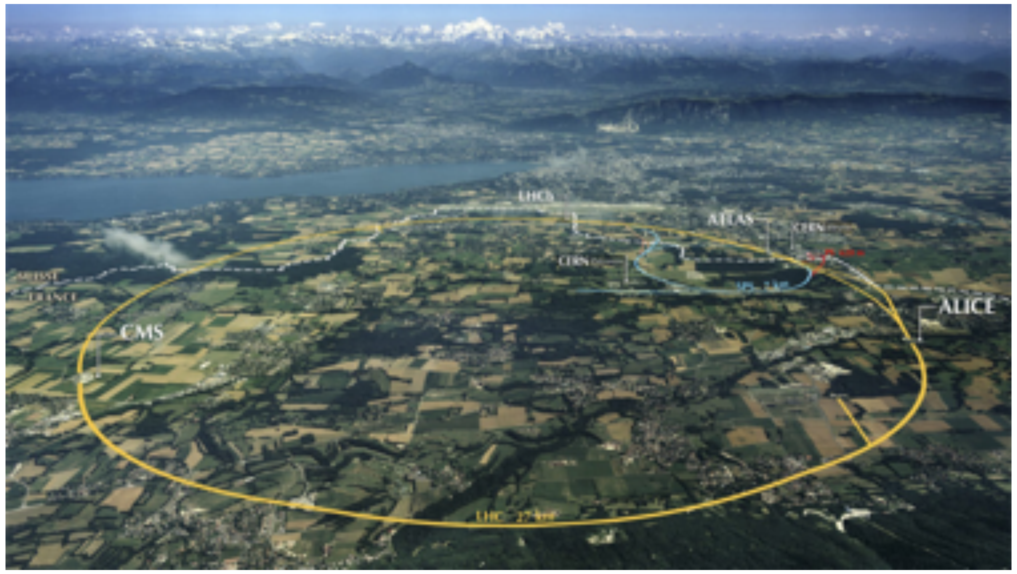
- But first of all, what are PDFs and why do we care?

Parton Distribution Functions

The LHC

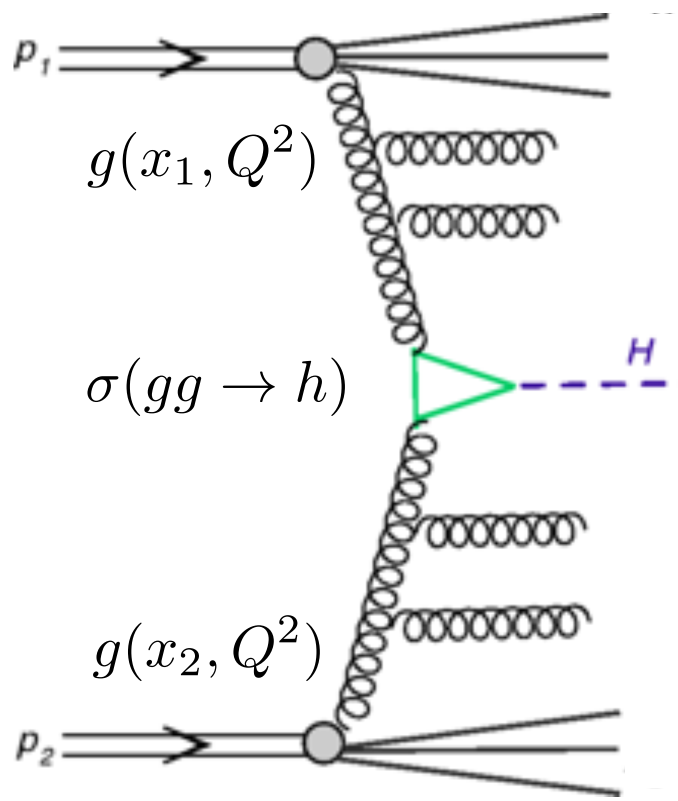
- The **Large Hadron Collider** (LHC) is the most powerful accelerator ever built, with unique sensitivity to the Higgs sector and physics within and beyond the Standard Model.
- It is also (predominantly) a **proton-proton** collider.

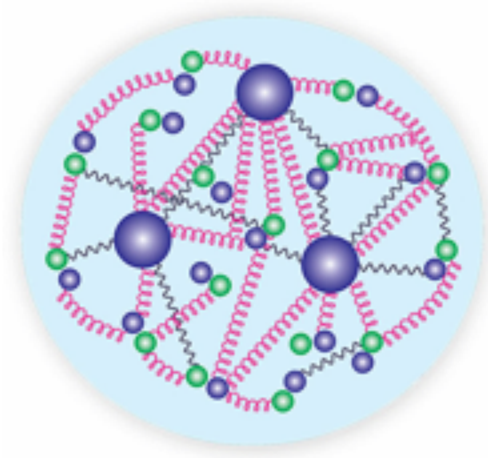




An LHC collision

- How do we model an LHC collision? Proton is composite - collision involves quarks/gluons:





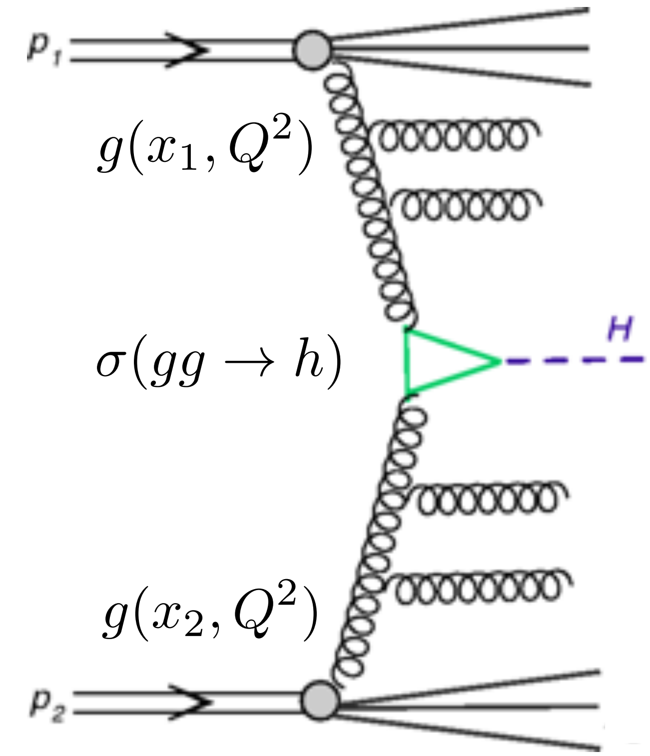
- The `**parton model**' - proton-proton cross section is convolution of **parton-level cross section** and **Parton Distribution Functions** (PDFs)

$$\sigma(pp\to h+X)\sim\sigma(gg\to h)\otimes g(x_1,Q^2)\otimes g(x_2,Q^2)\ ,$$

$$f(x)\otimes g(x)\sim\int \mathrm{d}y\, f(x)g(x/y)\ ,$$

Parton Distribution Functions

$$\sigma(pp \to h + X) \sim \sigma(gg \to h) \otimes g(x_1, Q^2) \otimes g(x_2, Q^2) \ ,$$



- Cross section given in terms of:

$\sigma(gg \to h)$: parton-level cross section. $\alpha_S(m_h) \ll 1 \Rightarrow$ perturbative expansion in α_S :

$$\sigma(gg \to h) = \alpha_S(m_h)^2(\sigma_0 + \alpha_S(m_h)\sigma_1 + \cdots)$$

$g(x, Q^2)$: PDF for gluon

x - proton longitudinal **momentum fraction**.

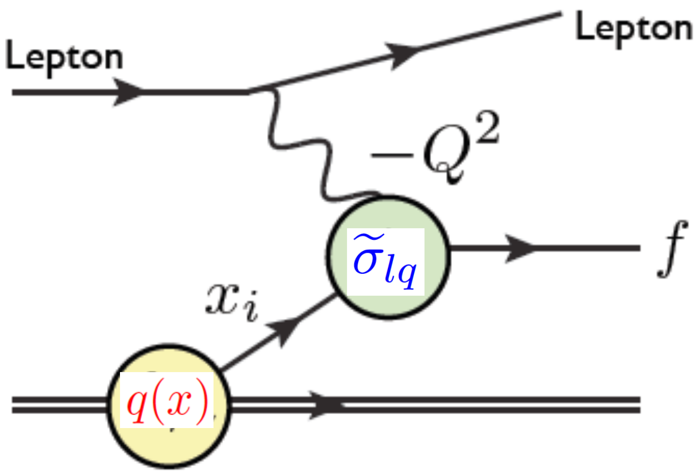
Q - **factorization scale** ~ energy of quark/gluon collision ~ inverse of resolution length.

- At lowest order PDF is probability of finding gluon in the proton carrying momentum fraction x .

DGLAP

- Quark/gluons like to radiate $\Rightarrow$ PDFs depend on resolution scale.

Formally, **factorization** in QCD requires introduction of a scale μ_F



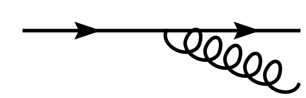
$$\sigma^{lp} \sim \sigma^{lq}(\mu_F) \otimes q(x, \mu_F)$$

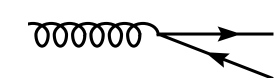
μ_F : factorization scale

- Requiring that physical cross section is independent of this to calculated order in α_S gives **DGLAP** evolution equation, e.g.

Dokshitzer-Gribov-Lipatov-Altarelli-Parisi

$$\frac{\mathrm{d}\sigma^{lp}}{\mathrm{d}\mu_F} = 0 + \text{higher orders} \rightarrow \frac{\partial q(x,\mu)}{\partial \mu} = P_{qq} \otimes q(x,\mu) + P_{qg} \otimes g(x,\mu)$$





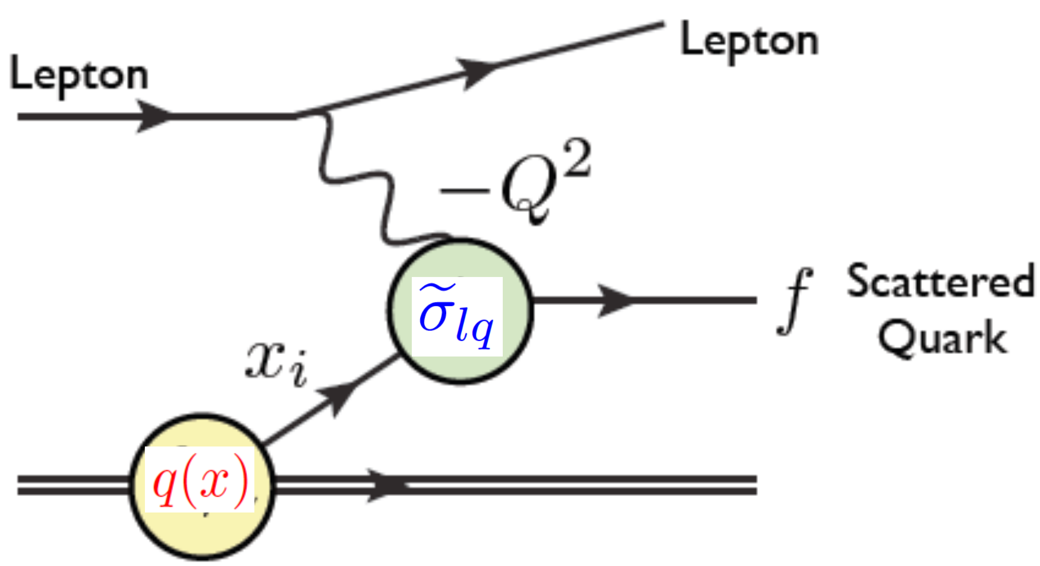
- **DGLAP** $\Rightarrow$ PDFs at lower scale determine PDFs at higher scales.

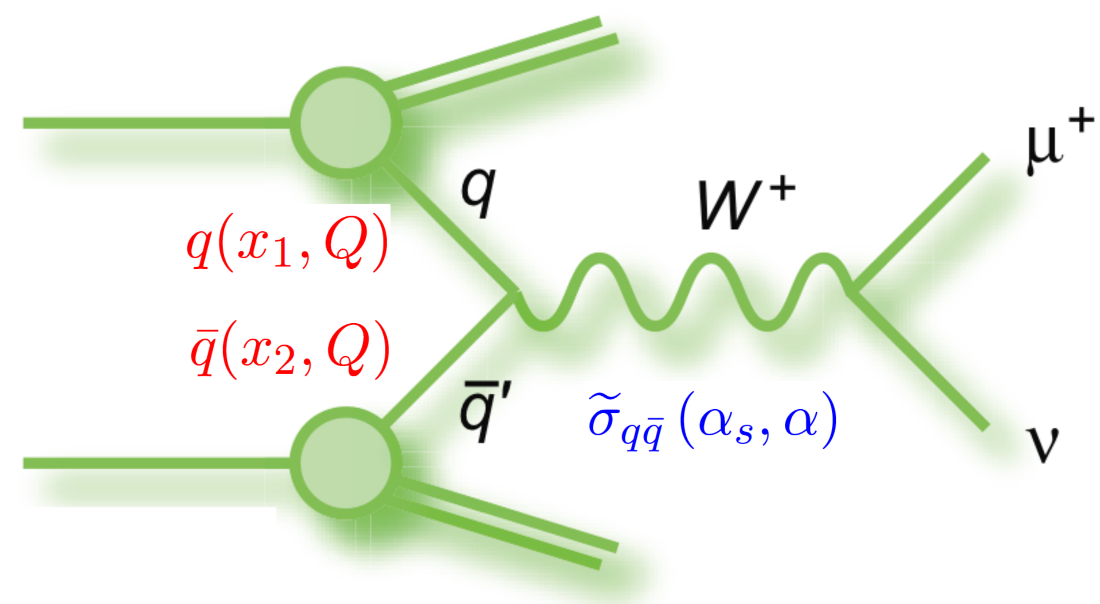
Thus fits parameterise at low scale Q_0 and fit to a range of energies.

Extracting PDFs

- QCD binding of quarks/gluons in the proton occurs at scale $\sim \Lambda_{\mathrm{QCD}} \Rightarrow$ cannot calculate using perturbative QCD.
- However **factorization** $\Longrightarrow$ PDFs are universal, e.g. for **Deep Inelastic Scattering** (DIS) and **Drell-Yan** (DY) production:

$$\sigma_{lp} \simeq \widetilde{\sigma}_{lq}\,(\alpha_s,\alpha) \otimes q(x,Q) \qquad \sigma_{pp} \simeq \widetilde{\sigma}_{q\bar{q}}\,(\alpha_s,\alpha) \otimes q(x_1,Q) \otimes \bar{q}(x_2,Q)$$





$$\text{Factorization} \Rightarrow q_{DIS}(x,Q^2) \equiv q_{DY}(x,Q^2)$$

$\longrightarrow$ **Fit** the PDFs to one dataset (DIS) to make predictions for another (DY).

Global fits - MMHT

- For LHC (and elsewhere) aim to constrain PDFs to high precision for all flavours (q, $\bar{q}$, g ...) over a wide x region.

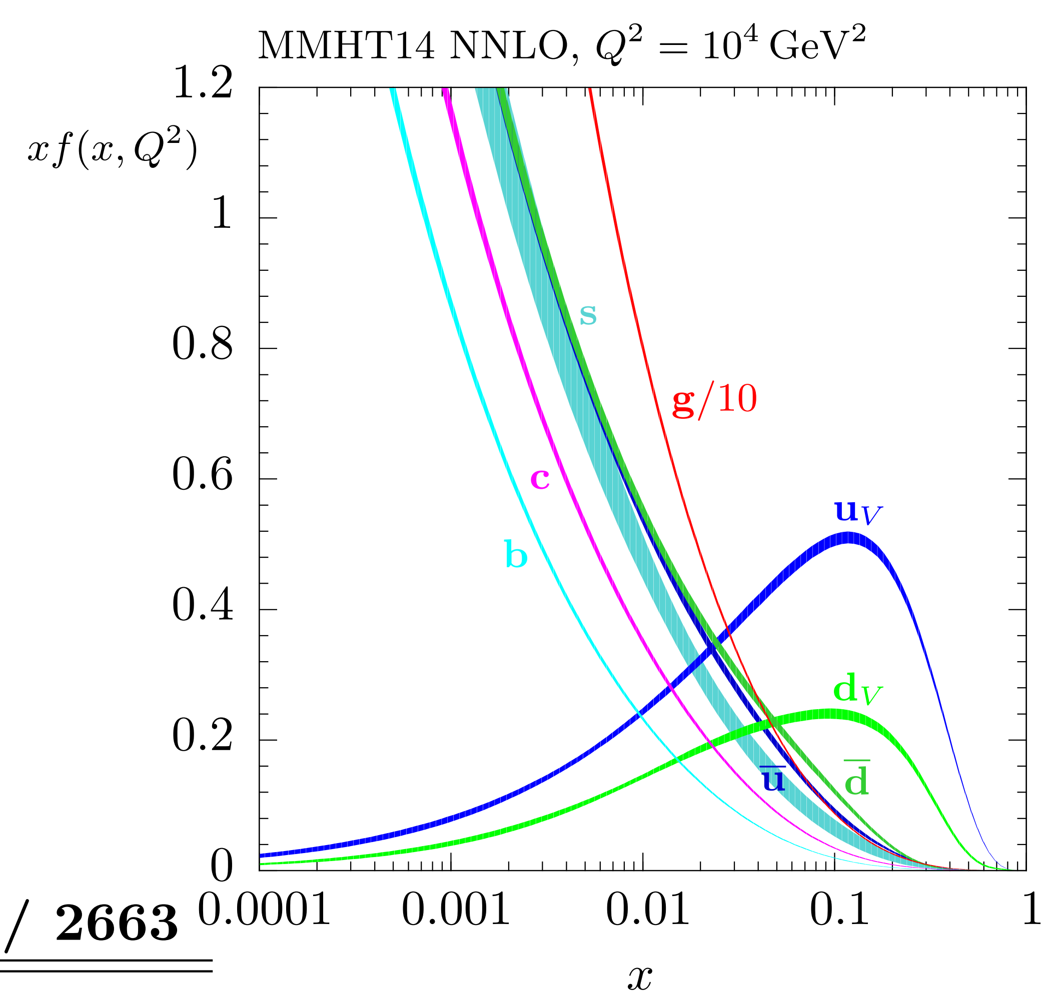
- Only so much can be done with DIS $\Rightarrow$ **MMHT** collaboration performs **global PDF fits** to wide range of data (DIS, fixed nuclear targets with l, ν beams, hadron collider data - jets, W, Z, $t\bar{t}$ ...).

- One of three major global fitters (**CT, MMHT, NNPDF**).

$\chi^2/\text{dof} \sim 1$

$\Rightarrow$ **Non-trivial check of QCD.**

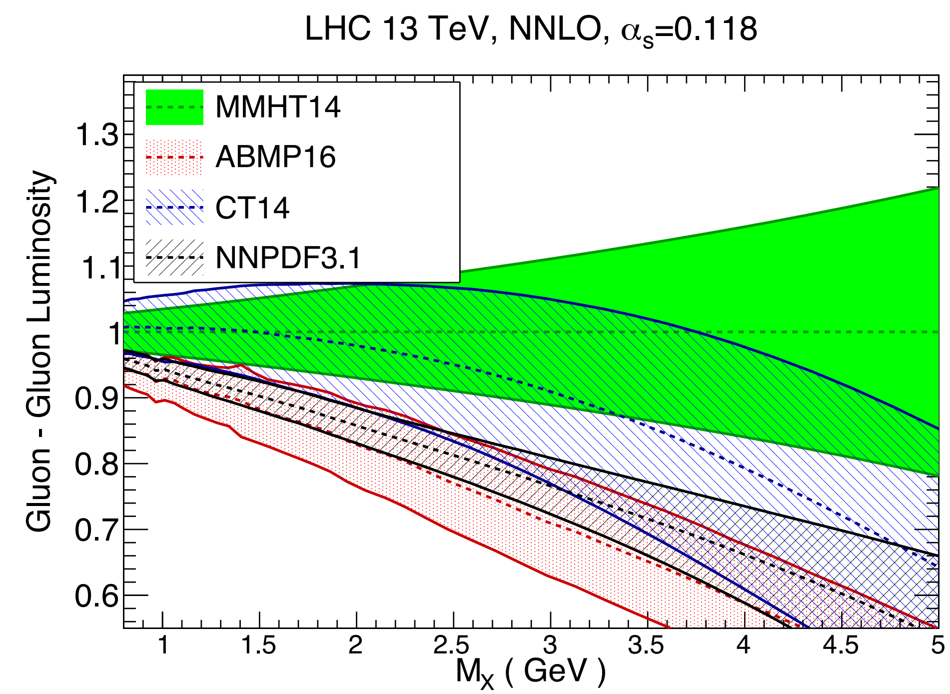
Data set	LO	NLO	NNLO
BCDMS μp F_2 [125]	162 / 153	176 / 163	173 / 163
BCDMS μd F_2 [19]	140 / 142	143 / 151	143 / 151
NMC μp F_2 [20]	141 / 115	132 / 123	123 / 123
NMC μd F_2 [20]	134 / 115	115 / 123	108 / 123
NMC $\mu n/\mu p$ [21]	122 / 137	131 / 148	127 / 148
E665 μp F_2 [22]	59 / 53	60 / 53	65 / 53
E665 μd F_2 [22]	52 / 53	52 / 53	60 / 53
SLAC ep F_2 [23, 24]	21 / 18	31 / 37	31 / 37
SLAC ed F_2 [23, 24]	13 / 18	30 / 38	26 / 38
NMC/BCDMS/SLAC/HERA F_L [20, 125, 24, 63, 64, 65]	113 / 53	68 / 57	63 / 57
E866/NuSea pp DY [88]	229 / 184	221 / 184	227 / 184
E866/NuSea pd/pp DY [89]	29 / 15	11 / 15	11 / 15
NuTeV νN F_2 [29]	35 / 49	39 / 53	38 / 53
CHORUS νN F_2 [30]	25 / 37	26 / 42	28 / 42
NuTeV νN xF_3 [29]	49 / 42	37 / 42	31 / 42
CHORUS νN xF_3 [30]	35 / 28	22 / 28	19 / 28
CCFR $\nu N \to \mu\mu X$ [31]	65 / 86	71 / 86	76 / 86
NuTeV $\nu N \to \mu\mu X$ [31]	53 / 40	38 / 40	43 / 40
HERA e^+p NC 820 GeV[61]	125 / 78	93 / 78	89 / 78
HERA e^+p NC 920 GeV[61]	479 /330	402 /330	373/ 330
HERA e^-p NC 920 GeV [61]	158/ 145	129/ 145	125 /145
HERA e^+p CC [61]	41 / 34	34 / 34	32 / 34
HERA e^-p CC [61]	29 / 34	23 / 34	21 / 34
HERA ep F_2^{charm} [62]	105 /52	72 / 52	82 / 52
H1 99–00 e^+p incl. jets [126]	77 / 24	14 / 24	—
ZEUS incl. jets [127, 128]	140/60	45 / 60	—
DØ II $p\bar{p}$ incl. jets [119]	125 / 110	116 / 110	119 / 110
CDF II $p\bar{p}$ incl. jets [118]	78 / 76	63 / 76	59 / 76
CDF II W asym. [66]	55 / 13	32 / 13	30 / 13
DØ II $W \to \nu e$ asym. [67]	47 / 12	28 / 12	27 / 12
DØ II $W \to \nu\mu$ asym. [68]	16 / 10	19 / 10	21 / 10
DØ II Z rap. [90]	34 / 28	16 / 28	16 / 28
CDF II Z rap. [70]	95 / 28	36 / 28	40 / 28
ATLAS W^+, W^-, Z [10]	94/30	38/30	39/30
CMS W asymm $p_T > 35$ GeV [9]	10/11	7/11	9/11
CMS asymm $p_T > 25$ GeV, 30 GeV[77]	7/24	8/24	10/24
LHCb $Z \to e^+e^-$[79]	76/9	13/9	20/9
LHCb W asymm $p_T > 20$ GeV[78]	27/10	12/10	16/10
CMS $Z \to e^+e^-$ [84]	46/35	19/35	22/35
ATLAS high-mass Drell-Yan [83]	42/13	21/13	17/13
CMS double diff. Drell-Yan [86]	—	372/132	149/132
Tevatron, ATLAS, CMS $\sigma_{t\bar{t}}$ [91]–[97]	53/13	7/13	8/13
ATLAS jets (2.76 TeV+7 TeV)[108, 107]	162/116	106/116	—
CMS jets (7 TeV) [106]	150/133	138/133	—
All data sets	**3706 / 2763**	**3267 / 2996**	**2717 / 2663**



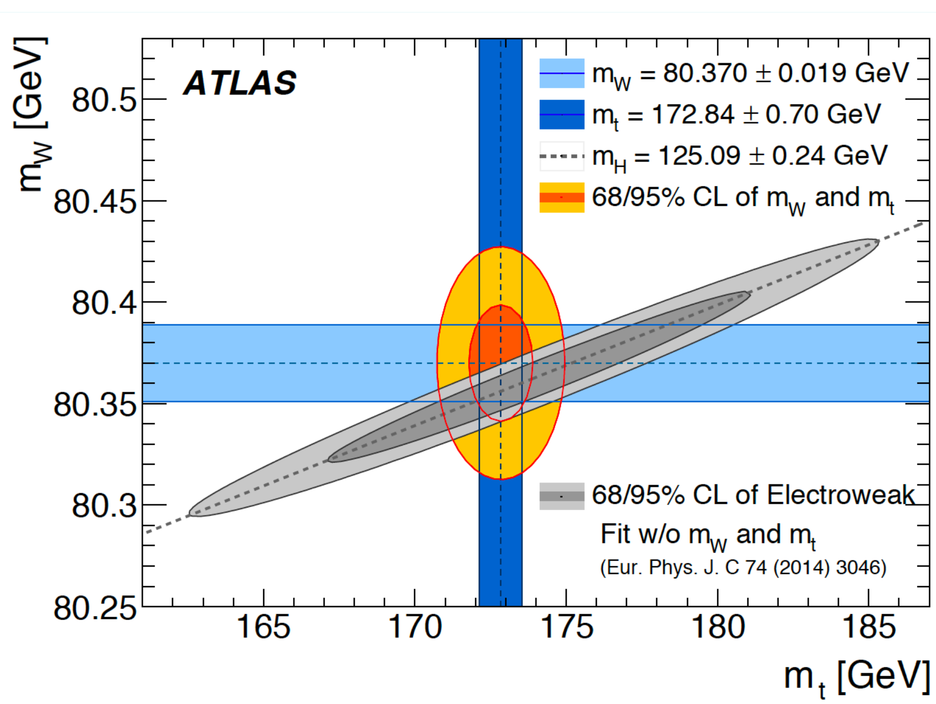
LHL et al., Eur. Phys. J. C75 (2015) no.5 204

Precise PDFs for the LHC

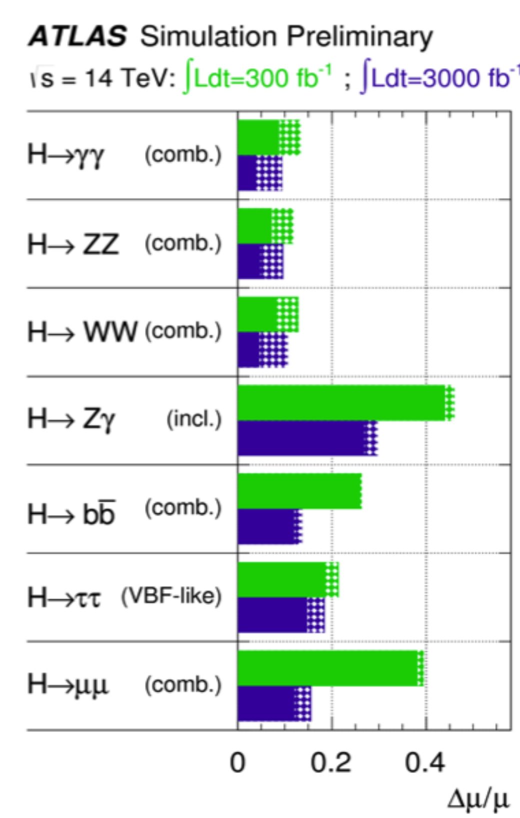
- Ultimate reach of LHC limited by knowledge of PDFs.



- **High mass searches** - PDFs in high region (currently constraints poor)



- **Higgs couplings** → need to model SM production precisely.



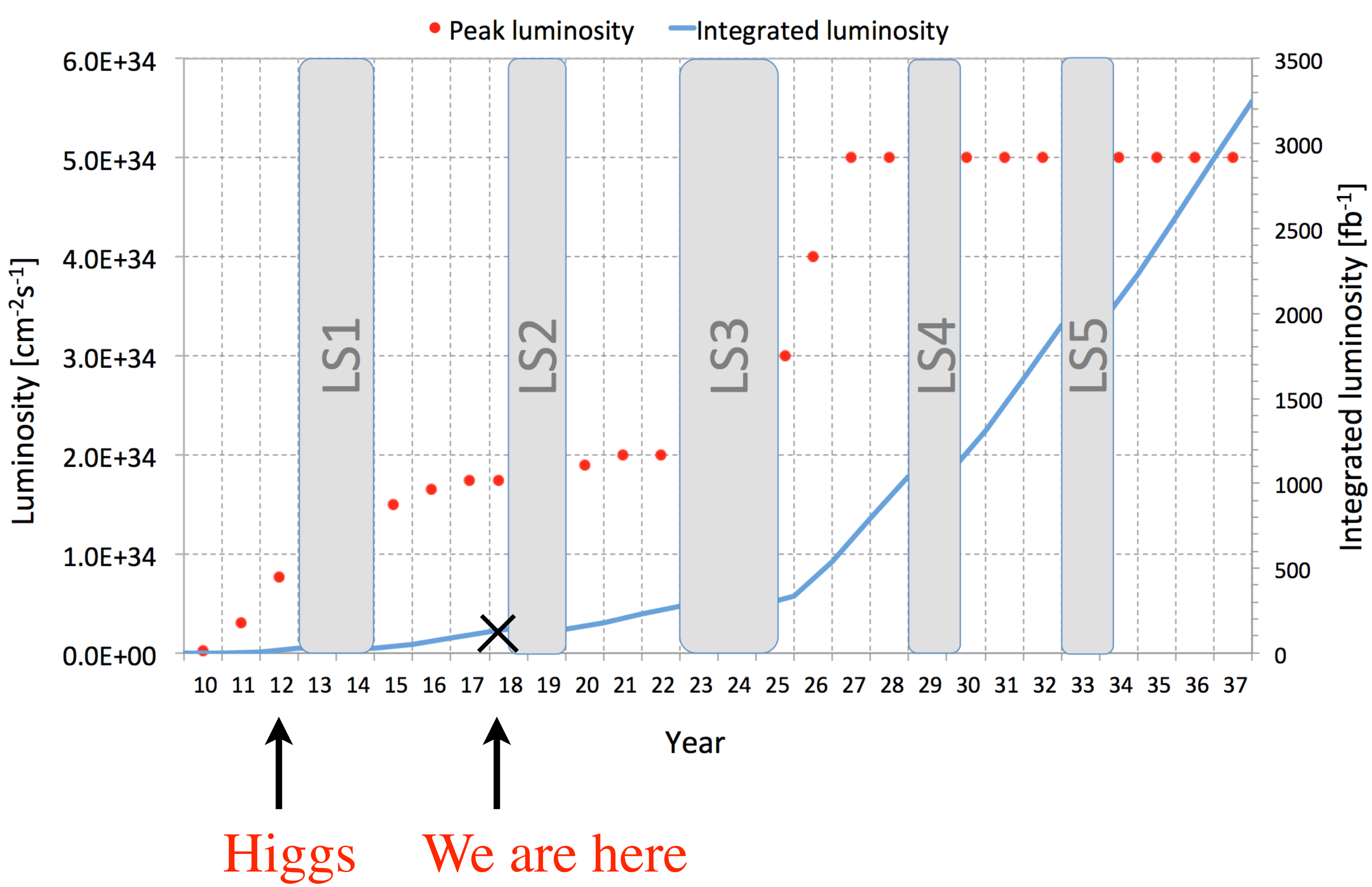
Combined categories	Value [MeV]	Stat. Unc.	Muon Unc.	Elec. Unc.	Recoil Unc.	Bckg. Unc.	QCD Unc.	EW Unc.	PDF Unc.	Total Unc.
m_T-p_T^ℓ, $W^\pm$, e-μ	80369.5	6.8	6.6	6.4	2.9	4.5	8.3	5.5	9.2	18.5

- **Precision SM** measurements - PDFs dominant uncertainty for e.g. W mass.

LHC: The Future

- We are at a very **early stage** in LHC running: so far only a few percent of the final projected data sample collected.

$\longrightarrow$ Precision requirements at the LHC rapidly increasing.



Precise Theory

- Past years has seen an explosion in calculations for LHC processes at Next-to-Next-to-Leading-Order (**NNLO**) in the strong couplings ($\sim$ % level precision).
- Thus, precision in data and theory at unprecedented level. As we will see, provides **opportunities** and **challenges** for PDF fitters.

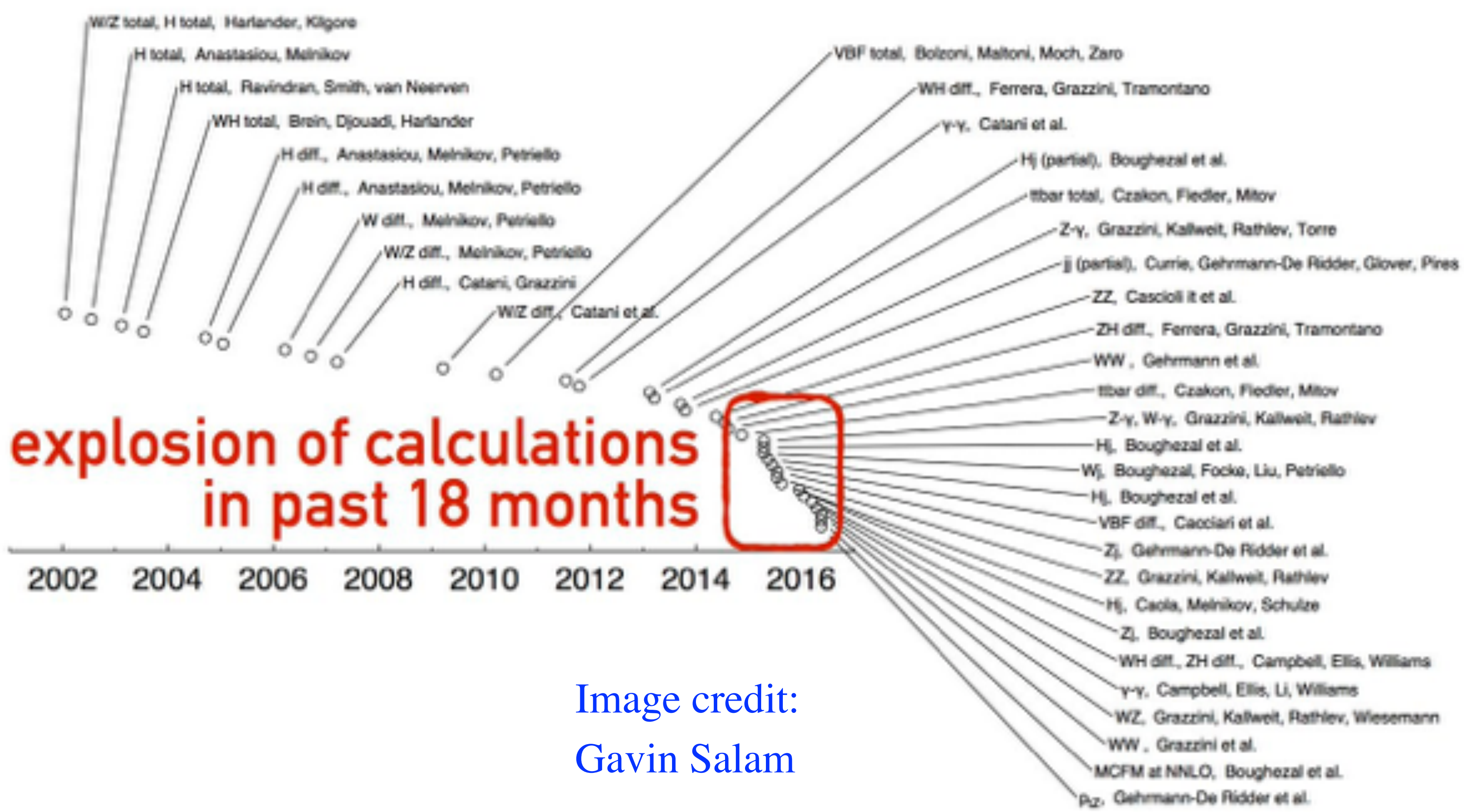


Image credit:
Gavin Salam

Confronting LHC Data

LHC Data

- Global groups busily updating fits to include the plentiful and precise new LHC data. **ABMP16, NNPDF3.1** released, **MMHT18** and **CT17** on their way.
- Many studies ongoing - I will described **two examples** in some detail here.

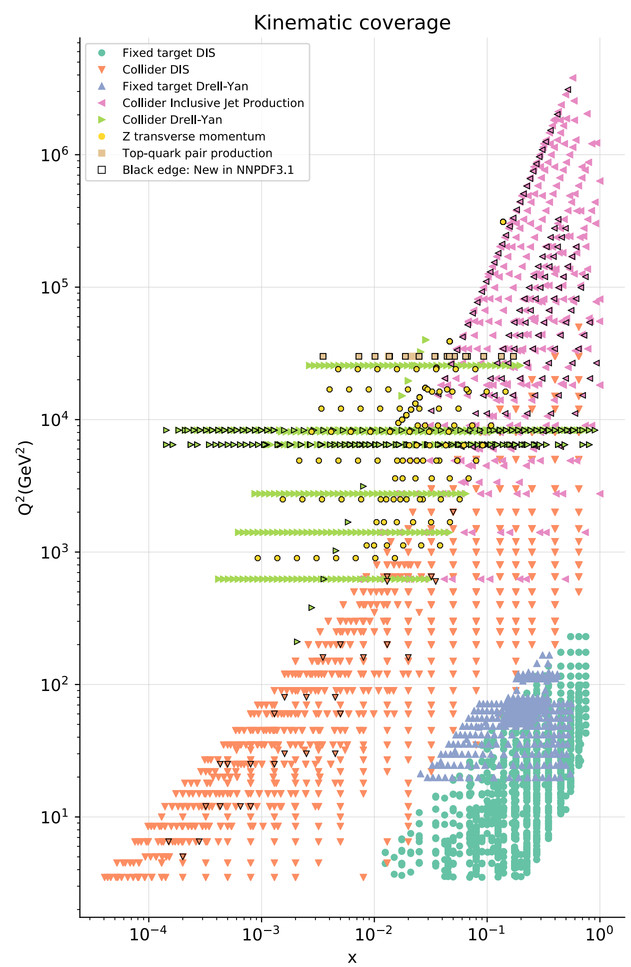
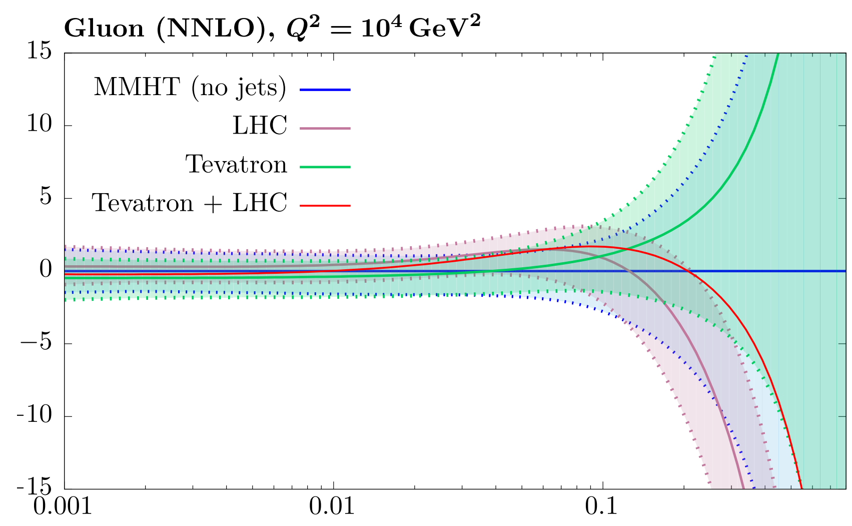
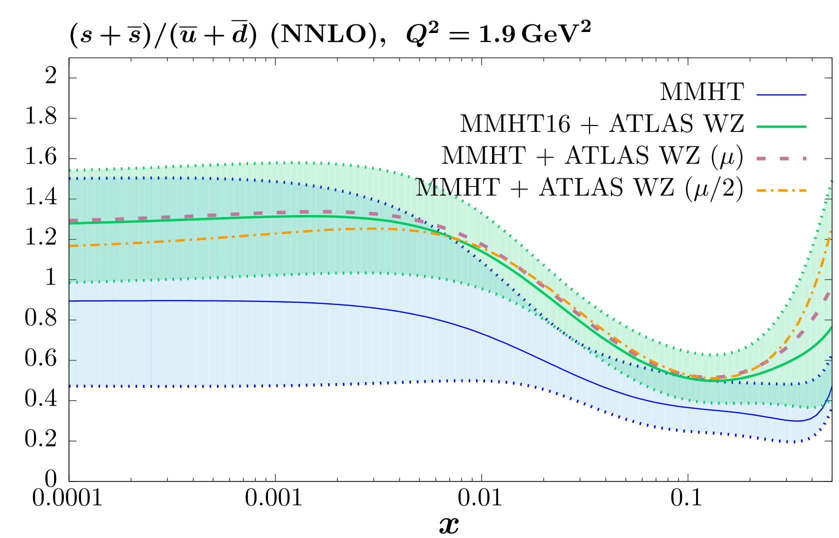


Figure 2.1: The kinematic coverage of the NNPDF3.1 dataset in the (x, Q^2) plane.



CT17p — data to be included

- Previous LHC and HERA 1 data included in CT14 will be superseded by updated Run 1 and HERA 1+2 data; adding new LHC data, especially on Z boson p_T and top quark differential distributions
  - Combined HERA1+2 DIS [1506.06042] update
  - LHCb 7 TeV Z, W muon rapidity dist. [1505.07024] update
  - LHCb 8 TeV Z rapidity dist. [1503.00963] update
  - ATLAS 7 TeV inclusive jet [1410.8857] update
  - CMS 7 TeV inclusive jet (extended y range)[1406.0324] update
  - ATLAS 7 TeV Z pT dist. [1406.3660] new
  - LHCb 13 TeV Z rapidity dist. [1607.06495] update
  - CMS 8 TeV Z pT and rapidity dist. (double diff.) [1504.03511] new
  - CMS 8 TeV W, muon asymmetry dist. [1603.01803] update
  - ATLAS 7 TeV W/Z, lepton(s) rapidity dist. [1612.03016] update
  - CMS 7,8 TeV tT differential distributions new
  - ATLAS 7,8 TeV tT differential distributions new



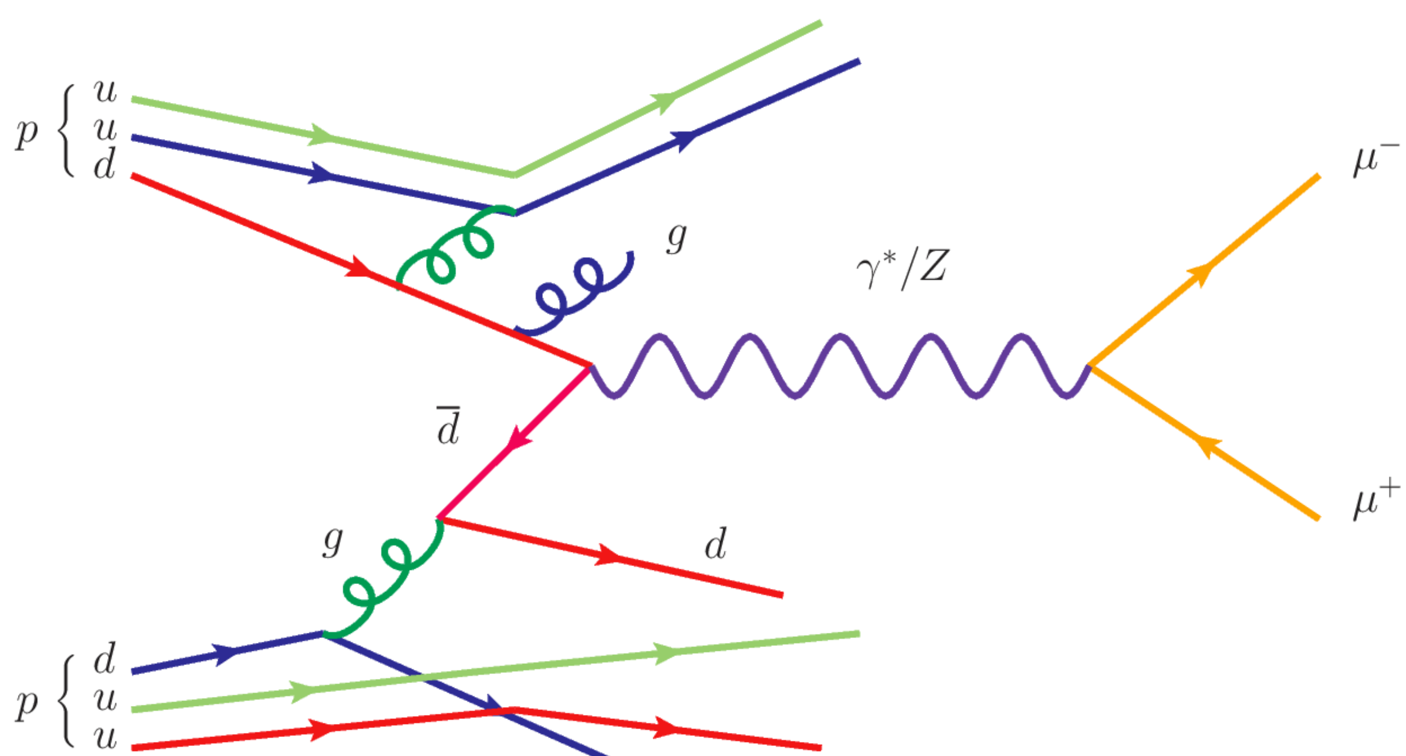
Example 1 - ATLAS W,Z and the proton strangeness

Vector Bosons and Proton Strangeness

- Vector boson (W, Z) production proceeds via range of channels:

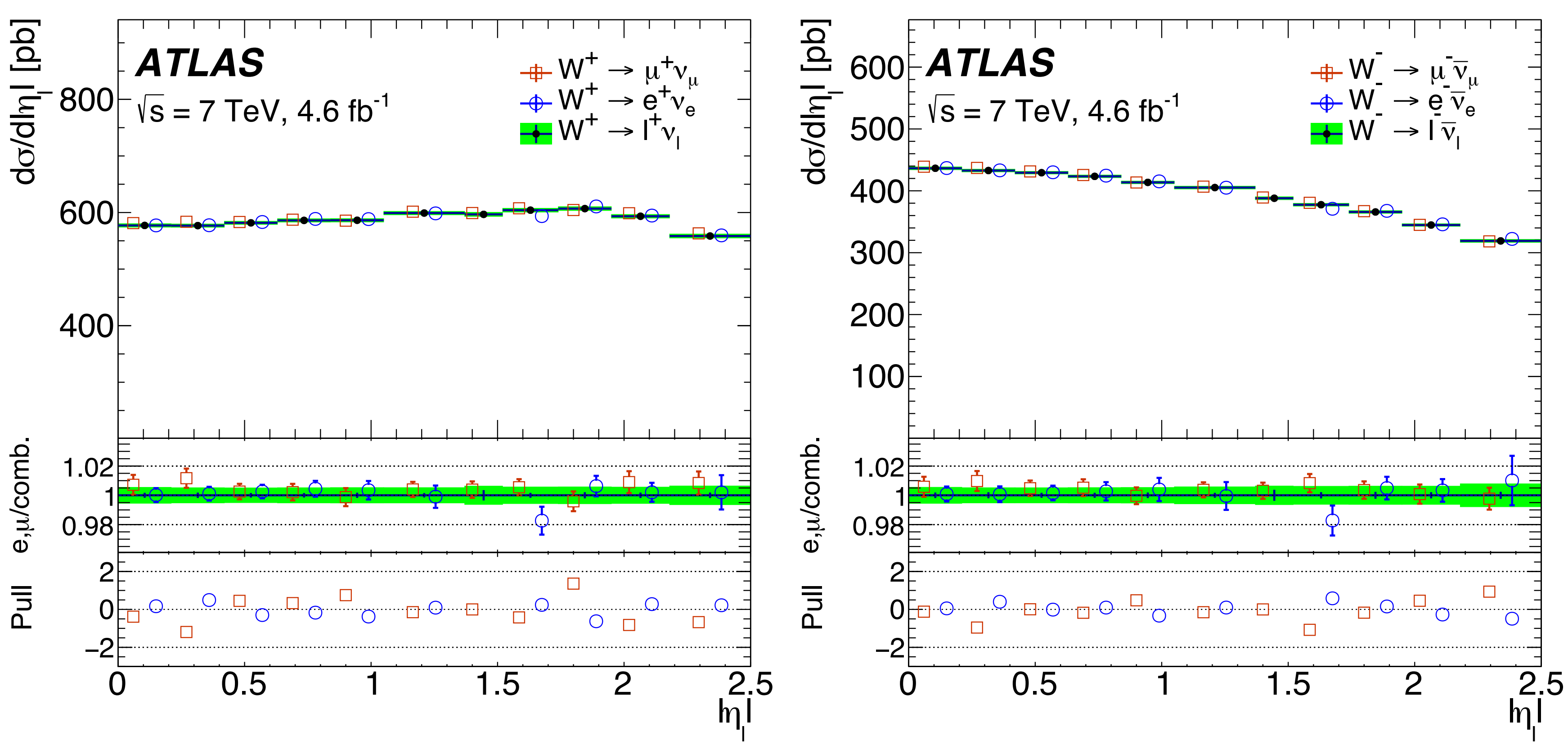
$$u\bar{d},\, c\bar{s} \quad (u\bar{s},\, c\bar{d}) \to W^+ \;,$$
$$d\bar{u},\, s\bar{c} \quad (s\bar{u},\, d\bar{c}) \to W^- \;,$$
$$q\bar{q} \to Z/\gamma^* \;,$$

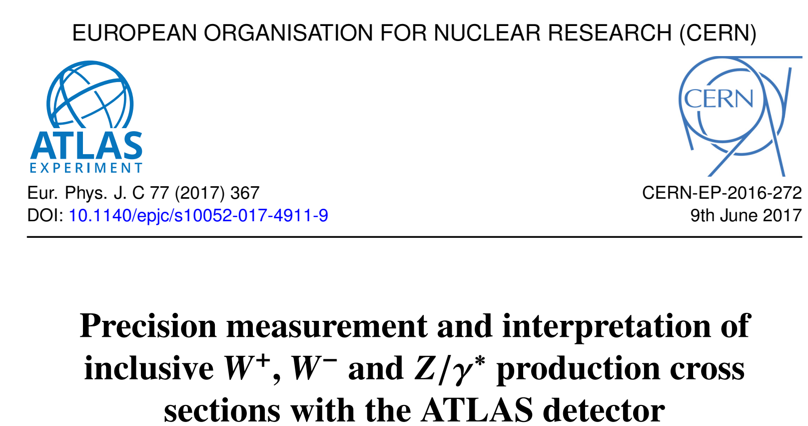
- Least constrained involves initial state s, $\bar{s}$ (no valence s) $\to$ sensitive to **proton strangeness**.
- Only in principle: small contribution, requires **precise data** to pin down.



ATLAS data

- Such data now available - highest ever precision measurement of W, Z production by the **ATLAS** collaboration at the LHC.
- Data uncertainties at the sub-% level. Statistical errors negligible completely dominated by systematics (common theme).
- Uses *7* TeV dataset taken in **2011**. Understanding these systematic errors as well as possible has taken many years.




EUROPEAN ORGANISATION FOR NUCLEAR RESEARCH (CERN)

Eur. Phys. J. C 77 (2017) 367
DOI: 10.1140/epjc/s10052-017-4911-9

CERN-EP-2016-272
9th June 2017

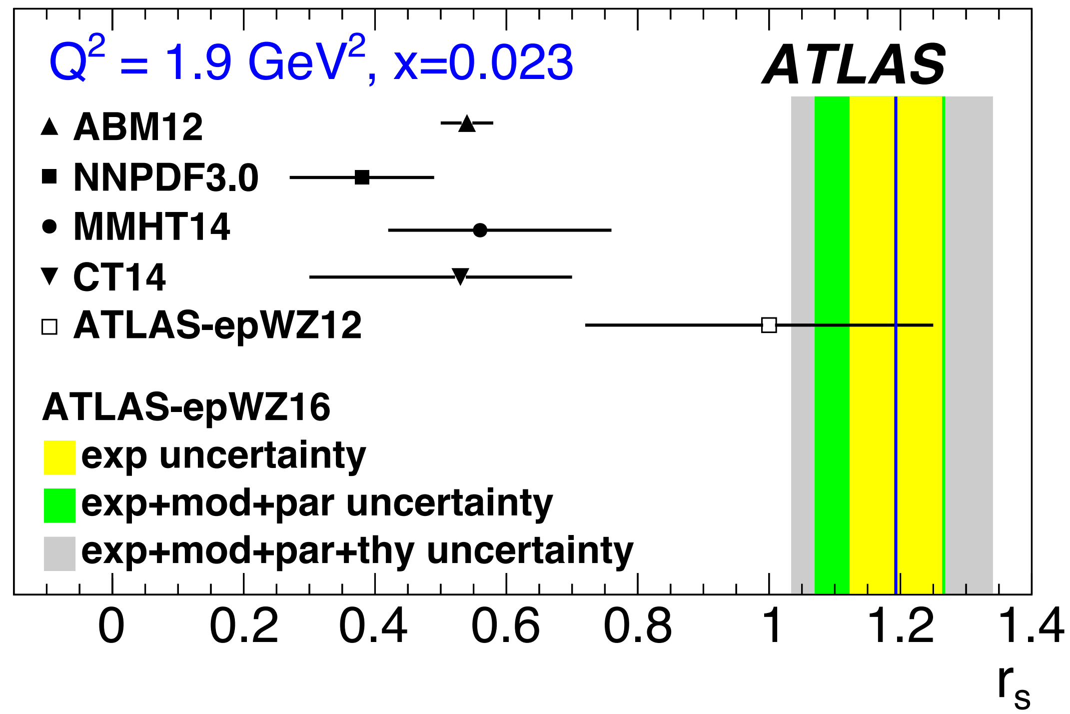
Precision measurement and interpretation of inclusive W^+, W^- and Z/γ^* production cross sections with the ATLAS detector

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

Impact on Strangeness

- ATLAS prefer a proton strangeness ratio $R_s \sim 1$, **higher** than global fits (constraints from e.g. $\overline{\nu}s \to lc$ DIS).

$$R_s = \frac{s+\bar{s}}{\bar{u}+\bar{d}}$$



- However description of the data **poor**, with:

Data set	n.d.f.	ABM12	CT14	MMHT14	NNPDF3.0	ATLAS-epWZ12
Total χ^2	61	136|222	103|290	118|396	147|351	113|159

$$\chi^2 \sim \frac{(D-T)^2}{\sigma^2}$$

$$\chi^2/\mathrm{d.o.f} \sim 2$$

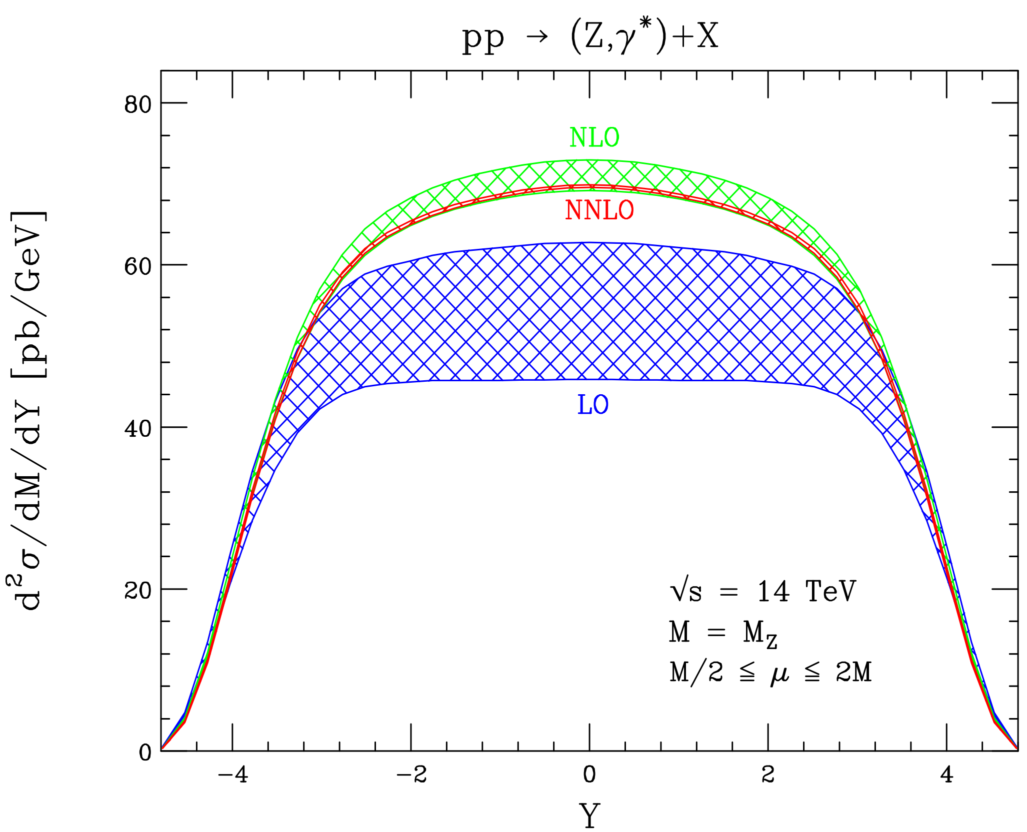
- **Challenge**: what about impact on PDFs? Can we still use these data reliably?

Perturbative Theory vs. Data

- Theory prediction is given as a **perturbative expansion** in α_S, e.g.*

$$\sigma^{q\bar{q}\to Z}(\mu_F,\mu_R) = \sigma_0 + \alpha_S(\mu_R)\sigma_1(\mu_F,\mu_R) + \alpha_S^2(\mu_R)\sigma_2(\mu_F,\mu_R) + \cdots$$

- This is truncated to a given order- precision of the result limited by **missing higher orders**. For NNLO case above ~% level.
- ATLAS data has a similar/higher level of precision $\Rightarrow$ good description not guaranteed!
- How to quantify this mis-match?



*Only showing $q\bar{q}$ channel for simplicity. Beyond LO have gq, gg...

C. Anastasiou, Phys. Rev. D69 (2004) 094008

Perturbative Theory vs. Data

- Proton-level cross section:

$$\sigma^{pp\to Z+X}(\mu_F,\mu_R) \sim \sigma(q\bar{q}\to Z)(\mu_F,\mu_R)\otimes q(x,\mu_F)\otimes \bar{q}(x,\mu_F)$$

An 'all-order' calculation (= the right answer) cannot depend on artificial scales $\Rightarrow$ dependence of σ on $\mu_{F,R}$ is at next order up

$$\sigma \sim O(\alpha_S^n) \Rightarrow \frac{\mathrm{d}\sigma}{\mathrm{d}\mu_F} \sim O(\alpha_S^{n+1})$$

$\longrightarrow$ varying μ_F, μ_R gives estimate of **uncertainties** from **higher orders**.

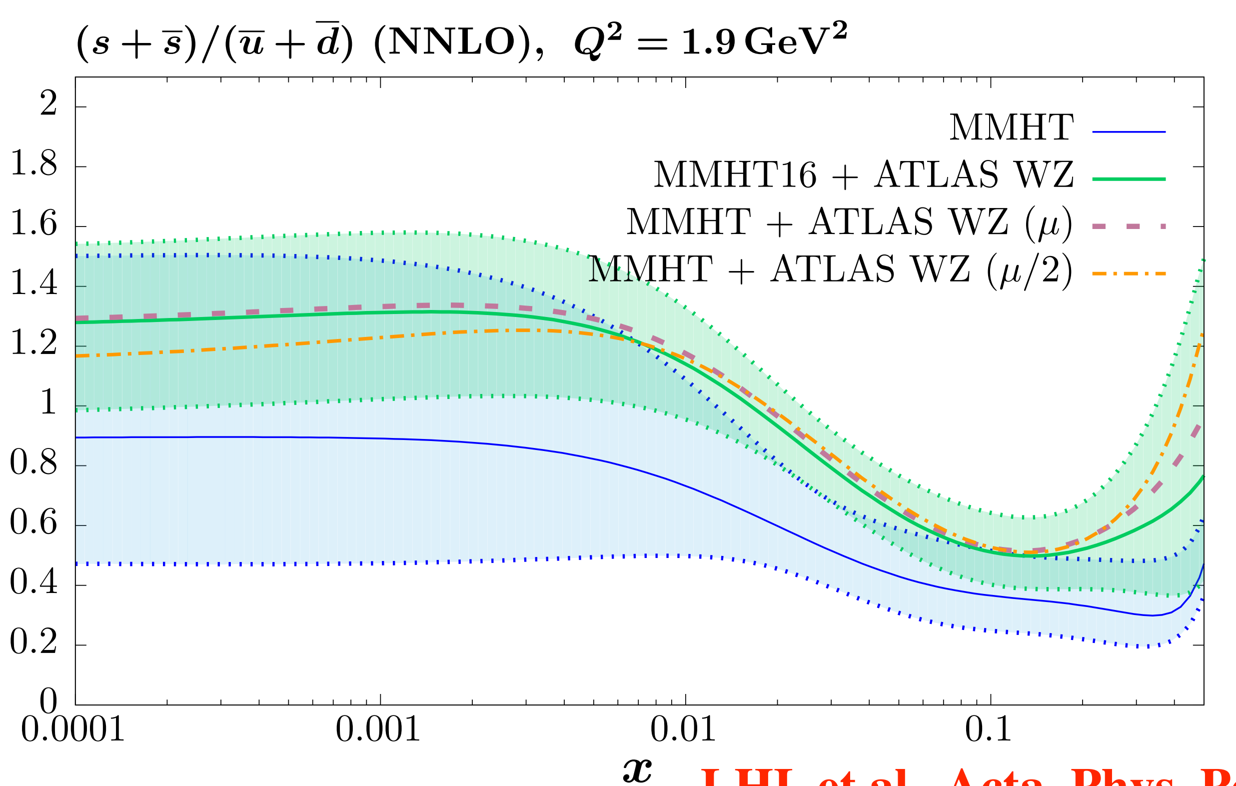
- Comparison to ATLAS data made for one default choice of $\mu_{R,F}$. But free to take others.
- Varying $\mu_{R,F}$ between $(\mu/2, 2\mu)$ ATLAS find that χ^2/dof **improves** from ~ 2 to ~1.5 per point by taking $\mu/2$.
- Should this concern us? What about **PDFs**?

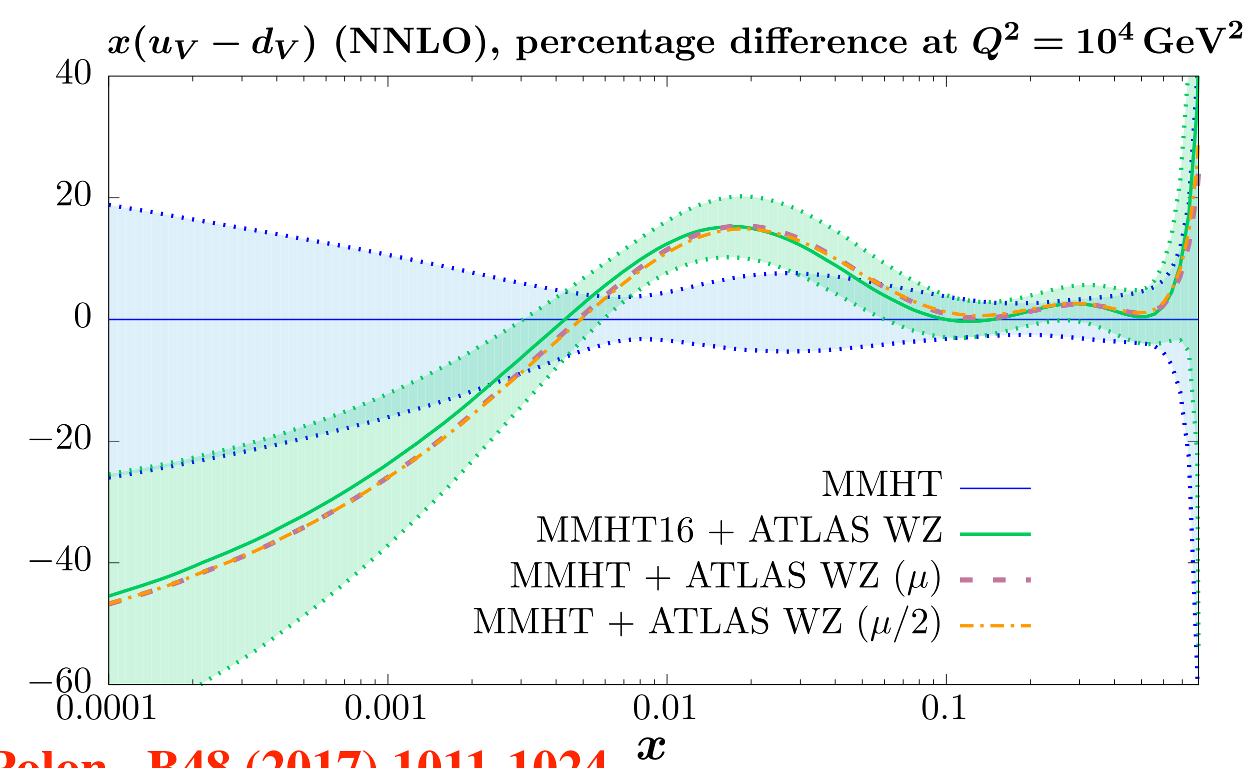
$^*\mu$ is taken as M_{ll} in the $Z/\gamma*$ case and M_W for the W

Theoretical Uncertainty

- **MMHT** study- include ATLAS data within global fit. Find higher strangeness but consistent within PDF uncertainties.
- Taking $\mu/2$ leads to $\chi^2/\mathrm{dof} \sim 2.17 \to 1.77$, definite **improvement**. However find that impact on extracted **PDFs** is **very small**.

$\longrightarrow$ Fixed order uncertainty may not be obstacle to reliable PDF determination. However, **first step** on long road: need to address question systematically (work ongoing).



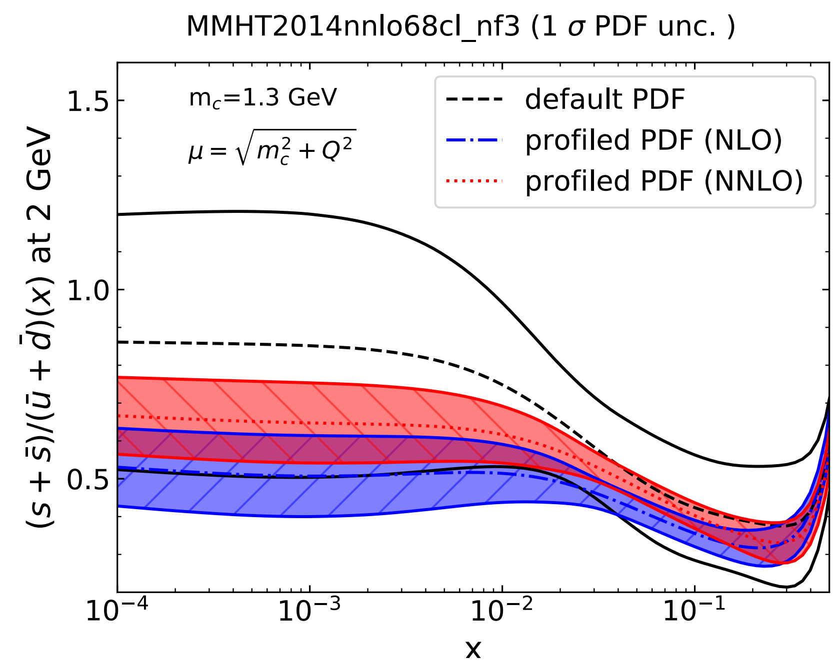


LHL et al., Acta. Phys. Polon., B48 (2017) 1011-1024

Other Effects/Futher Work

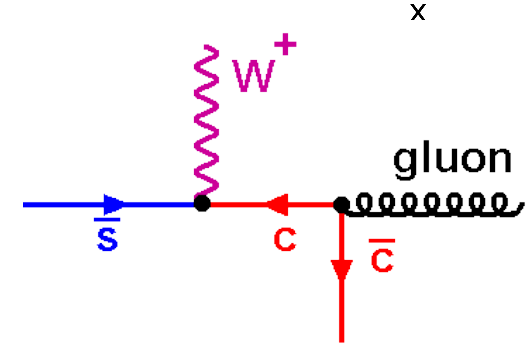
- Other open issues related to ATLAS data and proton strangeness:

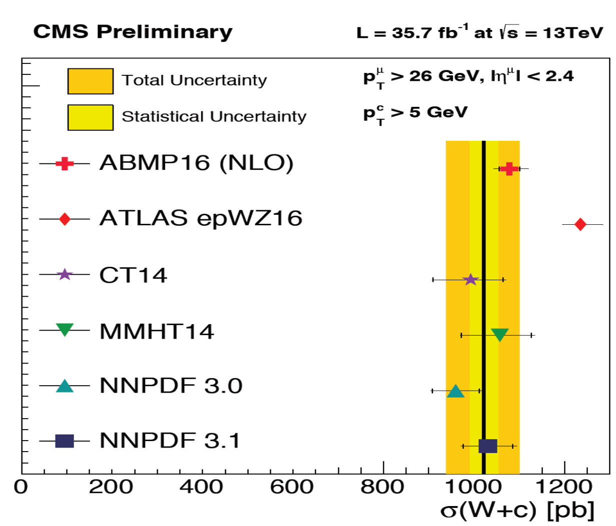
★ ATLAS data globally consistent, but pulls in different direction to ν-induced charm DIS. Recent NNLO calculation should help this.

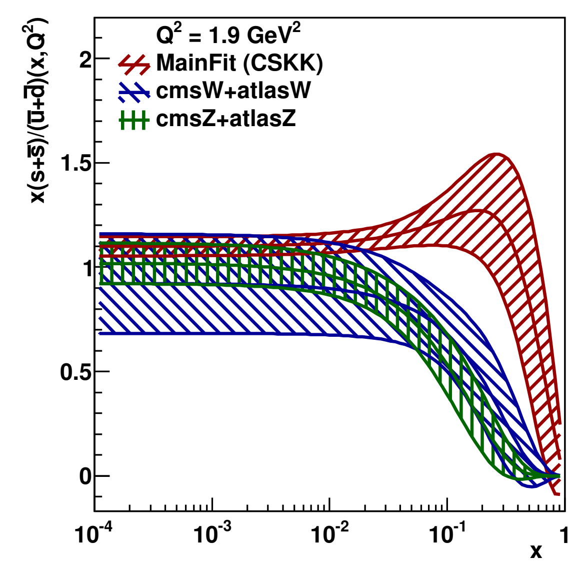


J. Gao, JHEP 1802 (2018) 026

- New combined ATLAS + CMS study of 7/8 TeV data find pull consistent, with W + Z largest (correlations $\Rightarrow$ more information).
- Excluding ATLAS Z low/high M_{ll} : little effect (but χ^2 better)
- On the other hand - CMS W + c data prefer lower strangeness.







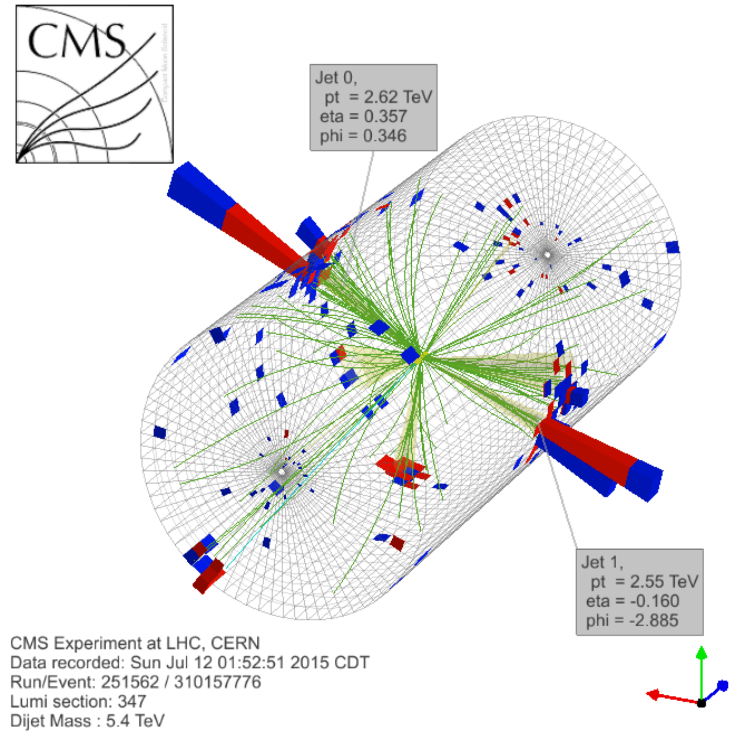
A.M. Cooper-Sarkar, K. Wichmann, arXiv: 1803.00968

More work needed!

B. Bilin, DIS18

Example 2 - LHC Jets and the Gluon PDF

Jet production and PDFs



- At the LHC, **jet** production is dominated by the **gluon-initiated** parton-level processes:

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

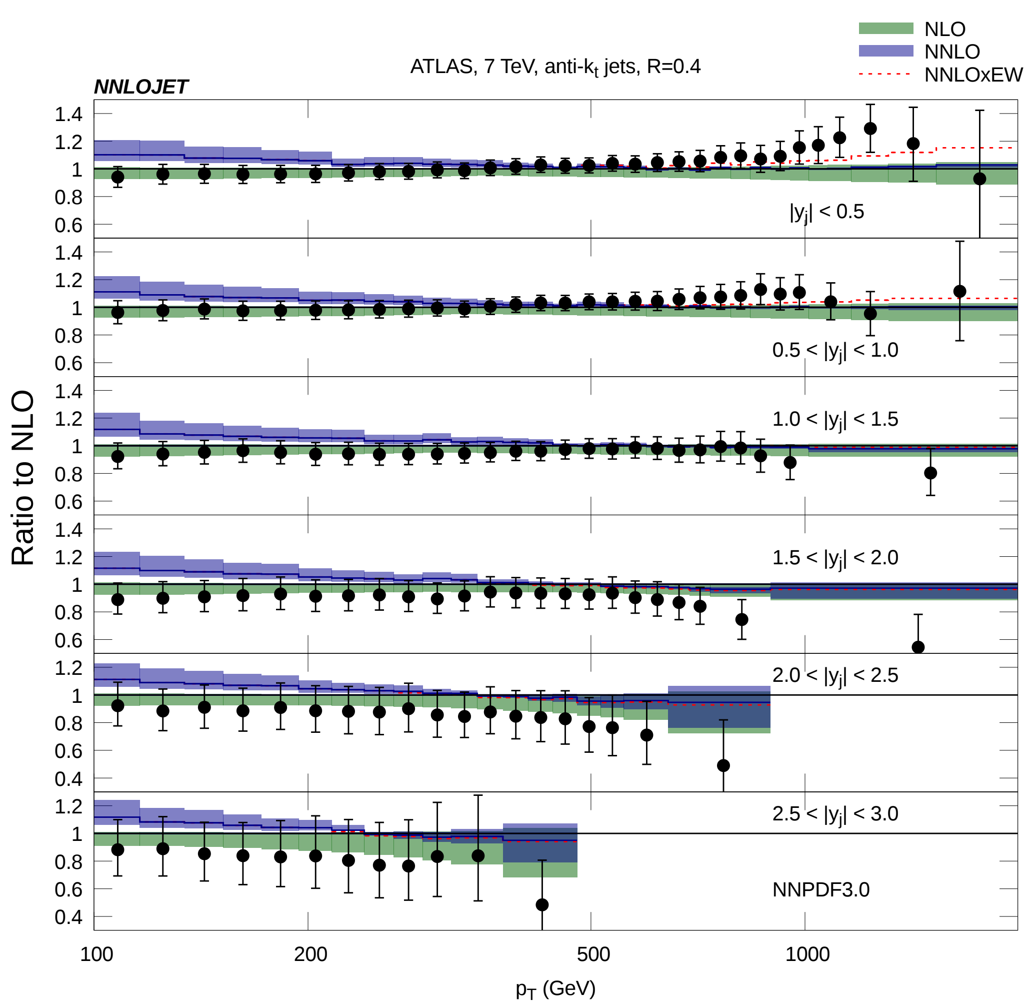
- Kinematics: $x_1 = \frac{p_T}{\sqrt{s}}(e^{y_1} + e^{y_2}),\quad x_2 = \frac{p_T}{\sqrt{s}}(e^{-y_1} + e^{-y_2})\ ,$

$\longrightarrow$ Data on jets at high transverse momenta, $p_\perp$, sensitive to **gluon PDF** at high x .

- Gluon at high x is both important for **BSM searches** and quite **poorly constrained** from DIS $\Rightarrow$ LHC data such as jet production plays crucial role in PDF determination.

NNLO jet calculation

- Full **NNLO** calculation for inclusive jet production in hadron-hadron collisions now **available**. Completion of large scale, long term project.
- Combined with availability of **high precision jet data** from ATLAS/CMS $\rightarrow$ can consider the impact on a NNLO fit for first time!



NNLO QCD predictions for single jet inclusive production at the LHC

J. Currie[a], E.W.N. Glover[a], J. Pires[b]
[a] *Institute for Particle Physics Phenomenology, University of Durham, Durham DH1 3LE, England*
[b] *Max-Planck-Institut für Physik, Föhringer Ring 6 D-80805 Munich, Germany*

We report the first calculation of fully differential jet production in all partonic channels at next-to-next-to leading order (NNLO) in perturbative QCD and compare to the available ATLAS 7 TeV data. We discuss the size and shape of the perturbative corrections along with their associated scale variation across a wide range in jet transverse momentum, p_T, and rapidity, y. We find significant effects, especially at low p_T, and discuss the possible implications for Parton Distribution Function fits.

J. Currie et al., Phys.Rev.Lett. 118 (2017) no.7, 072002

- **Recent study** on this:

arXiv:1711.05757

IPPP/17/85
November 24, 2017

The Impact of LHC Jet Data on the MMHT PDF Fit at NNLO

L. A. Harland–Lang[1,2], A. D. Martin[3], R. S. Thorne[2],

[1] *Rudolf Peierls Centre for Theoretical Physics, University of Oxford, 1 Keble Road, OX1 3NP, UK*
[2] *Department of Physics and Astronomy, University College London, WC1E 6BT, UK*
[3] *Institute for Particle Physics Phenomenology, University of Durham, DH1 3LE, UK*

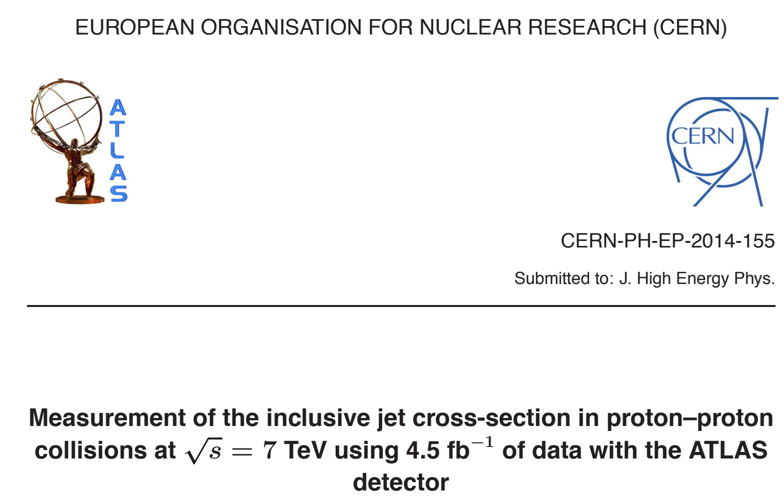
Abstract

We investigate the impact of the high precision ATLAS and CMS 7 TeV measurements of inclusive jet production on the MMHT global PDF analysis at next–to–next–to–leading order (NNLO). This is made possible by the recent completion of the long–

EPJC 78 (2018) no.3, 248

ATLAS jet data - a Challenge

- ATLAS jet data at 7 TeV- extends over ~ **11 orders of magnitude**, with by eye a successful QCD description!
- However, devil is in detail: **fit quality**, $\chi^2/\text{dof} \sim 2-3$, is actually very **poor**.
- Similar effect seen in 8, 13 TeV data. What is going on*? And what about PDFs?

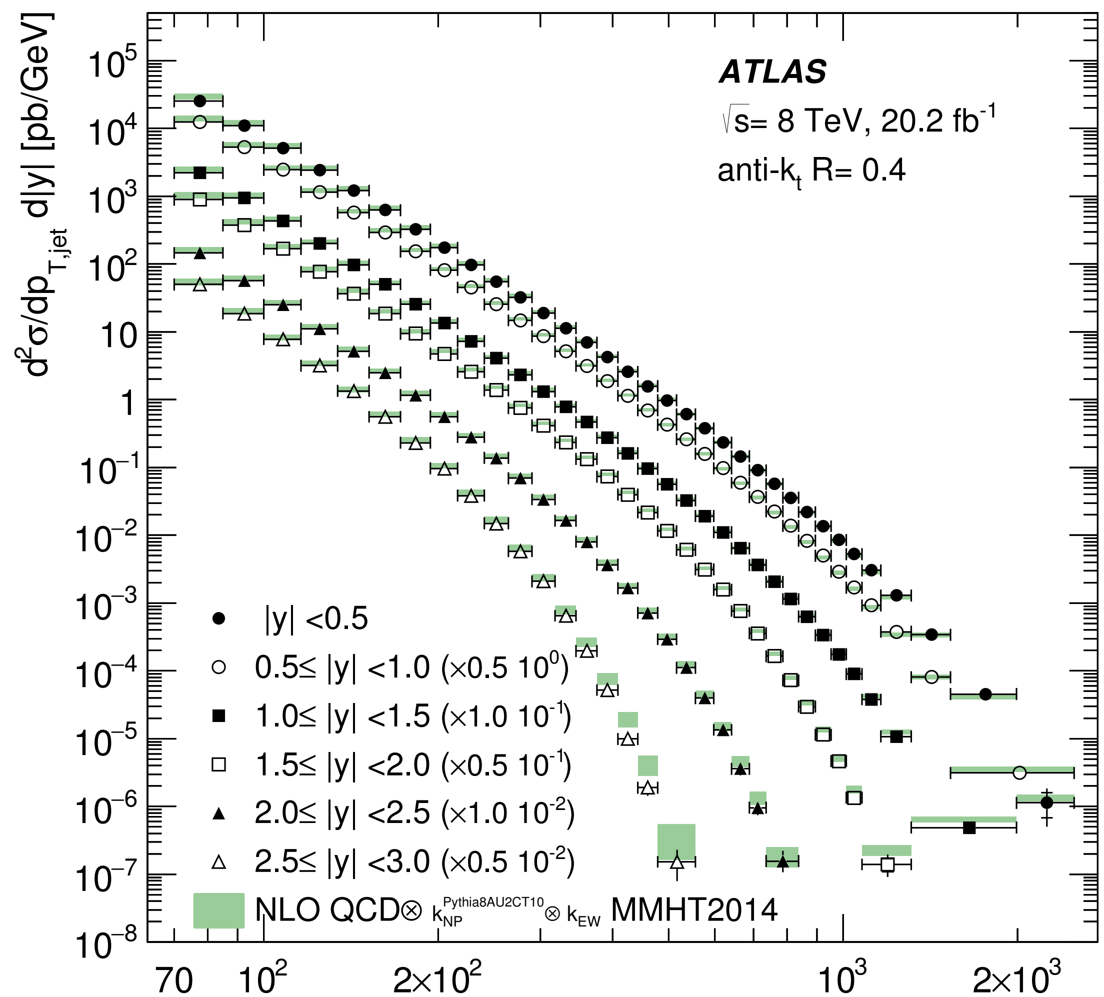

EUROPEAN ORGANISATION FOR NUCLEAR RESEARCH (CERN)

CERN-PH-EP-2014-155

Submitted to: J. High Energy Phys.

Measurement of the inclusive jet cross-section in proton–proton collisions at $\sqrt{s} = 7$ TeV using 4.5 fb^{-1} of data with the ATLAS detector

χ^2/ndf	$p_T^{\text{jet,max}}$ $R = 0.4$	$R = 0.6$
$p_T > 70$ GeV		
CT14	349/171	398/171
HERAPDF2.0	415/171	424/171
NNPDF3.0	351/171	393/171
MMHT2014	356/171	400/171



*Similar poor description in fact also seen by CMS prior to further internal error decorellation.

Boughezal et al., JHEP 1707 (2017) 130

Dealing with Correlated Errors

- Simple case of statistical (uncorrelated) errors only: $\chi^2 = \sum_{i=1}^{N_{\mathrm{pts}}} \left(\frac{D_i - T_i}{\sigma_i^{\mathrm{uncorr}}} \right)^2$
- Adding in N_{corr} **correlated systematics**, this becomes:

$$\chi^2 = \sum_{i=1}^{N_{\mathrm{pts}}} \left(\frac{D_i + \sum_{k=1}^{N_{\mathrm{corr}}} r_k \sigma_{k,i}^{\mathrm{corr}} - T_i}{\sigma_i^{\mathrm{uncorr}}} \right)^2 + \sum_{k=1}^{N_{\mathrm{corr}}} r_k^2,$$

i.e. the **data points** allowed to **shift** by $D_i \to D_i + \sum_{k=1}^{N_{\mathrm{corr}}} r_k \sigma_{k,i}^{\mathrm{corr}}$

with penalty of r_k^2 due to size of shift for each source of correlated error, calculated (analytically) to achieve the smallest overall χ^2.

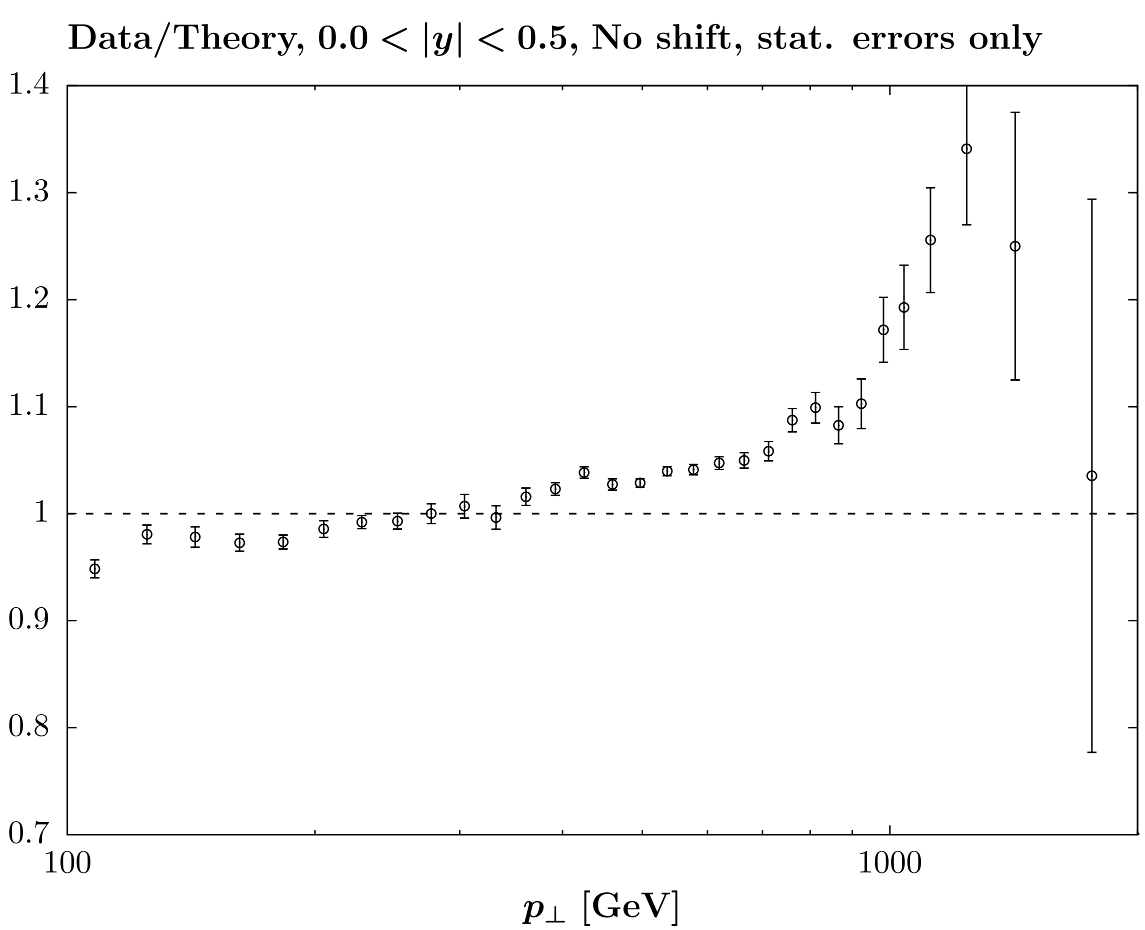
- Simplest example - an overall **normalization** (e.g. luminosity):

$$\sigma_{\mathrm{lumi},i}^{\mathrm{corr}} = \delta_{\mathrm{lumi}} D_i \qquad \delta_{\mathrm{lumi}} : \text{fractional lumi. uncertainty}$$

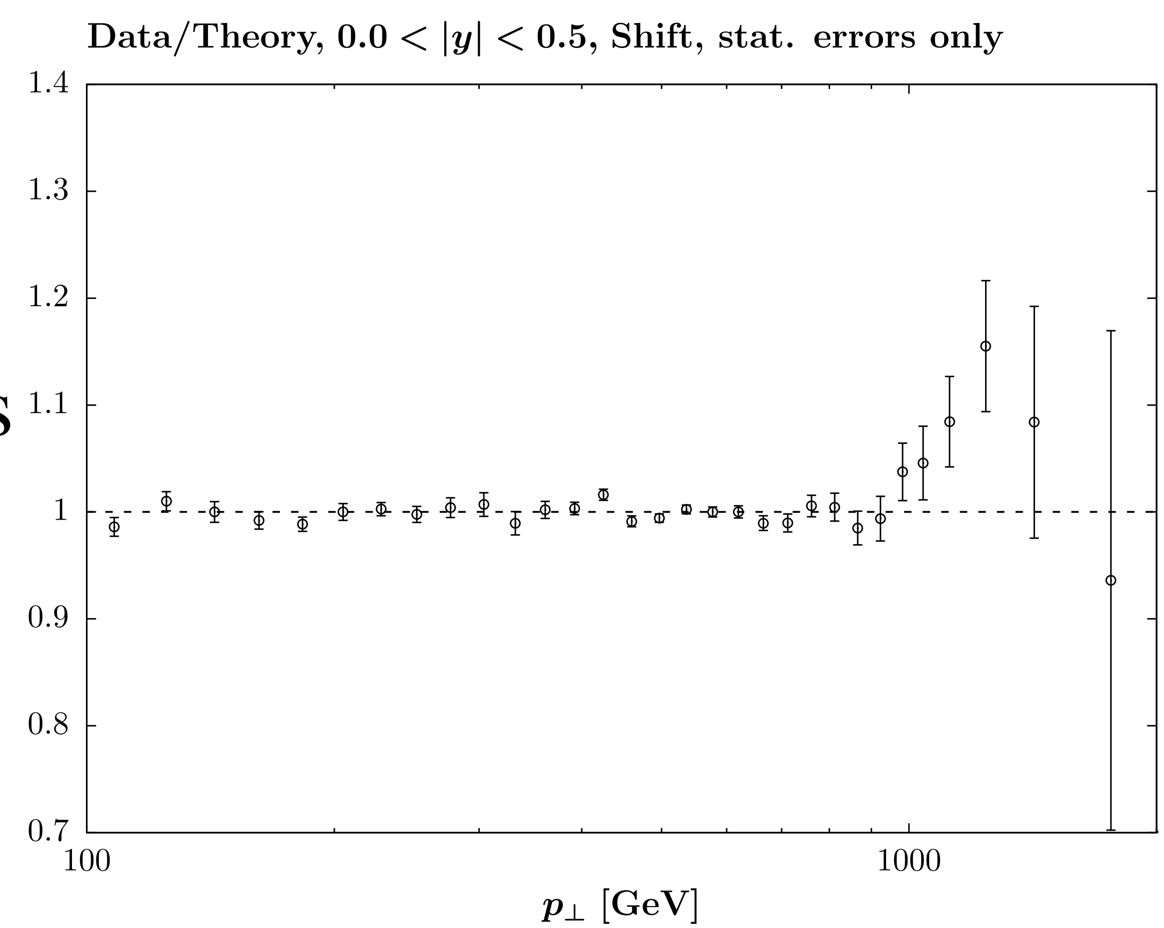
simply shifts data uniformly up/down by some fraction $1 + \delta_{\mathrm{lumi}}$, with single penalty term.

ATLAS jets - systematics

- The ATLAS **systematic errors** are **completely dominant** over the statistical in most regions. The shifts from these determine whether the theory description is good or not. $$D_i \to D_i + \sum_{k=1}^{N_{\rm corr}} r_k \sigma_{k,i}^{\rm corr}$$
- Plot Data/Theory before and after shift (**71 individual sources** in total):

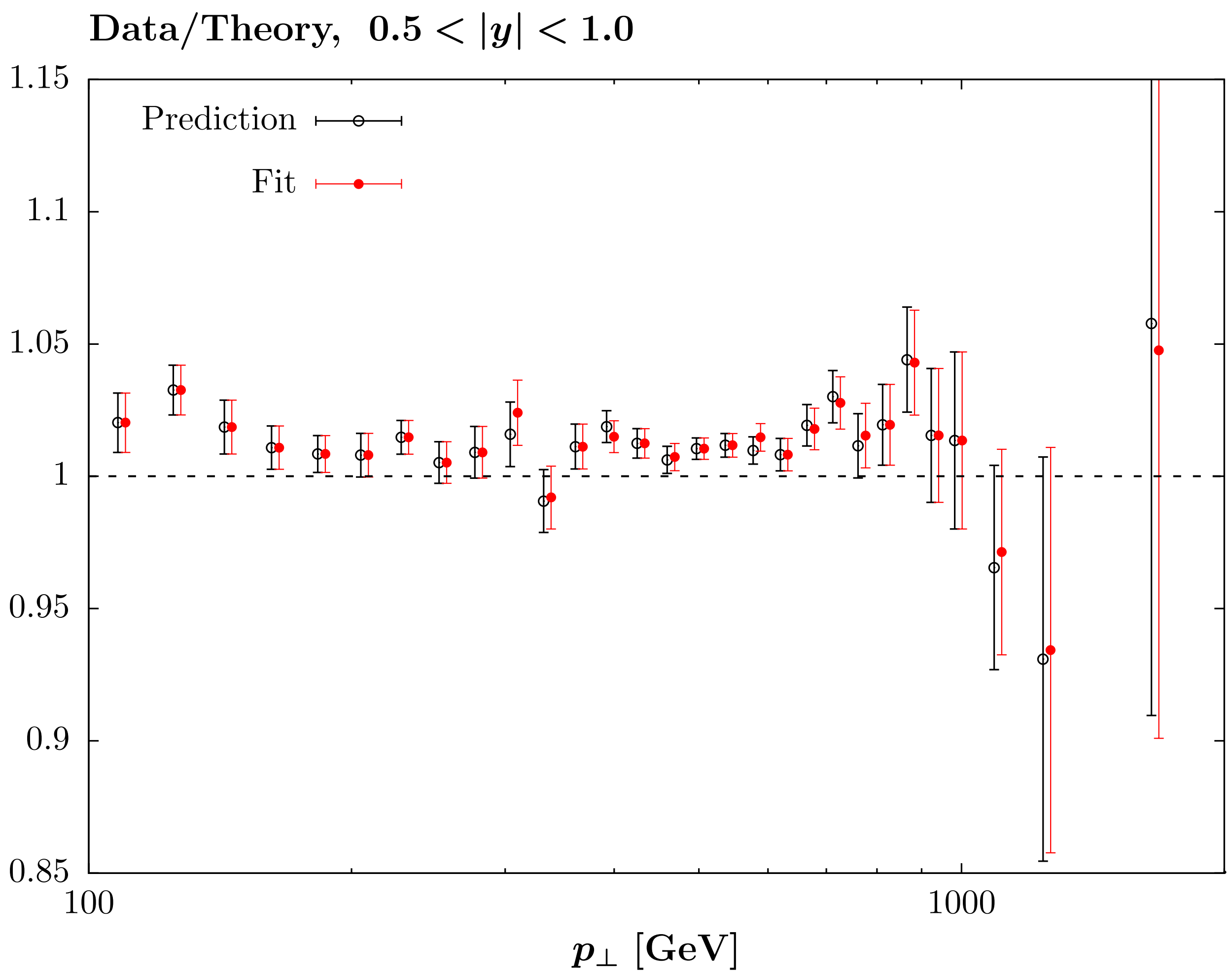


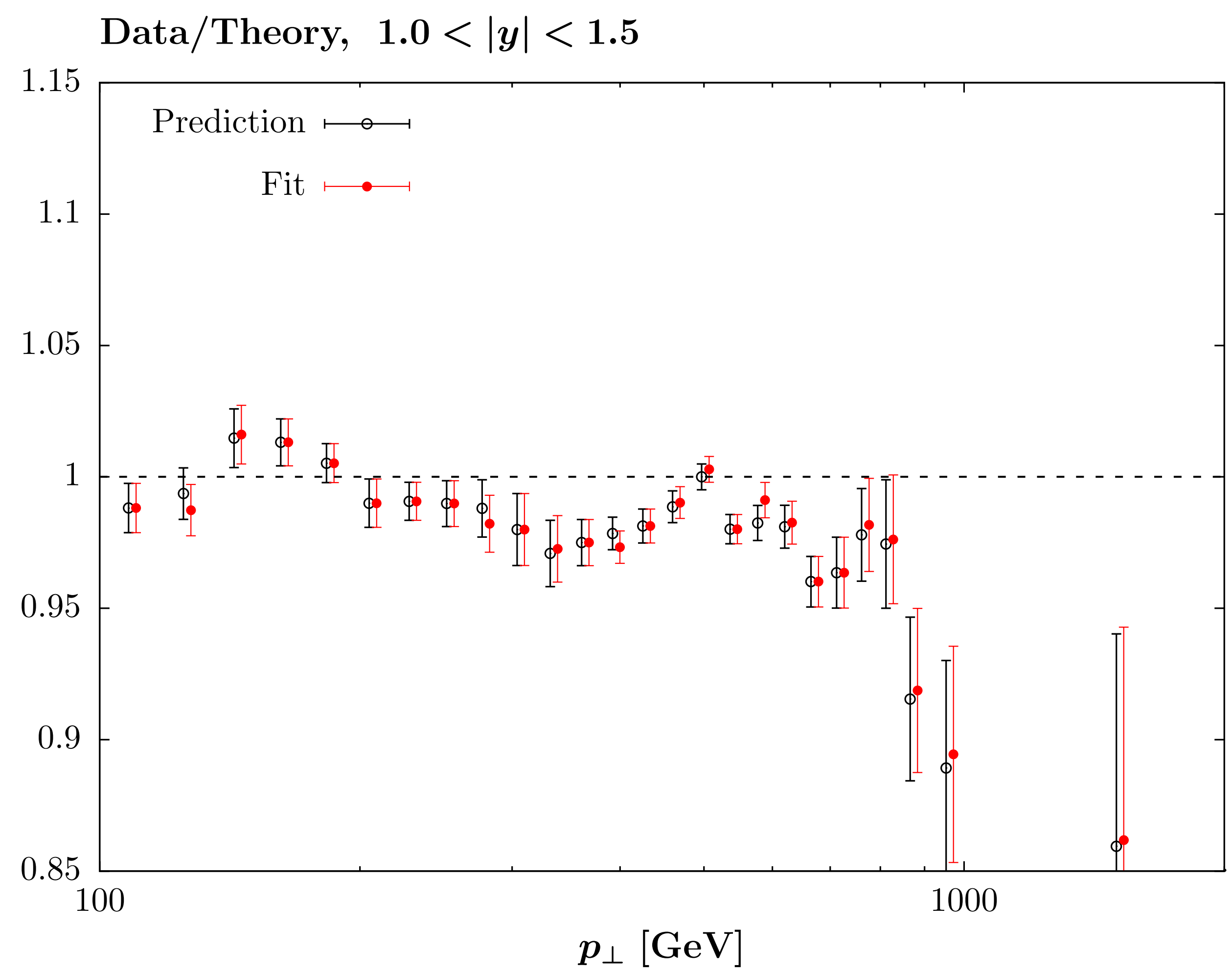
71
shifts
→



- Data presented in different jet rapidity bins. $y_j = -\ln\tan\left(\frac{\theta_j}{2}\right)$

- Systematic shifts allow good description of individual bins, but **not all together**:





$\longrightarrow$ Mismatch in one rapidity bin different in form to neighbouring bins but these are sensitive to PDFs of **same flavour** at **very similar** x and Q^2.

- Might this be due to overly constraining systematic correlations?

Decorrelation - simple approach

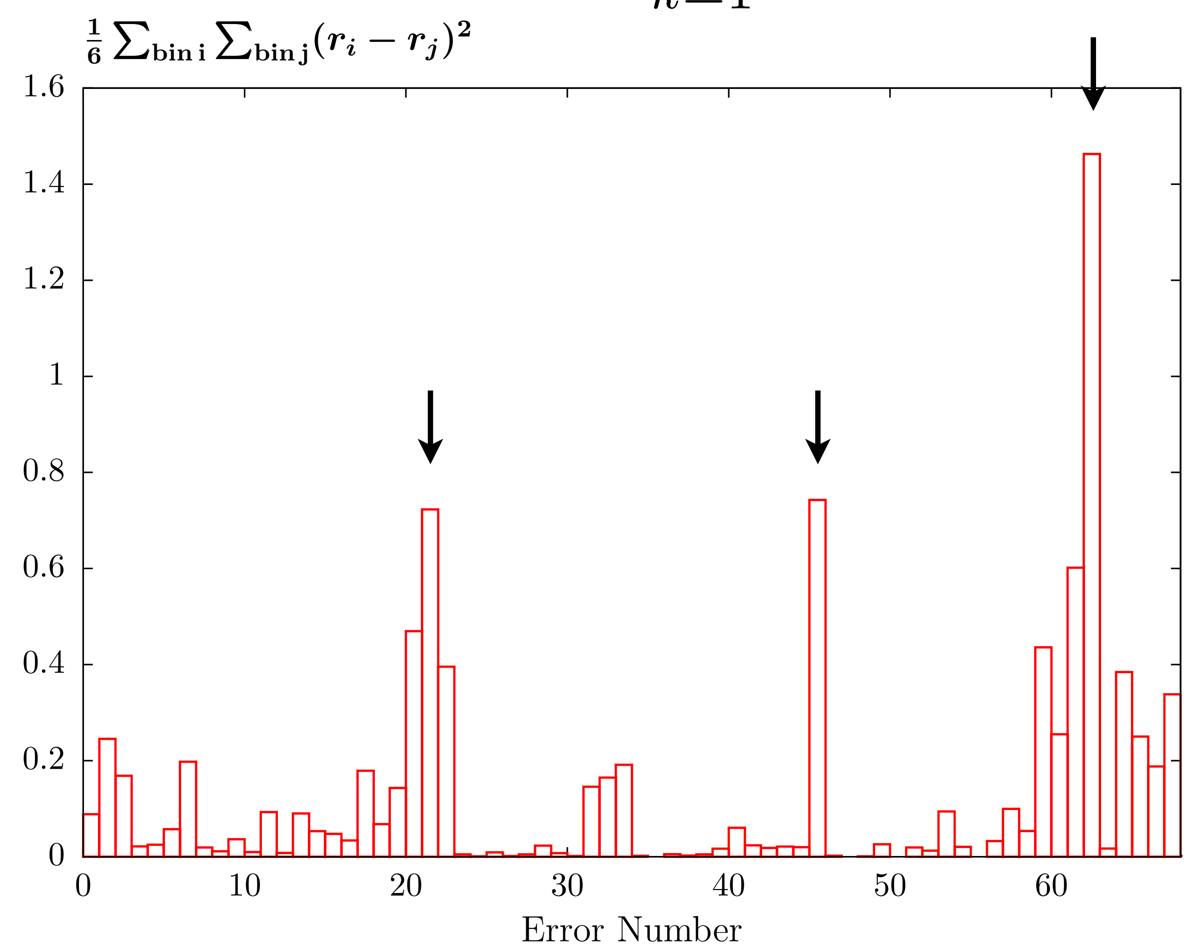
- **Our approach** - study this after the fact. Fit individual jet rapidity bins and see which **systematic shifts** want to go in **different directions**.
- Find in fact only a small number of offenders.
- Simple question - if we loosen correlations, **are PDFs affected**?

$$D_i \to D_i + \sum_{k=1}^{N_{\rm corr}} r_k \sigma_{k,i}^{\rm corr}$$

- Simply allowing a pair of these* to change freely between rapidity gives a ~ **180 point improvement** in χ^2 !

	Full	21	62	21,62
$\chi^2/N_{\rm pts.}$	2.85	1.58	2.36	1.27

- What about the **PDF impact**?



LHL et al., arXiv:1711.05757

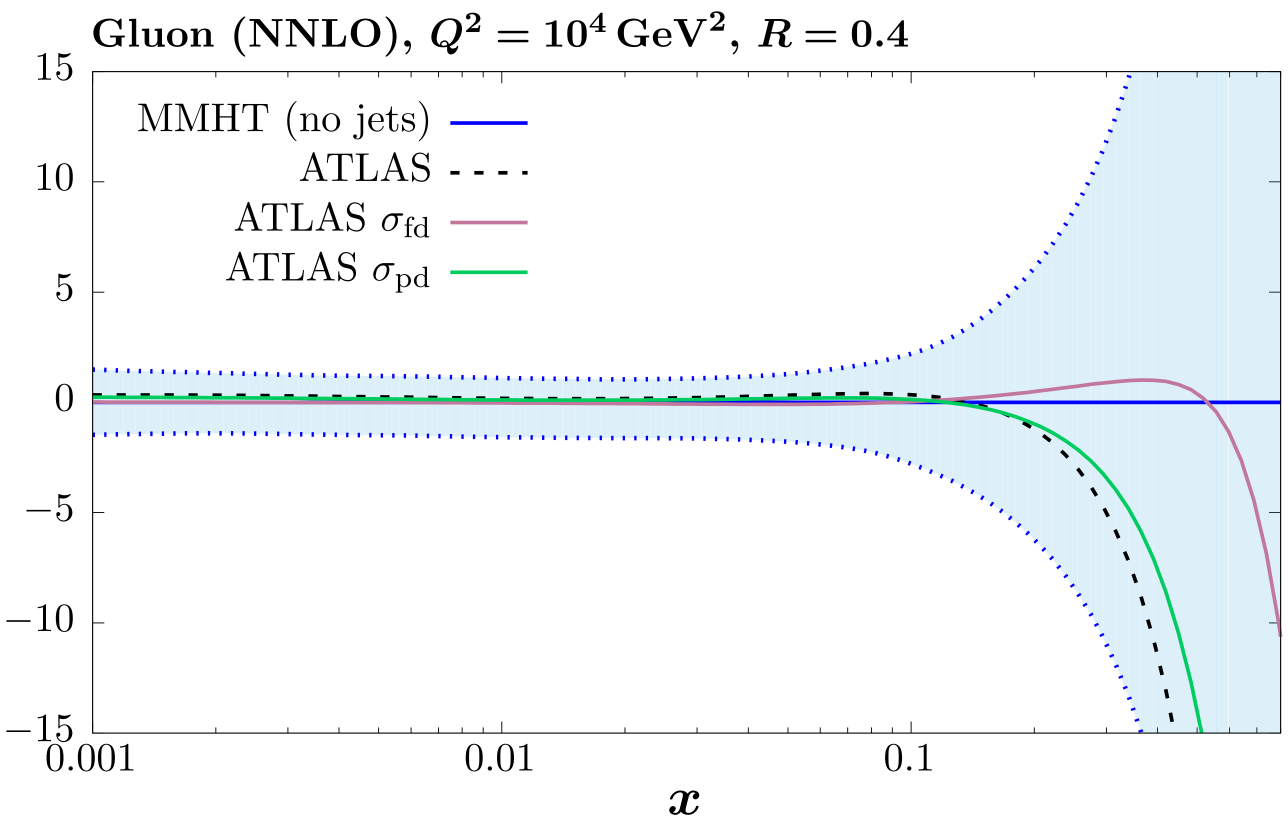
*Checking with ATLAS that it is not completely unreasonable for this choice.

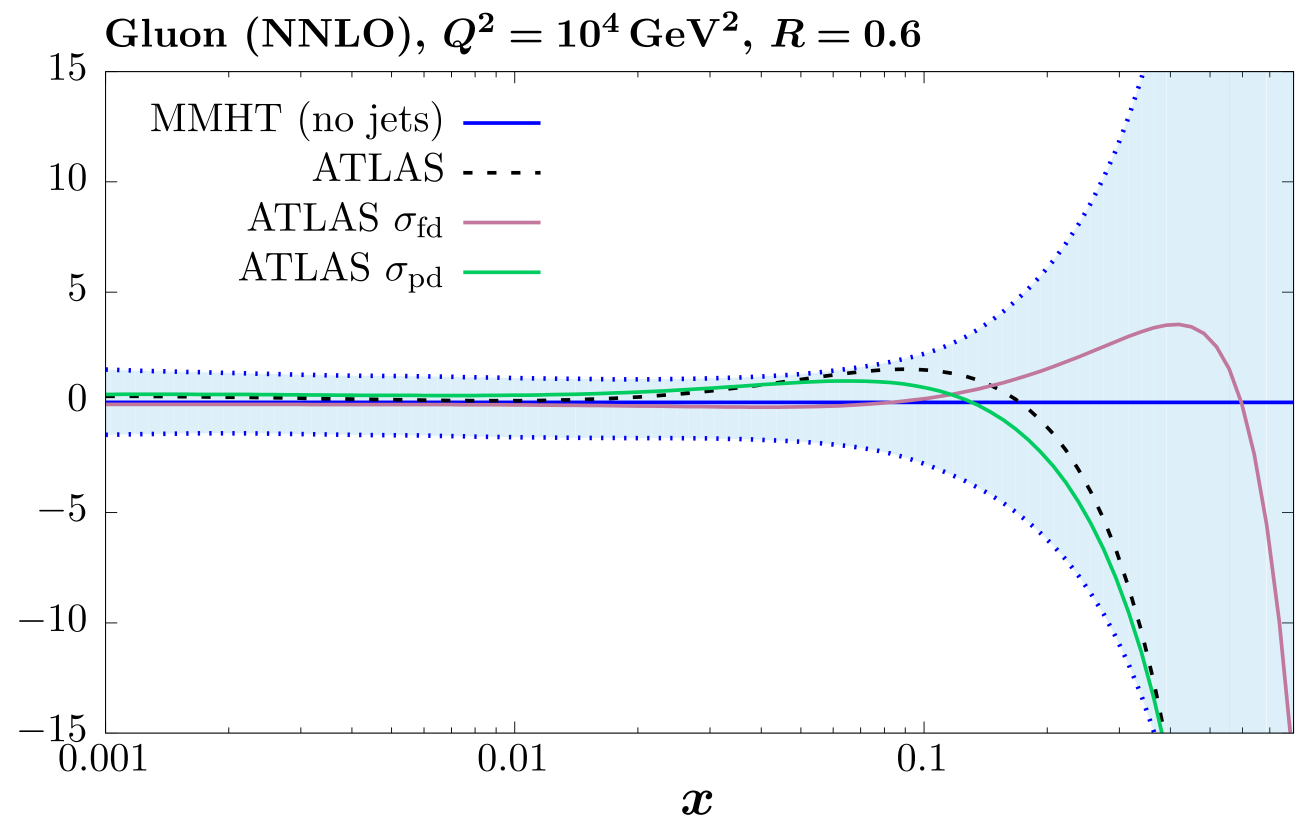
Impact on Gluon

- Despite **huge impact** on χ^2, allowing this extra freedom has **very little impact on the gluon** itself! PDF fit robust (from evolution, presence of other data sets...).

$\longrightarrow$ Despite initial issues with χ^2, can **reliably include in fit.**

	Full	21	62	21,62
$\chi^2/N_{\text{pts.}}$	2.85	1.58	2.36	1.27





Impact on Gluon

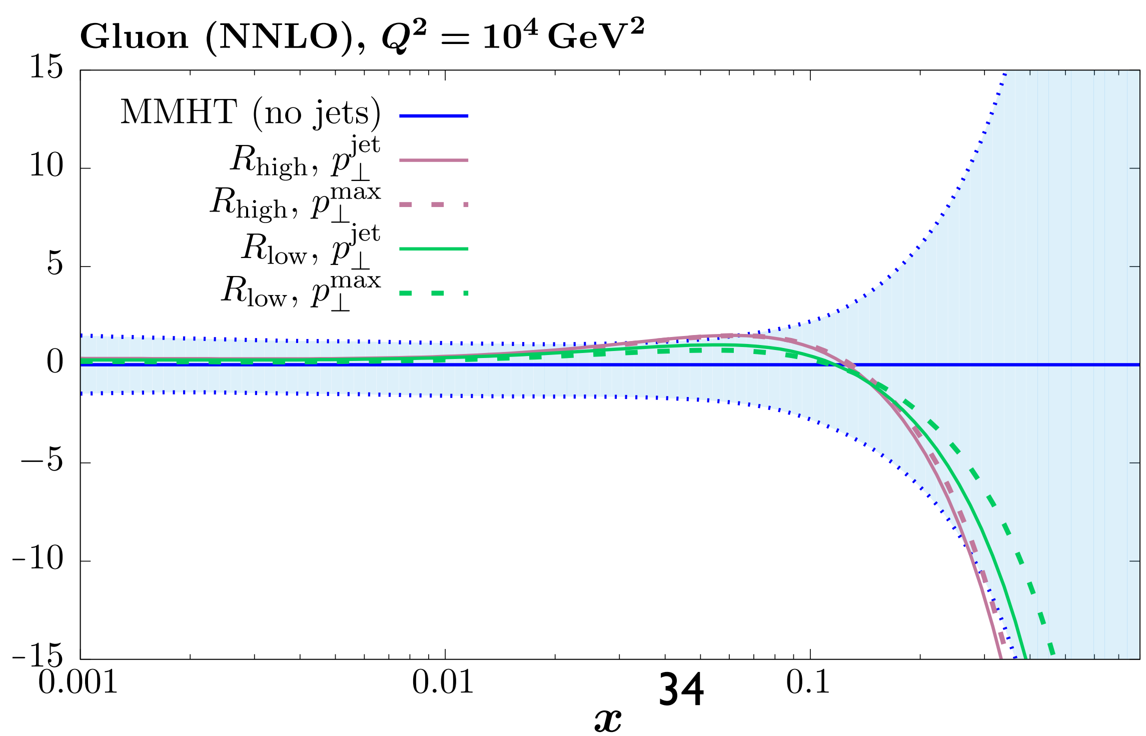
- We find many other **encouraging results**. Fit to ATLAS + CMS jets.

- **Improvement** in description from **NLO** to **NNLO** - pQCD working as it should.

χ^2

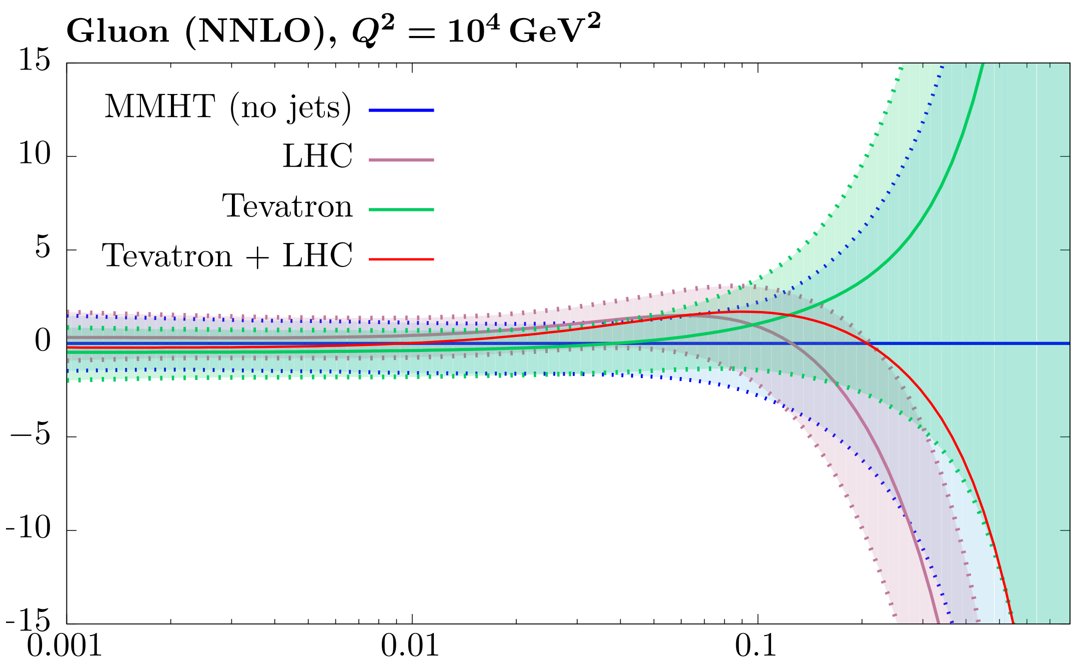
	NLO theory	NNLO
ATLAS, $R_{\rm low}$	215.3	172.3
ATLAS, $R_{\rm high}$	159.2	149.8
CMS, $R_{\rm low}$	194.2	177.8
CMS, $R_{\rm high}$	198.5	182.3

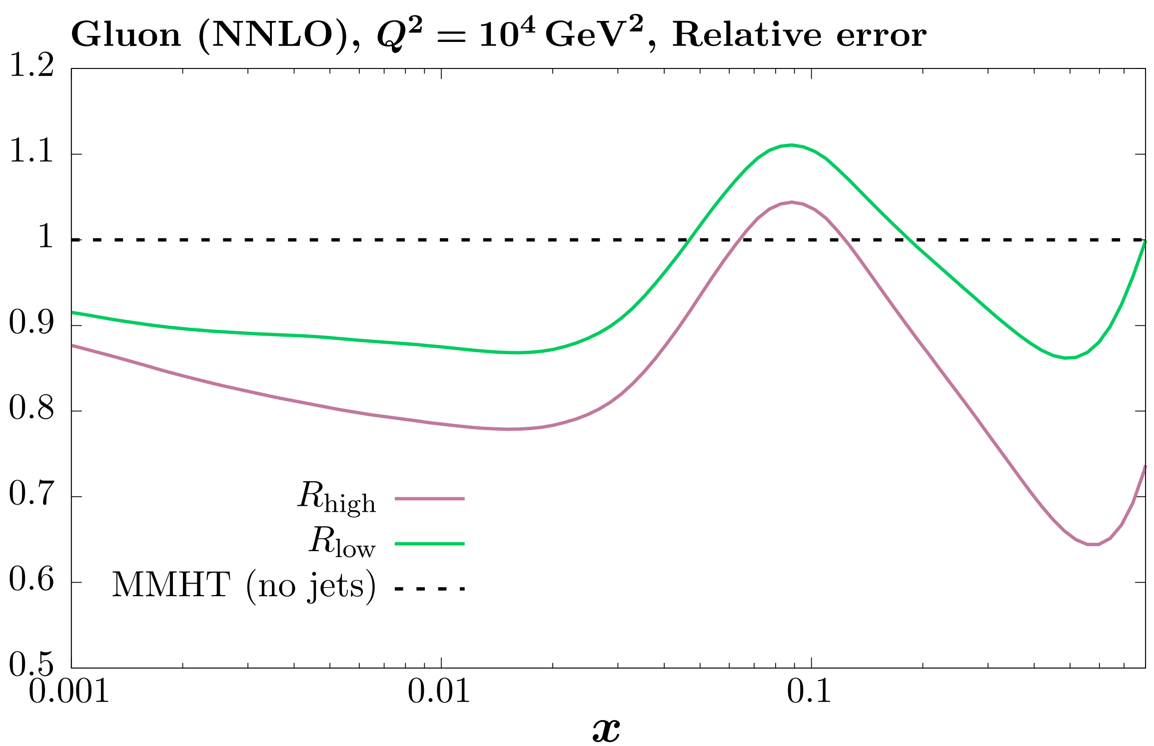
- Different choices of jet radius and the factorization/renormalization scale - either each jet $p_\perp$ or the maximum jet $p_\perp$ in event. Lead to quite different predictions. Find **gluon** quite **insensitive** to these choices:



Impact on Gluon

- Softer gluon at high x , opposite to pull of Tevatron jets. These apply approx. NNLO only $\longrightarrow$ will this change with full theory?





- Reduction in uncertainties over broad x region.

Outlook

- LHC jet data (7, 8, 13 TeV…) will play major role in NNLO fits. Fit stable w.r.t. theory input (scale choice) and jet radius.
- Question of fit quality when fitting all jet rapidity bins. However again stability in fit w.r.t. 'toy' error decorrelation. Important point: fit all data!
- More recent **ATLAS** study for 8 TeV jets: modest decorellations of various sources of **theory** and **experimental** uncertainty.

$\longrightarrow$ Essential for future fits. Our study indicates the fit may not be too sensitive to details, but to to be confirmed.

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
JES Flavour Response Opt 7		
JES MJB Fragmentation Opt 17		
JES Pile-up Rho topology Opt 18		
Scale variations Opt 20		
Alternative scale choice Opt 17		
Non-perturbative corrections Opt 7	261/159	260/159

Splitting option	Sub-component(s) definition(s), completed by complementary
1	$L(\ln(p_T[\text{TeV}]), \ln(0.1), \ln(2.5))\cdot$ uncertainty
2	$L(\ln(p_T[\text{TeV}]), \ln(0.1), \ln(2.5)) \cdot 0.5\cdot$ uncertainty
3	$L(p_T[\text{TeV}], 0.1, 2.5)\cdot$ uncertainty
4	$L(p_T[\text{TeV}], 0.1, 2.5) \cdot 0.5\cdot$ uncertainty
5	$L((\ln(p_T[\text{TeV}]))^2, (\ln(0.1))^2, (\ln(2.5))^2)\cdot$ uncertainty
6	$L((\ln(p_T[\text{TeV}]))^2, (\ln(0.1))^2, (\ln(2.5))^2) \cdot 0.5\cdot$ uncertainty
7	$L(\lvert y\rvert, 0, 3)\cdot$ uncertainty
8	$L(\lvert y\rvert, 0, 3) \cdot 0.5\cdot$ uncertainty
9	$L(\ln(p_T[\text{TeV}]), \ln(0.1), \ln(2.5)) \cdot L(\lvert y\rvert, 0, 3)\cdot$ uncertainty
10	$L(\ln(p_T[\text{TeV}]), \ln(0.1), \ln(2.5)) \cdot \sqrt{1 - L(\lvert y\rvert, 0, 3)^2}\cdot$ uncertainty
11	$L(\ln(p_T[\text{TeV}]), \ln(0.1), \ln(2.5)) \cdot L(\lvert y\rvert, 0, 3) \cdot 0.5\cdot$ uncertainty
12	$L(\ln(p_T[\text{TeV}]), \ln(0.1), \ln(2.5)) \cdot \sqrt{1 - L(\lvert y\rvert, 0, 3)^2} \cdot 0.5\cdot$ uncertainty
13	$L(\ln(p_T[\text{TeV}]), \ln(0.1), \ln(2.5)) \cdot \sqrt{1 - L(\lvert y\rvert, 0, 1.5)^2}\cdot$ uncertainty $L(\ln(p_T[\text{TeV}]), \ln(0.1), \ln(2.5)) \cdot L(\lvert y\rvert, 1.5, 3)\cdot$ uncertainty
14	$L(\ln(p_T[\text{TeV}]), \ln(0.1), \ln(2.5)) \cdot \sqrt{1 - L(\lvert y\rvert, 0, 1)^2}\cdot$ uncertainty $L(\ln(p_T[\text{TeV}]), \ln(0.1), \ln(2.5)) \cdot L(\lvert y\rvert, 1, 3)\cdot$ uncertainty
15	$L(\ln(p_T[\text{TeV}]), \ln(0.1), \ln(2.5)) \cdot \sqrt{1 - L(\lvert y\rvert, 0, 2)^2}\cdot$ uncertainty $L(\ln(p_T[\text{TeV}]), \ln(0.1), \ln(2.5)) \cdot L(\lvert y\rvert, 2, 3)\cdot$ uncertainty
16	$\sqrt{1 - L(\ln(p_T[\text{TeV}]), \ln(0.1), \ln(2.5))^2} \cdot \sqrt{1 - L(\lvert y\rvert, 0, 1.5)^2}\cdot$ uncertainty $\sqrt{1 - L(\ln(p_T[\text{TeV}]), \ln(0.1), \ln(2.5))^2} \cdot L(\lvert y\rvert, 1.5, 3)\cdot$ uncertainty
17	$\sqrt{1 - L(\ln(p_T[\text{TeV}]), \ln(0.1), \ln(2.5))^2} \cdot \sqrt{1 - L(\lvert y\rvert, 0, 1)^2}\cdot$ uncertainty $\sqrt{1 - L(\ln(p_T[\text{TeV}]), \ln(0.1), \ln(2.5))^2} \cdot L(\lvert y\rvert, 1, 3)\cdot$ uncertainty
18	$\sqrt{1 - L(\ln(p_T[\text{TeV}]), \ln(0.1), \ln(2.5))^2} \cdot \sqrt{1 - L(\lvert y\rvert, 0, 2)^2}\cdot$ uncertainty $\sqrt{1 - L(\ln(p_T[\text{TeV}]), \ln(0.1), \ln(2.5))^2} \cdot L(\lvert y\rvert, 2, 3)\cdot$ uncertainty

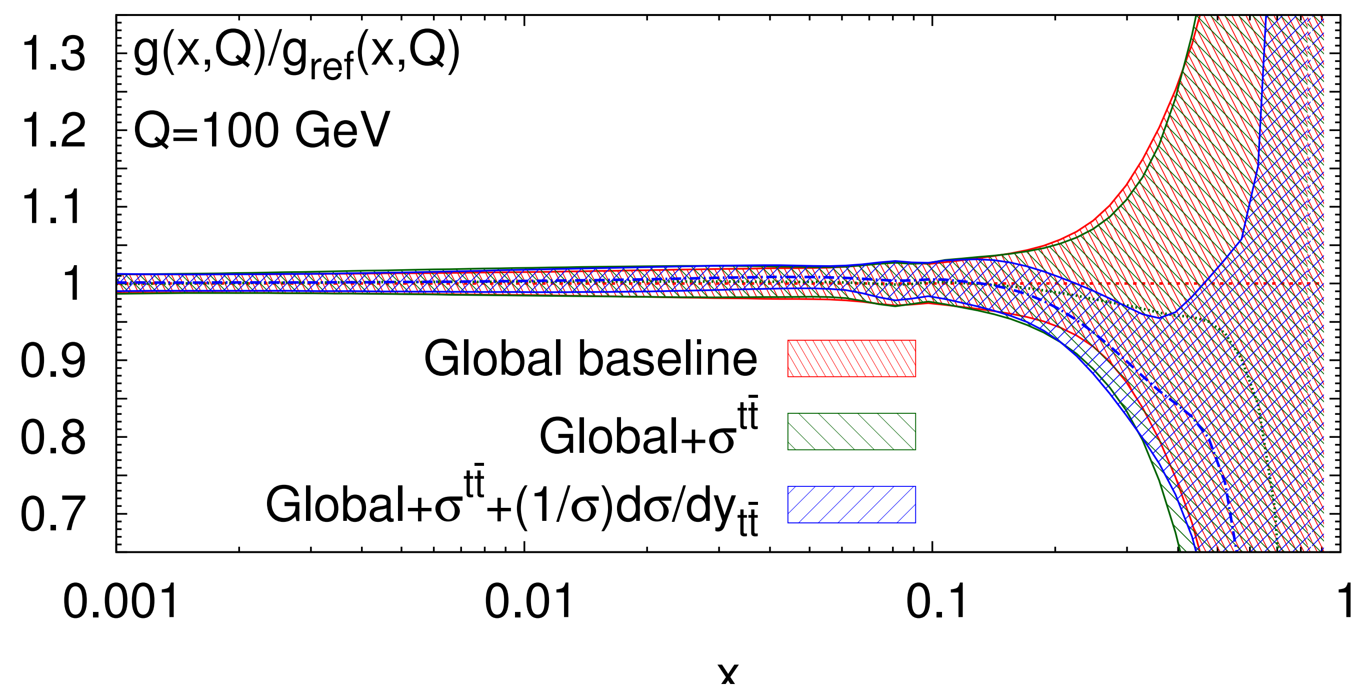
ATLAS, JHEP 09, 020 (2017)

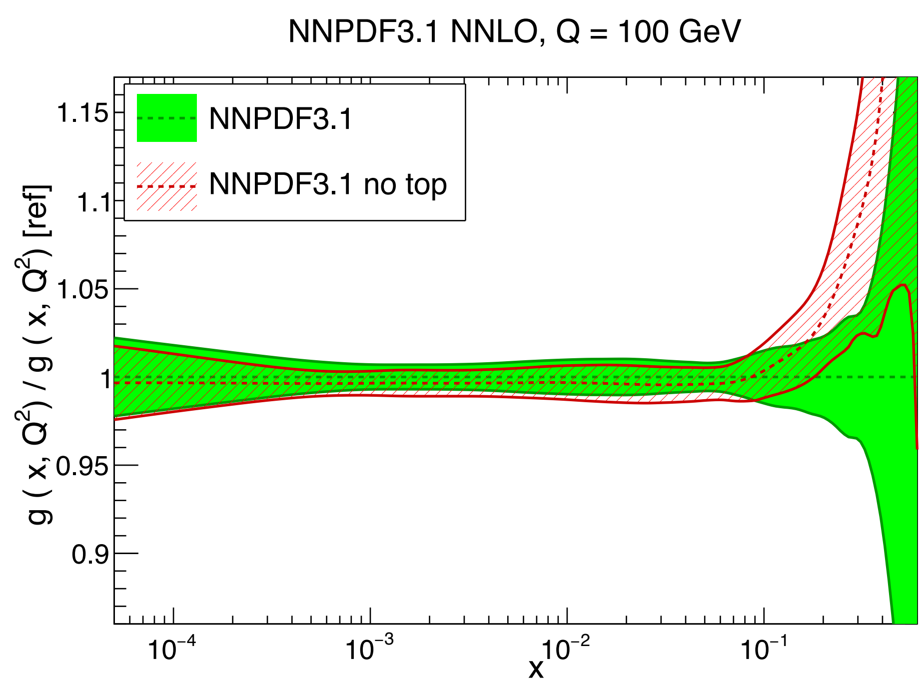
(More) LHC Impact

- LHC data, combined with new NNLO theory, are now playing a significant role in constraining the PDFs. Other examples:

- **Differential top**: increased sensitivity vs. total cross section.
- Fits performed with latest NNLO theory. Impact on gluon at high x .
- Future: double differential?

M. Czakon et al., JHEP 1704 (2017) 044

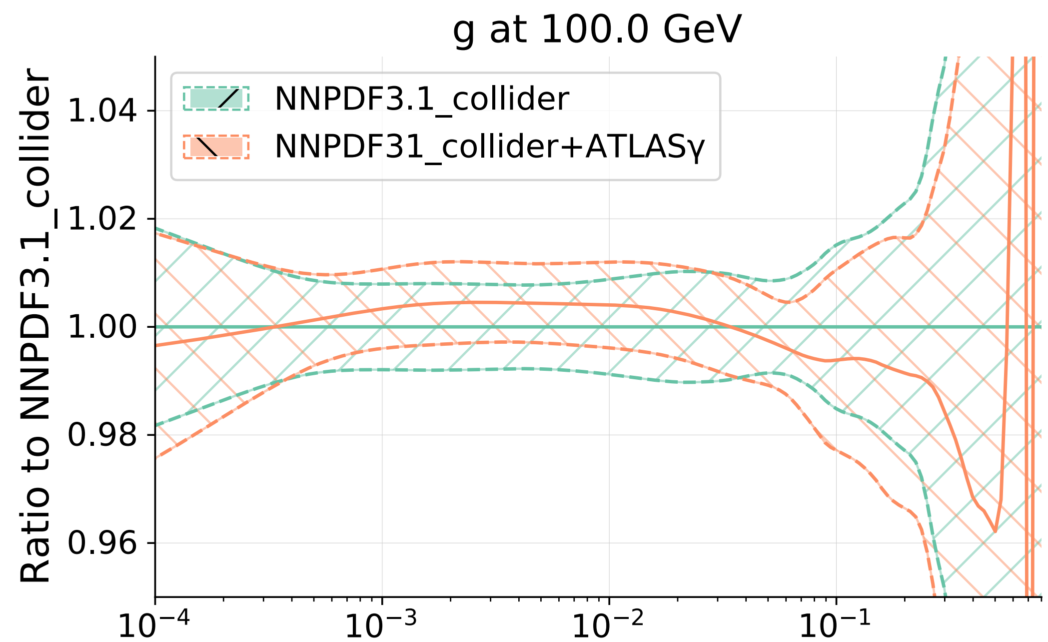


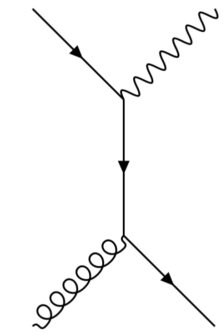


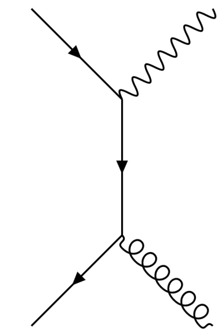
R.D. Ball et al., EPJC 77 (2017) no. 10, 663

See also http://www.precision.hep.phy.cam.ac.uk/results/ttbar-fastnlo/

- **Direct photon**: sensitivity to gluon at intermediate/high x.
- New NNLO calculations allows possibility for inclusion in precision fits.



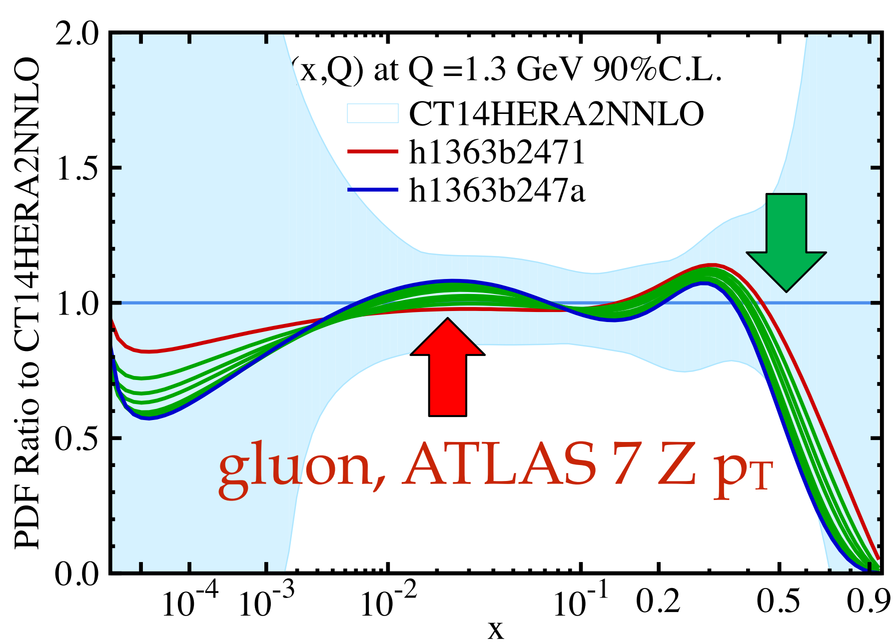




J.M. Campbell et al., arXiv:1802.03021

J. Gao, “Progress on CTEQ-TEA PDFs”, DIS2017

CT17p best-fit vs. CT14 HERA2



- **Z boson $p_\perp$ distribution**. Sensitive to gluon at high . New NNLO calculation allows constraints on PDFs at this order.

Boughezal et al., Phys. Rev. Lett. 116 (2016) no 15 152001

New Calculations- the Photon PDF

Electroweak (EW) Corrections

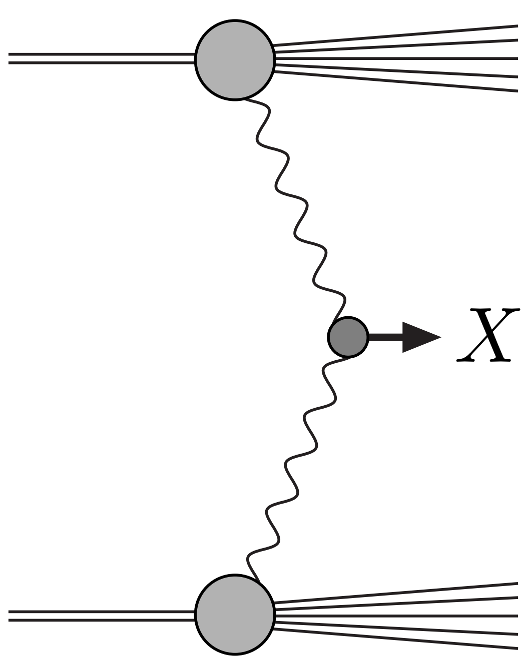
• In era of high precision phenomenology at the LHC: NNLO calculations rapidly becoming the 'standard'. However:

$$\alpha_S^2(M_Z) \sim 0.118^2 \sim \frac{1}{70} \qquad \alpha_{\rm QED}(M_Z) \sim \frac{1}{130}$$

$\longrightarrow$ **EW** and **NNLO QCD** corrections can be **comparable** in size.

• Thus at this level of accuracy, must consider a proper account of EW corrections. At LHC these can be relevant for a range of processes (W, Z, WH, ZH, WW, $t\bar{t}$, jets...).

• For consistent treatment of these, must incorporate QED in initial state: **photon-initiated** production.

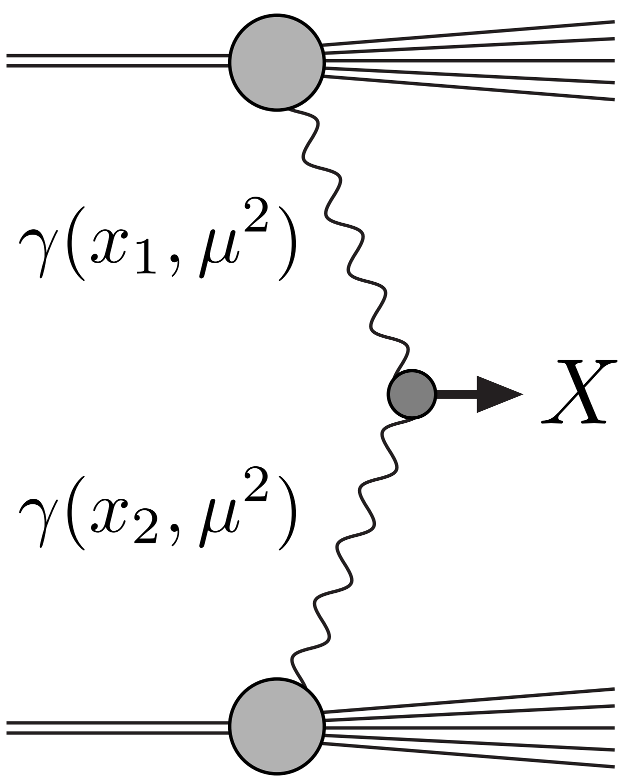


The Photon PDF

- Used to talking about the quarks and gluon within a proton, but what about **photons**?
- Consider $\gamma\gamma \to X$ process, with $X = W^+W^-, l^+l^- \quad \ldots$
- Then can write

$$\sigma^{pp\to X+\cdots} = \sigma^{\gamma\gamma\to X} \otimes \gamma(x_1,\mu^2) \otimes \gamma(x_2,\mu^2)$$

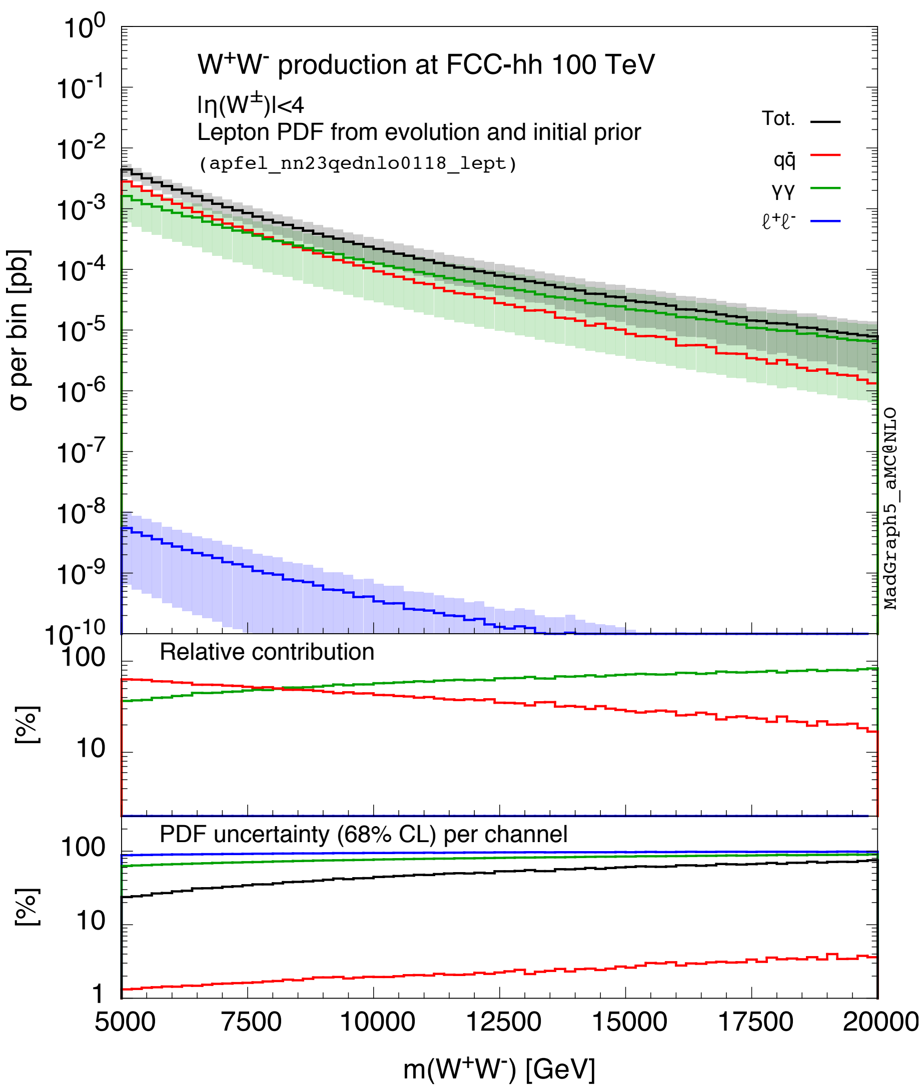
$\gamma(x,\mu^2)$: The **PDF of the photon** within the proton

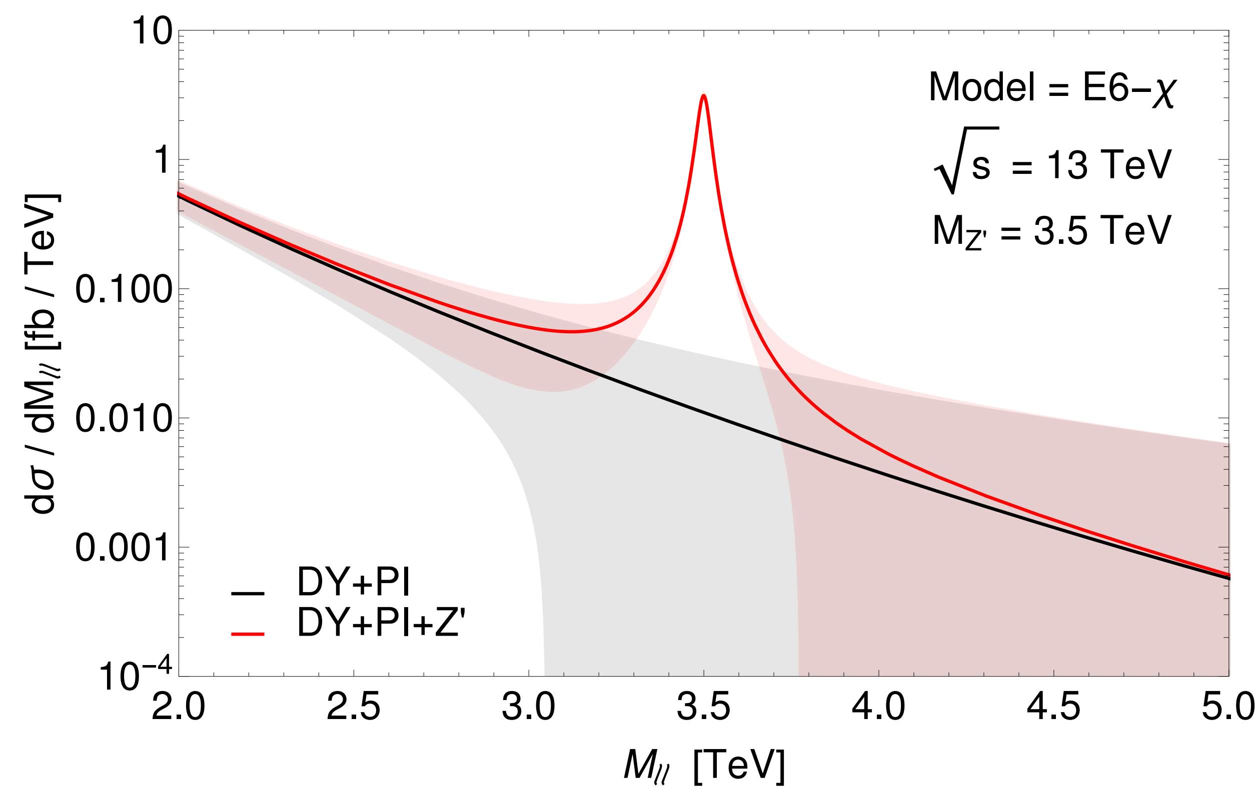


The Photon PDF - Recent Interest

- Recap: earlier studies indicated potentially **big contributions** to $l^+l^-, W^+W^-, t\bar{t}$ … production at large invariant masses, with **sizeable PDF uncertainties**.
- Should we worry?

M.L. Mangano et al., CERN Yellow Report (2017) no.3, 1-254



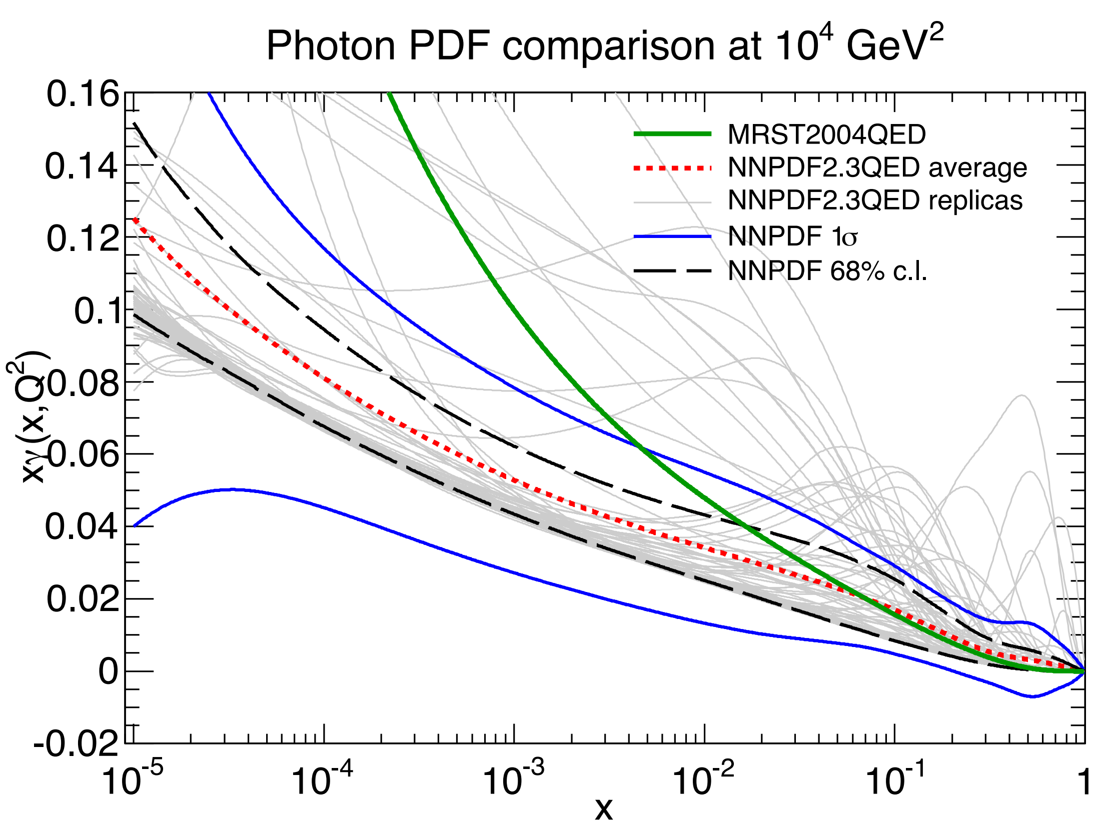


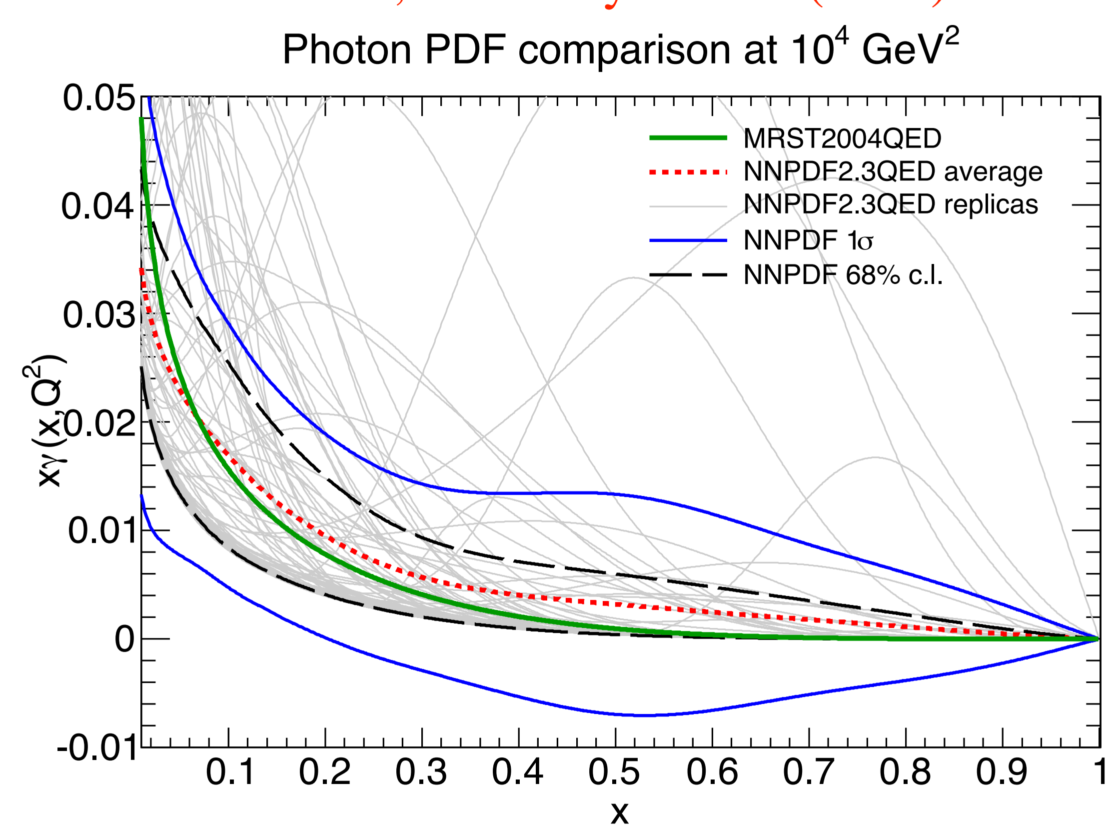
E. Accomando, Phys.Rev. D95 (2017) no.3, 035014

Extracting the Photon PDF

- How do we determine the photon PDF? One option, natural for PDF fitters: simply **parameterise and fit to data**.
- This was done by **NNPDF**. However impact on data (DIS and W, Z) generally small $\Rightarrow$ photon poorly determined. In fact precisely this effect which lead to findings in previous slides.
- Is this the **best we can do?**

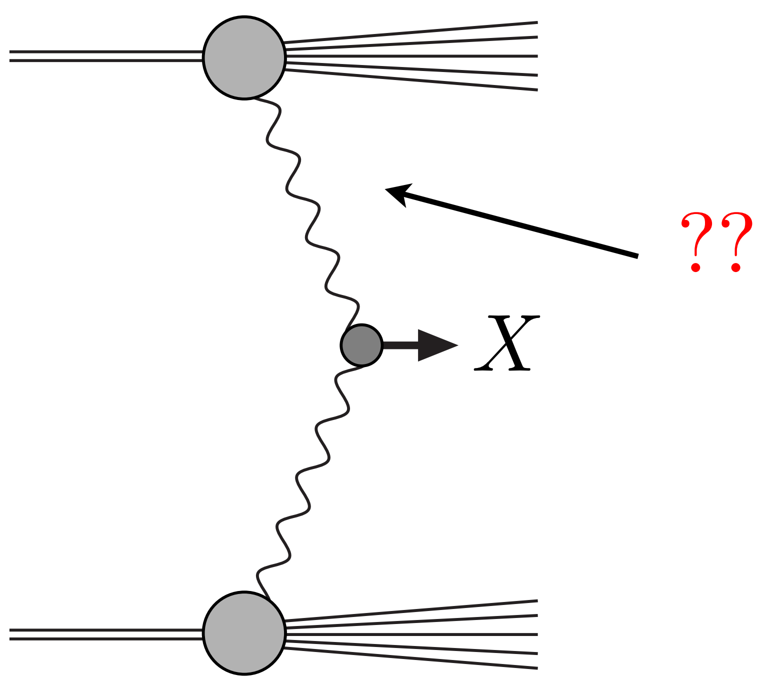
R. D. Ball et al., Nucl. Phys. B877 (2013) 290-320



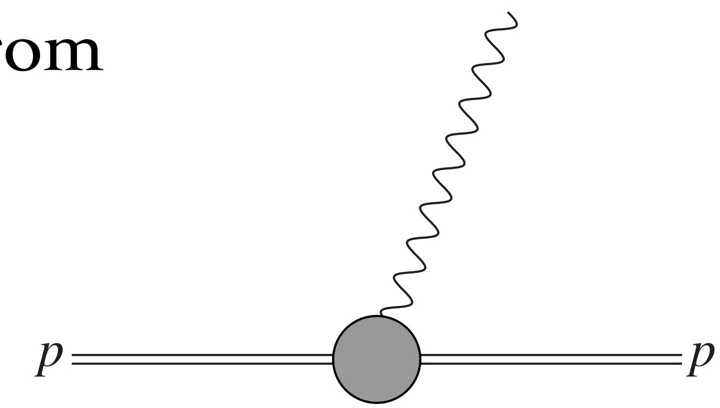


Recent Work

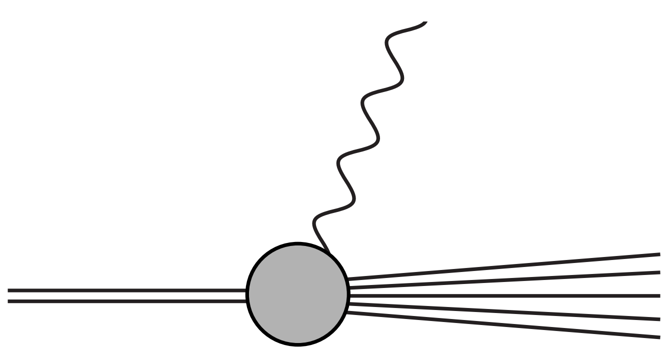
- Simply **fitting** in this way is **far too conservative**, as the photon is in fact quite distinct from the QCD partons $\Rightarrow$ **QED is a long range force**.
- Consider what can generate an initial-state photon at scale $Q_0 \sim 1\,\mathrm{GeV}$ (above this determined by usual DGLAP):



- Dominant component from simple **elastic** $p \to p\gamma$ emission:

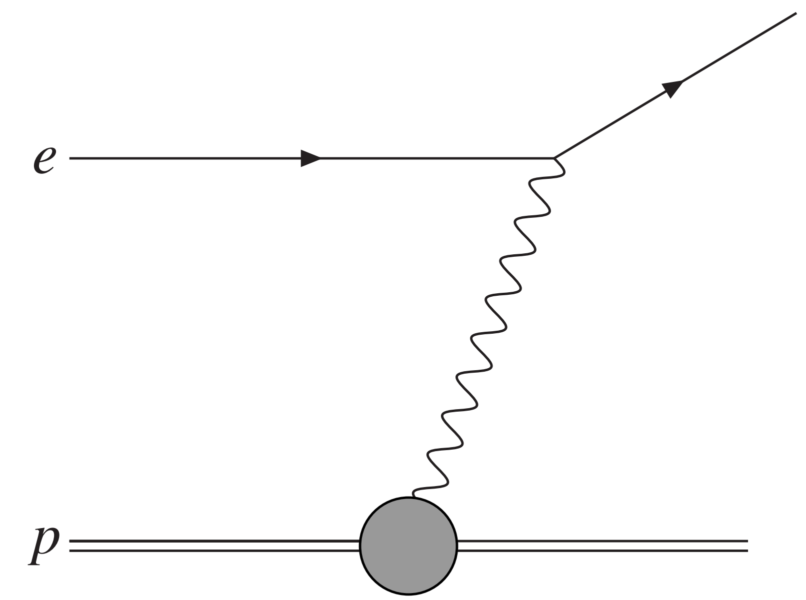


- In addition proton may break up (**inelastic**). Subleading.

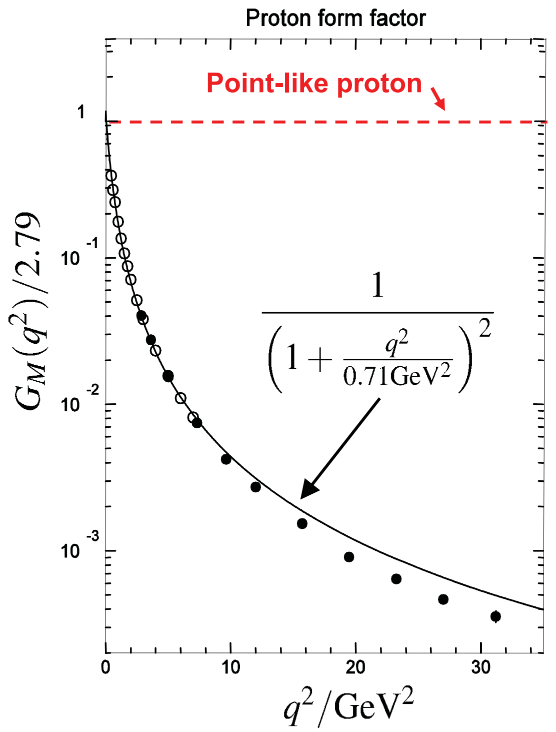


Elastic Emission

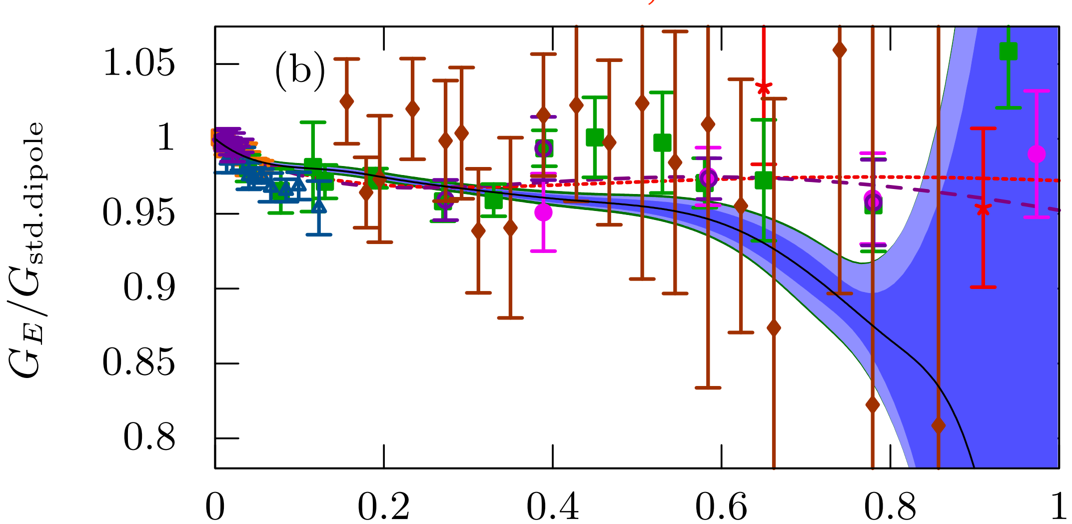
- What do we know about **elastic emission**?
- A lot - exactly the same process as in elastic ep scattering!



- The **form factors** for this are **very well measured** $\Rightarrow$ by thinking a bit more about physics of photon PDF can **constrain precisely**.



A1 Collaboration, arXiv:1307.6227



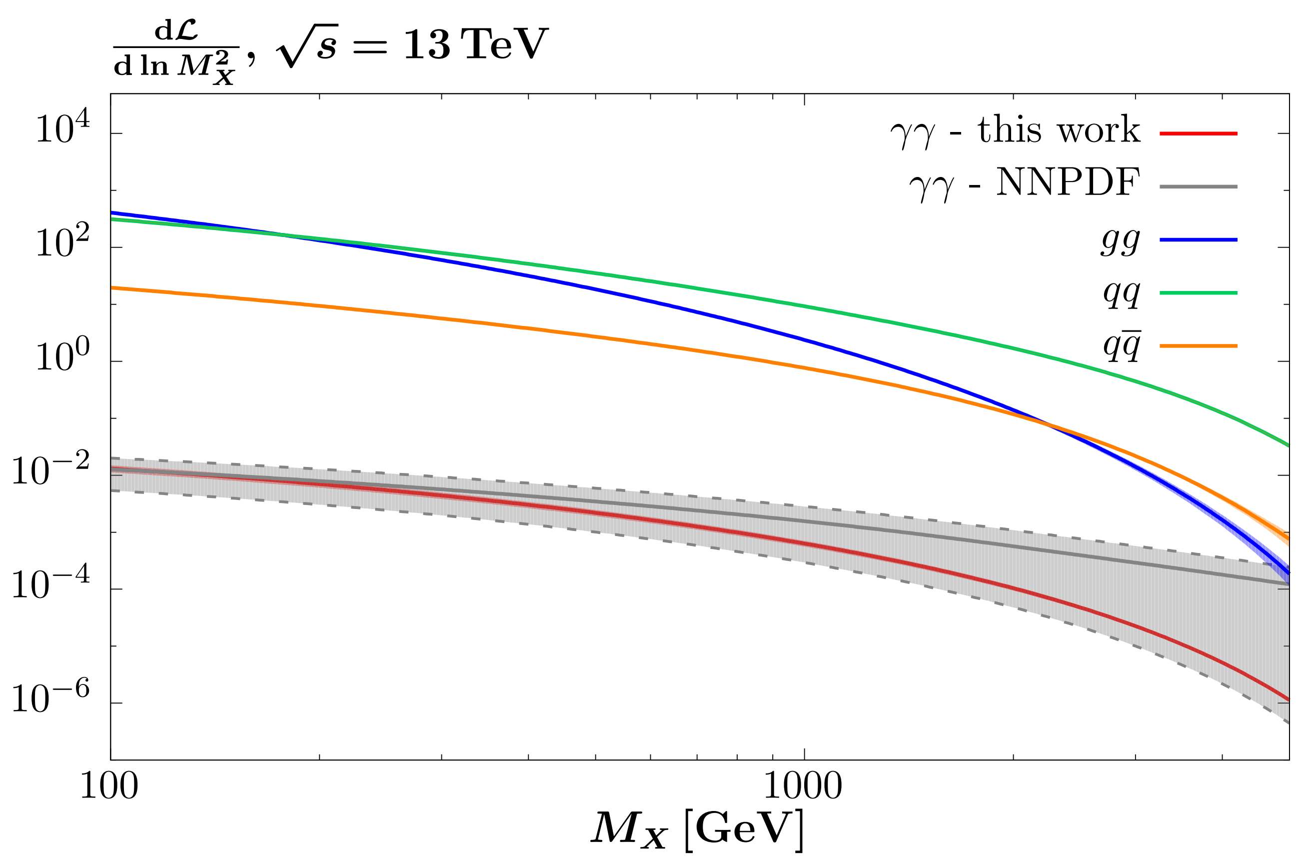
- We can write down **elastic component** of photon PDF:

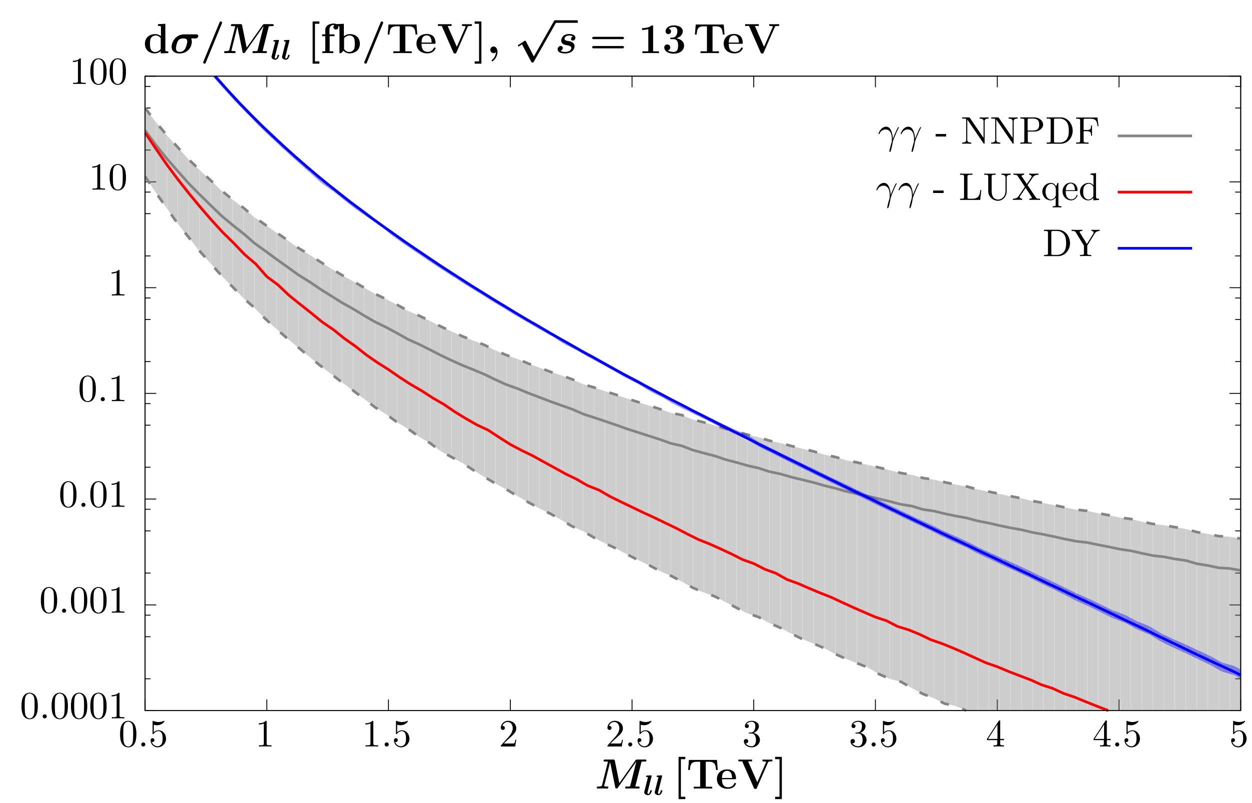
γ transverse mom.

$$\gamma_{\rm coh}(x,Q_0^2) = \frac{1}{x}\frac{\alpha}{\pi}\int_0^{Q^2<Q_0^2} \frac{{\rm d}q_t^2}{q_t^2+x^2m_p^2}\left(\frac{q_t^2}{q_t^2+x^2m_p^2}(1-x)F_E(Q^2)+\frac{x^2}{2}F_M(Q^2)\right)$$

G_E, G_M : proton electric, magnetic form factors

- With simple model for remaining (small) inelastic $p \to X\gamma$ component, get **precise predictions** for photon. **No room for large uncertainties**!

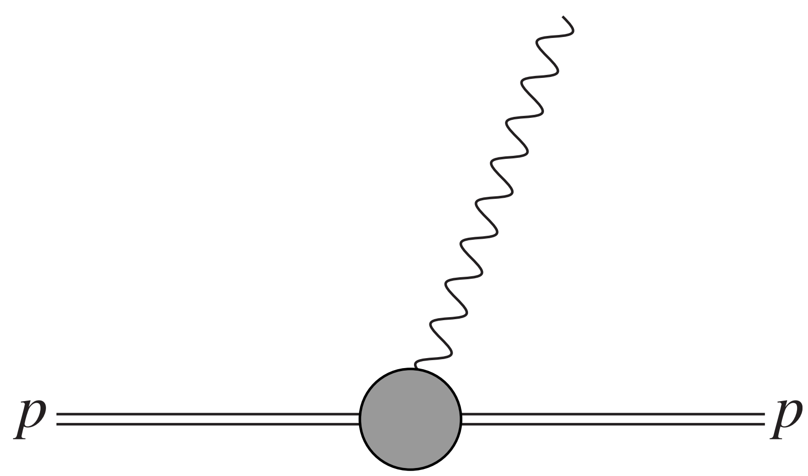




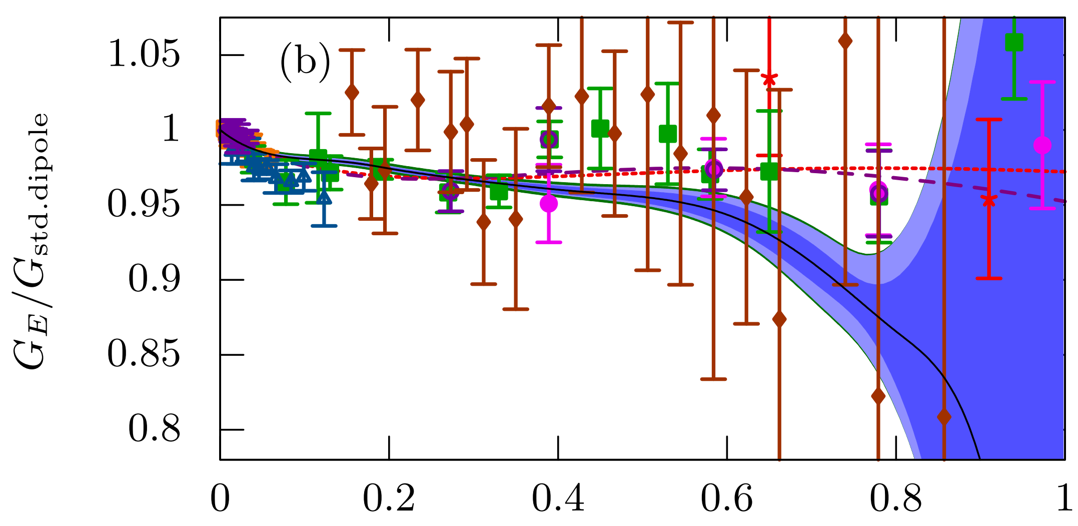
LHL eta al., Phys. Rev. D94 (2016) no.7, 074008

LUXqed

• Have discussed how dominant coherent $p \to p\gamma$ emission process is well constrained from **elastic** ep scattering.

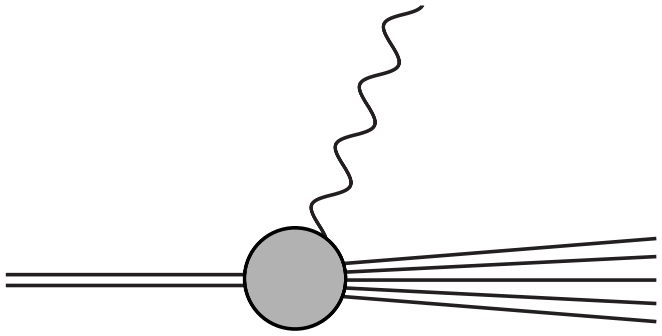


A1 Collaboration, arXiv:1307.6227



• What about inelastic component? Can we not also constrain this from well measured **inelastic** ep scattering?

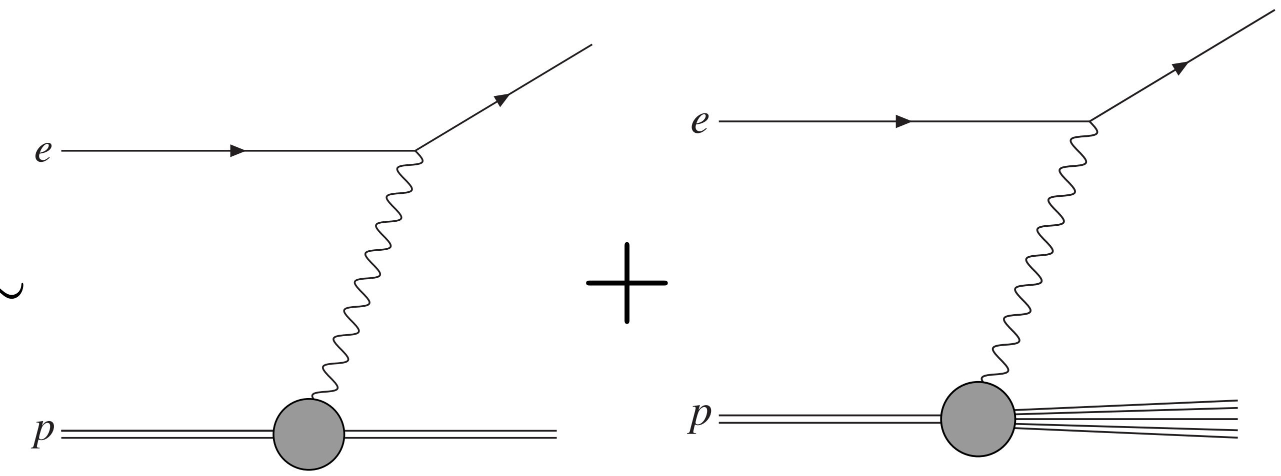
• Yes! $\longrightarrow$ **LUXqed** study shows precisely how this can be done.



- Treatment of Photon put on truly quantitative footing by **LUXqed**. Photon PDF **completely determined** in terms of F_2 and F_L structure functions.

$$x f_{\gamma/p}(x,\mu^2) = \frac{1}{2\pi\alpha(\mu^2)} \int_x^1 \frac{dz}{z} \left\{ \int_{\frac{x^2 m_p^2}{1-z}}^{\frac{\mu^2}{1-z}} \frac{dQ^2}{Q^2} \alpha^2(Q^2) \left[\left(z p_{\gamma q}(z) + \frac{2x^2 m_p^2}{Q^2} \right) F_2(x/z, Q^2) - z^2 F_L\left(\frac{x}{z}, Q^2\right) \right] - \alpha^2(\mu^2) z^2 F_2\left(\frac{x}{z}, \mu^2\right) \right\}, \quad (6)$$

$$\gamma(x,Q^2) \sim$$

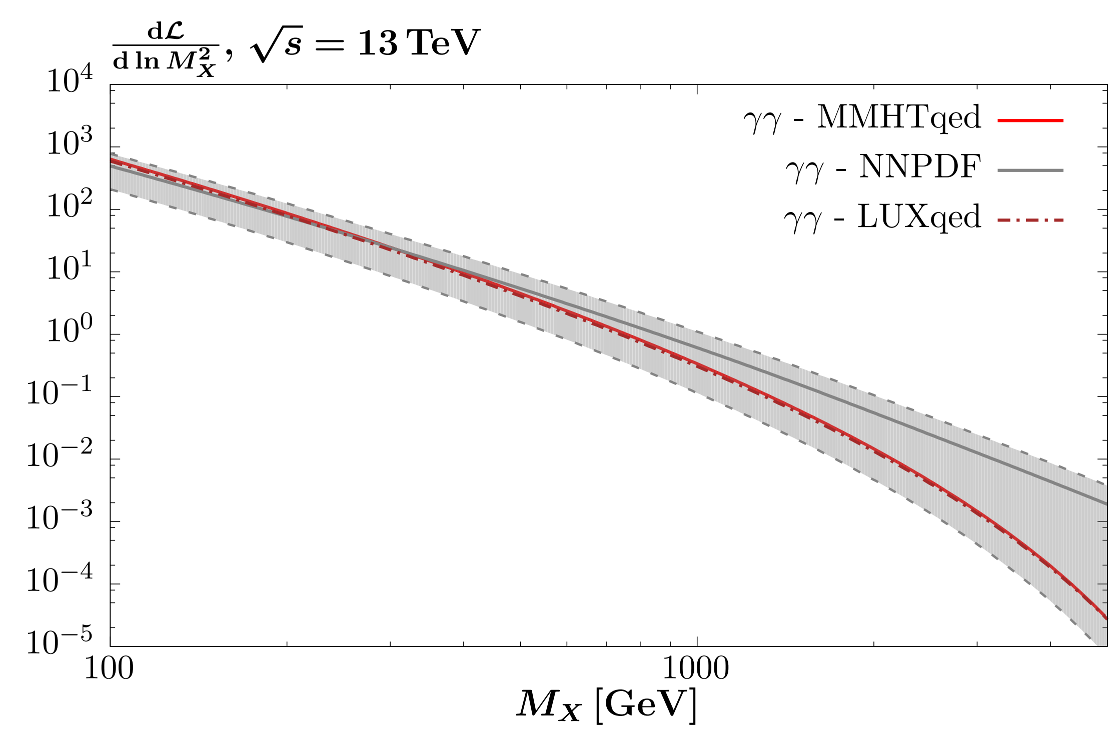


Elastic **Inelastic**

A. Manohar et al., JHEP 1712 (2017) 046

A. Manohar et al., Phys. Rev. Lett. 117 (2016) no.24, 100001

- Conclusion: photon PDF known to % level precision across relevant x .
- Moved beyond era of large photon PDF uncertainties. Photon has gone from being the poorest to the **best constrained** parton!



Implementing LUX

- Conclusion from above: photon has gone from being the poorest to the **best constrained** parton! However LUX formula not directly amenable to use in PDF fit:

★ Cross talk between q, g and γ?

★ Effect of refitting?

★ Neutron PDF?

$$xf_{\gamma/p}(x,\mu^2) = \frac{1}{2\pi\alpha(\mu^2)}\int_x^1 \frac{dz}{z}\left\{\int_{\frac{x^2m_p^2}{1-z}}^{\frac{\mu^2}{1-z}} \frac{dQ^2}{Q^2}\alpha^2(Q^2)\left[\left(zp_{\gamma q}(z)+\frac{2x^2m_p^2}{Q^2}\right)F_2(x/z,Q^2)-z^2F_L\left(\frac{x}{z},Q^2\right)\right] - \alpha^2(\mu^2)z^2F_2\left(\frac{x}{z},\mu^2\right)\right\}, \quad (6)$$

- Currently pursued by:

MMHTQED NNPDF3.1luxQED

- Also work in early stages for CT.
- In general, all future set should (will?) have photon included by default via high precision LUX determination. In more detail…

MMHTQED

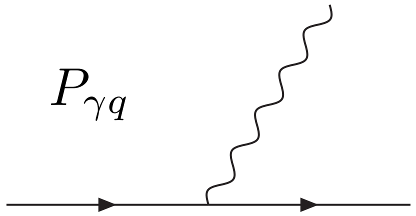
- Ongoing work towards **MMHTQED** - connect LUXqed to standard DGLAP approach at input scale Q_0.
- Breaking this down:

$$x f_{\gamma/p}(x,\mu^2) = \frac{1}{2\pi\alpha(\mu^2)}\int_x^1 \frac{dz}{z}\left\{ \int_{\frac{x^2 m_p^2}{1-z}}^{\frac{\mu^2}{1-z}} \frac{dQ^2}{Q^2}\alpha^2(Q^2) \left[\left(z p_{\gamma q}(z) + \frac{2x^2 m_p^2}{Q^2}\right) F_2(x/z,Q^2) - z^2 F_L\left(\frac{x}{z},Q^2\right)\right] - \alpha^2(\mu^2) z^2 F_2\left(\frac{x}{z},\mu^2\right)\right\}, \quad (6)$$

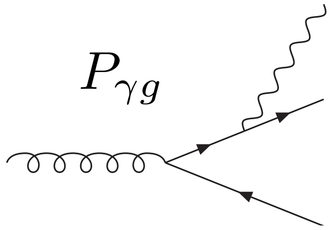
- Constraint at lower scale Q^2 comes from measured F_2. What about larger Q^2 ? Expression reduces to usual DGLAP:

$$\gamma(x,\mu^2) = \gamma(x,Q_0^2) + \int_{Q_0^2}^{\mu^2} \frac{\alpha(Q^2)}{2\pi}\frac{dQ^2}{Q^2}\int_x^1 \frac{dz}{z}\left(P_{\gamma\gamma}(z)\gamma(\frac{x}{z},Q^2) + \sum_q e_q^2 P_{\gamma q}(z) q(\frac{x}{z},Q^2) + P_{\gamma g}(z) g(\frac{x}{z},Q^2)\right),$$

$P_{\gamma q}$



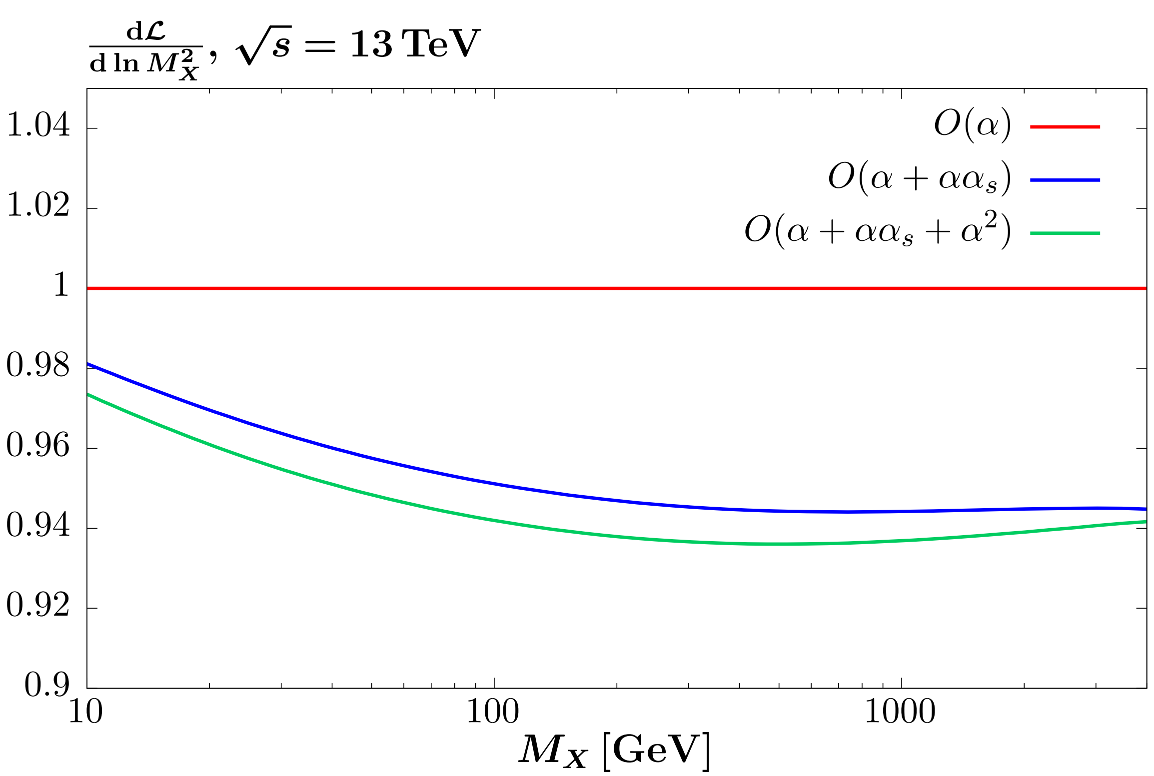
$P_{\gamma g}$

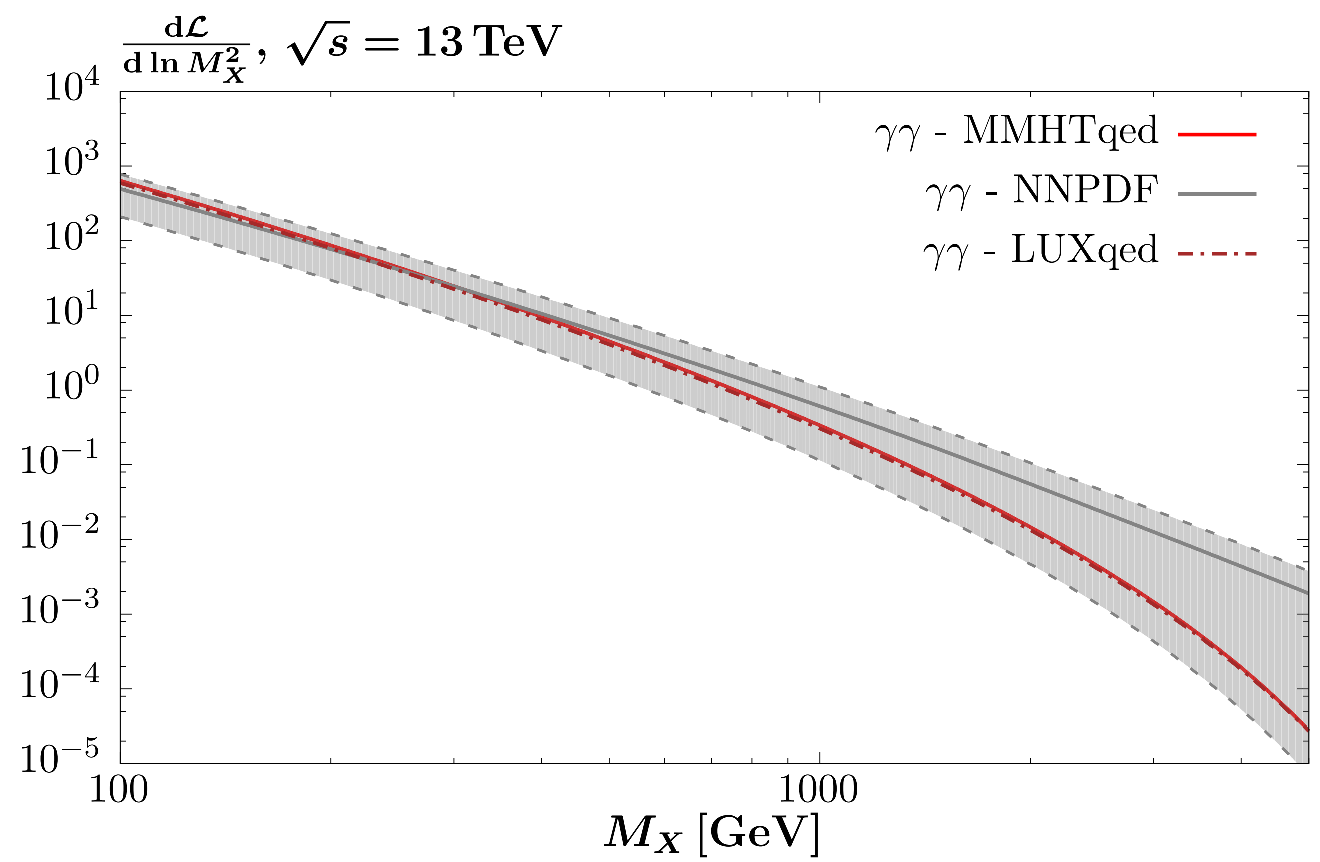


$\longrightarrow$ Photon due (as we expect) to $q \to q\gamma$ emission. Given in terms of known q, g PDFs. Must **merge** with PDFs from **global fit**.

MMHTQED

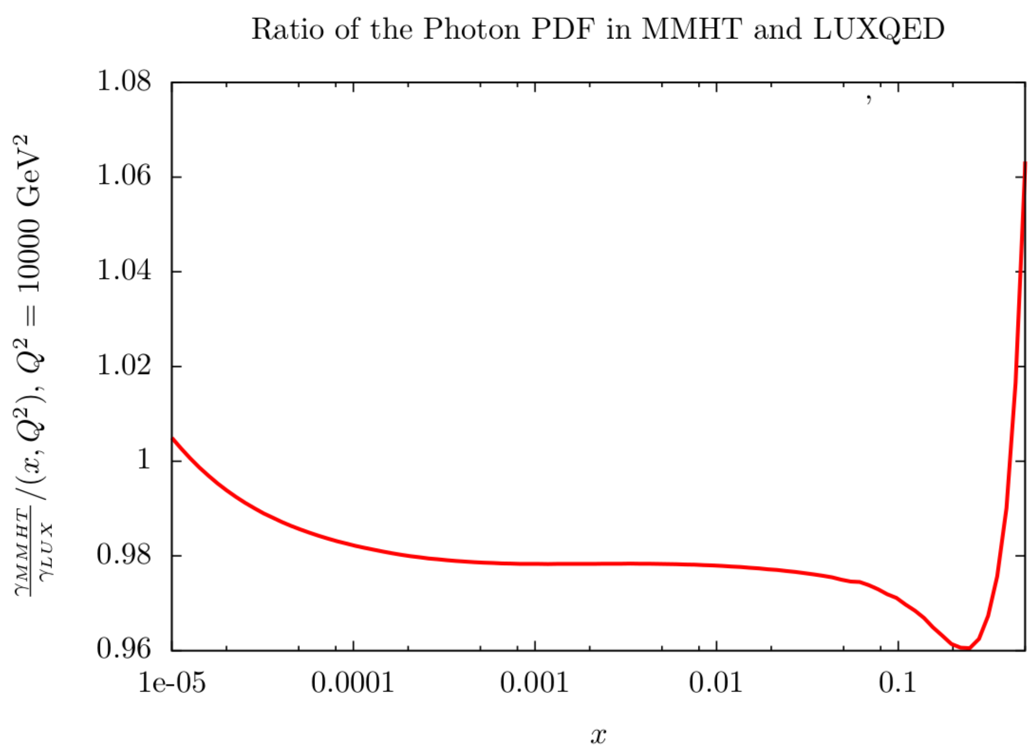
- Ongoing work towards **MMHTQED** - connect LUXqed to standard DGLAP approach at input scale Q_0.
- $Q < Q_0$: input photon based on LUXqed formula ($\gamma(x, Q_0) \leftrightarrow F_2, F_L$), including uncertainties as in LUX.
- $Q > Q_0$: apply standard ($\alpha\alpha_s + \alpha^2$) DGLAP*.
- Results (as expected) close to LUXqed. Release coming soon.

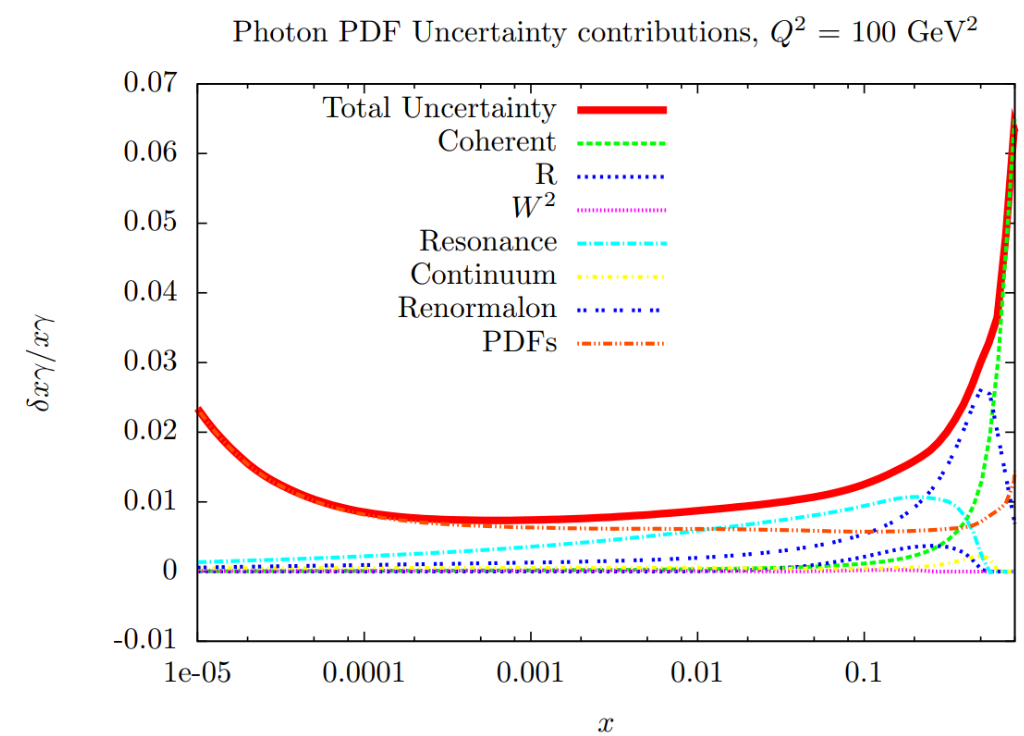




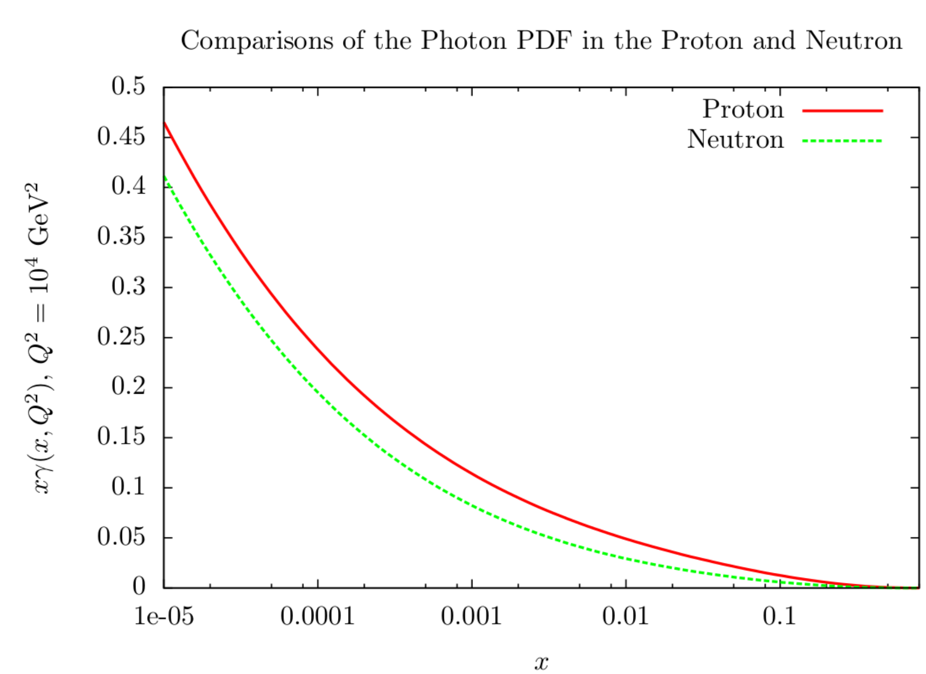
*In fact apply corrections for $\gamma_{el}(x, Q > Q_0)$, target mass corrections…

- Result close to LUXqed, but with some differences due q, g PDF input.
- **Uncertainties** at the **% level**.



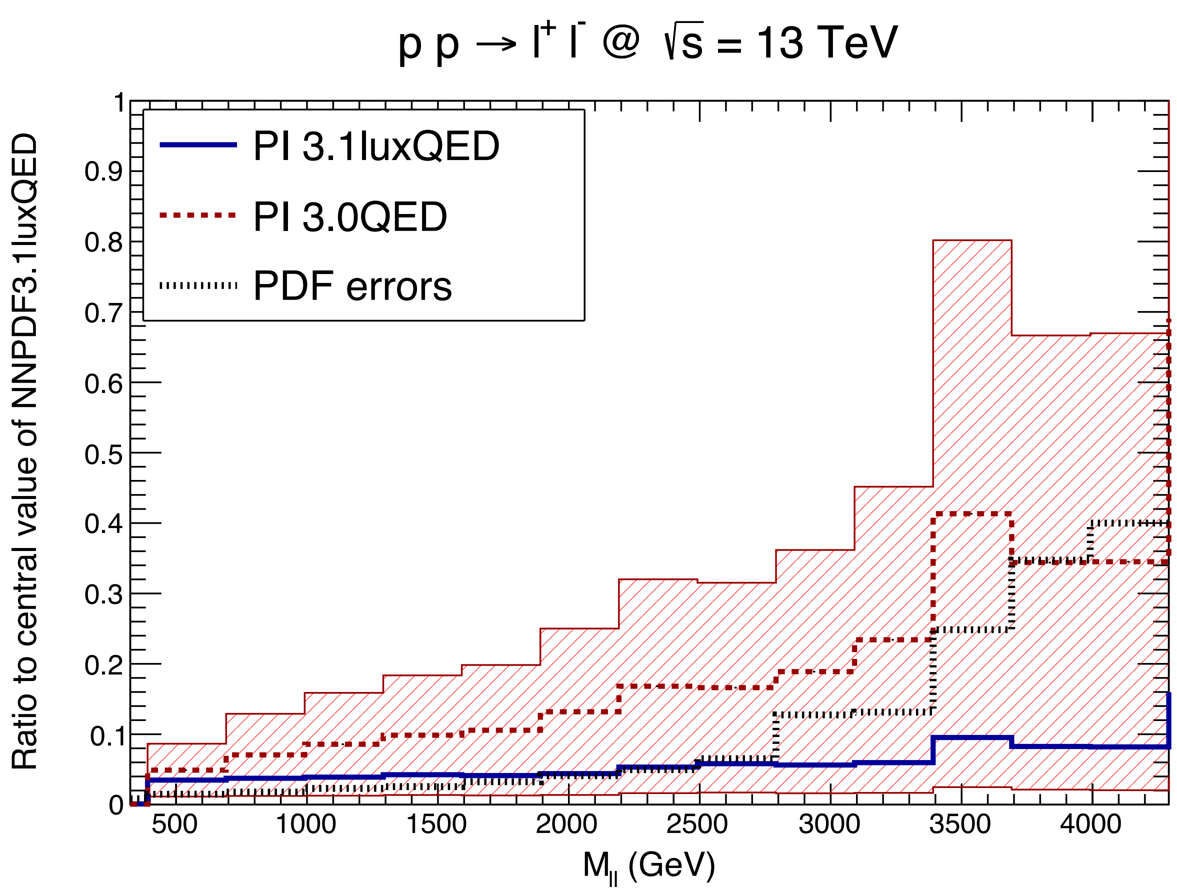


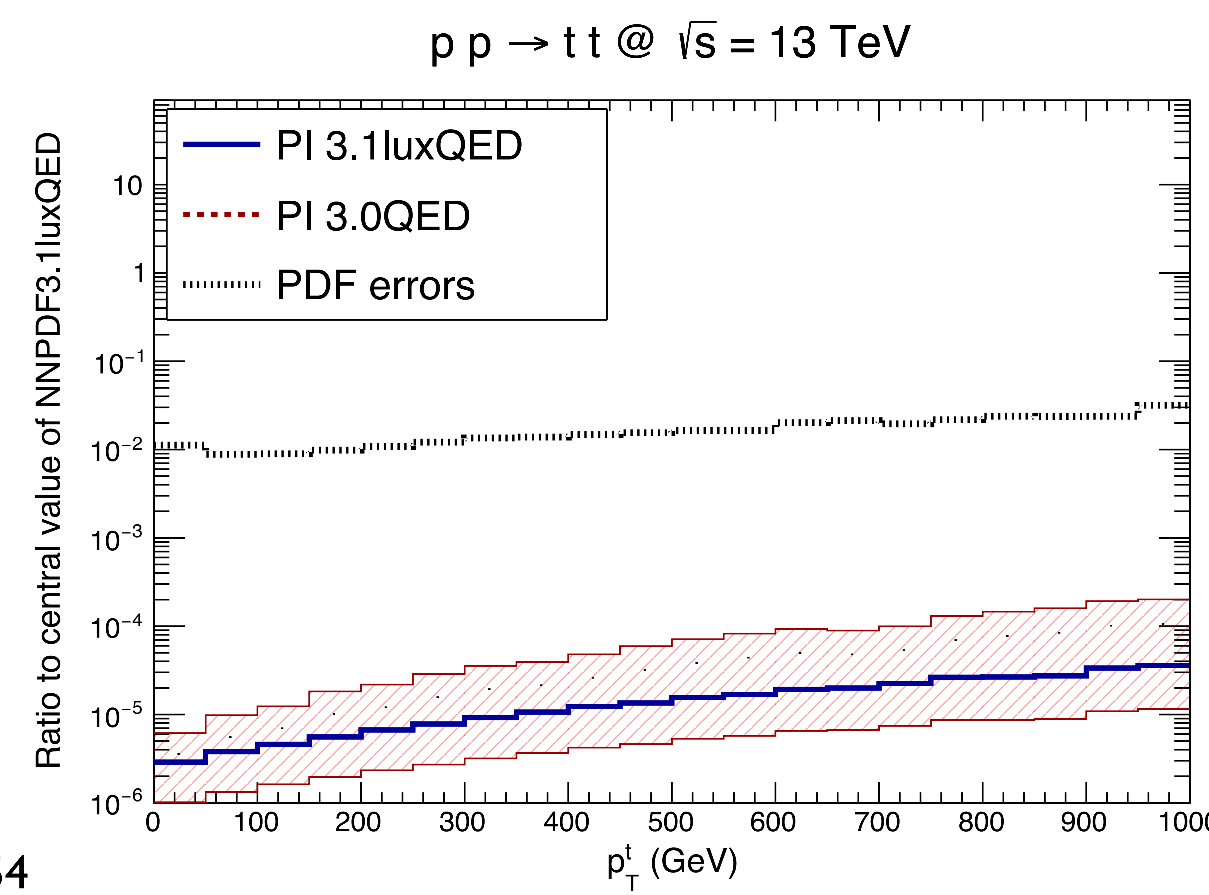
- QED effects/photon PDF for **neutron** (c.f. fixed target deuteron scattering in fit…) also included.



Photon-initiated production at the LHC

- What are the pheno. implications of this? From NNPDF study:
  - **Drell-Yan**: PI ~ a few %. Comparable to (larger than) PDF uncertainties at low (high) mass.
  - W^+W^-: as high as ~ 30 % for $M_{WW} \sim 3\,\mathrm{TeV}$, but ~ 1 % at $p_\perp^W \sim 0.3\,\mathrm{TeV}$. In both cases comparable to/larger than PDF errors.
  - $t\bar{t}$: at most at permille level, even at highest $m_{t\bar{t}}$. Well below PDF uncertainties.
  - HW: can be ~ 5% and larger than PDF uncertainties.





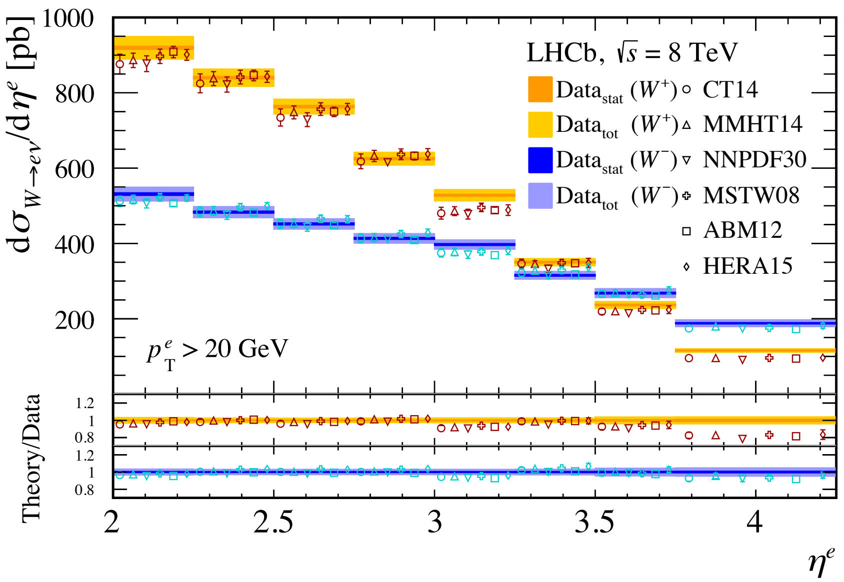
Work in Progress

Work in Progress - MMHT18 PDFs

- Work towards the **next generation** of MMHT PDFs is ongoing:

★ All **LHC Run I** data:

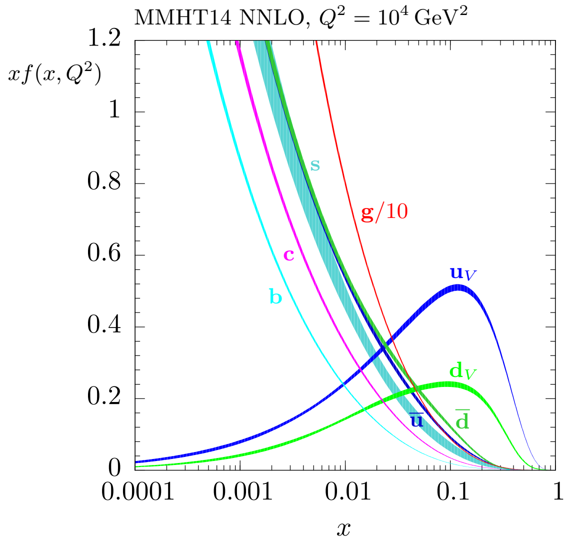
- Inclusive W, Z.
- Jets.
- Differential top.
- W+c
- …



★ Final HERA I+II inclusive and heavy flavour structure data.

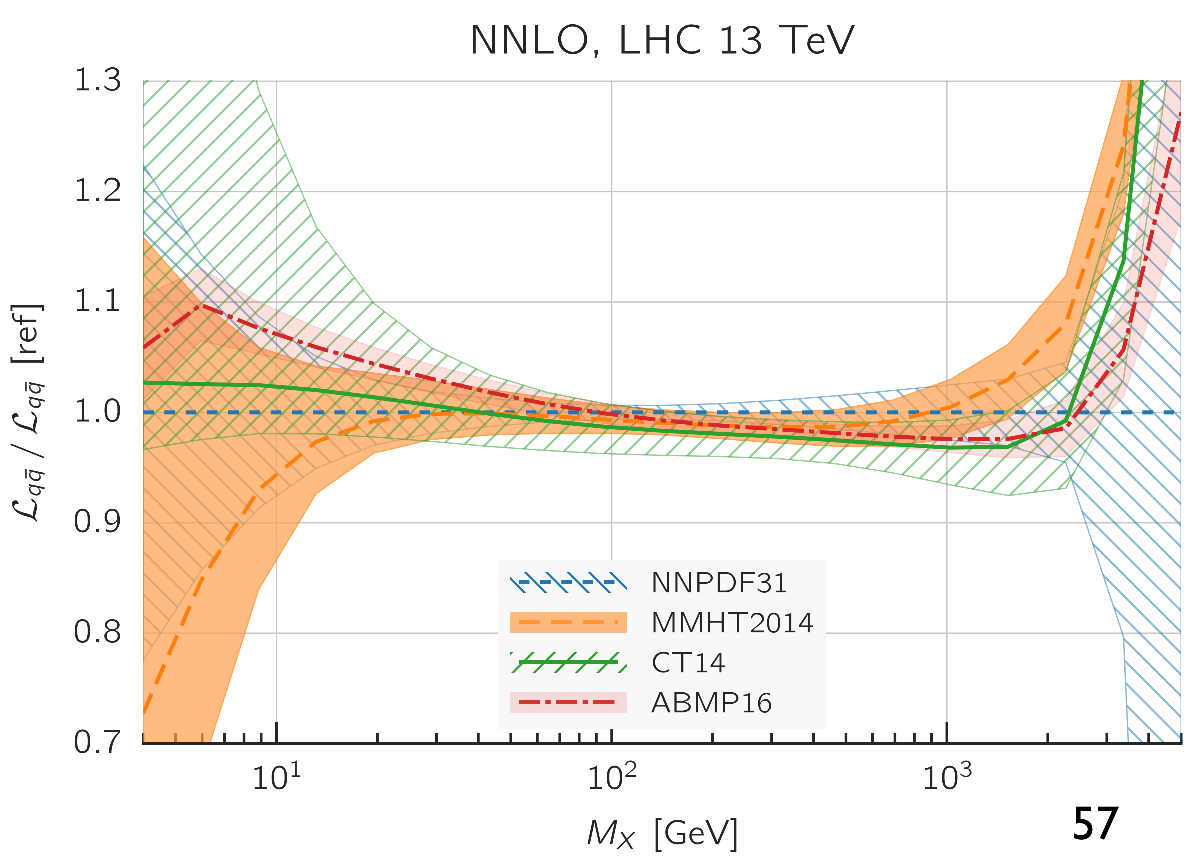
★ All with **updated theory**, in most cases NNLO (W,Z, jets, top, neutrino-induced DIS…).

★ Precise photon included by default.



Looking to Future - Ultimate PDFs

- The HL-LHC will provide a vast range of data with a direct impact on the PDFs (in particular in poorly known high x region).
- **Question:** what exactly can we expect that impact to be?
- To address this, collaborative effort to produce '**Ultimate**' PDF set ongoing: final precision that can be expected from the HL-LHC (w/ possible extension to HE-LHC).
- Produced via pseudo-data generated according to final expected kinematic coverage and experimental precision we can expect to reach.



?

Summary

• Understanding of proton structure is an essential element of the LHC precision program - encoded in the Parton Distribution Functions (PDFs). Have presented **a few selective examples**.

• High precision LHC data represents both a **opportunity and challenge** for PDF fitters.

• **Opportunity** - the highest ever precision measurements of standard candle SM processes playing significant role in PDF fit

• **Challenge** - confronting theory with such data not always simple. Delicate issues related to e.g. theoretical uncertainties and experimental systematics coming to the fore.

• Much progress being made in other areas. One example - the **photon PDF**. New theoretical insight has lead to very precise determination.

• **MMHT18** on its way - keep tuned!