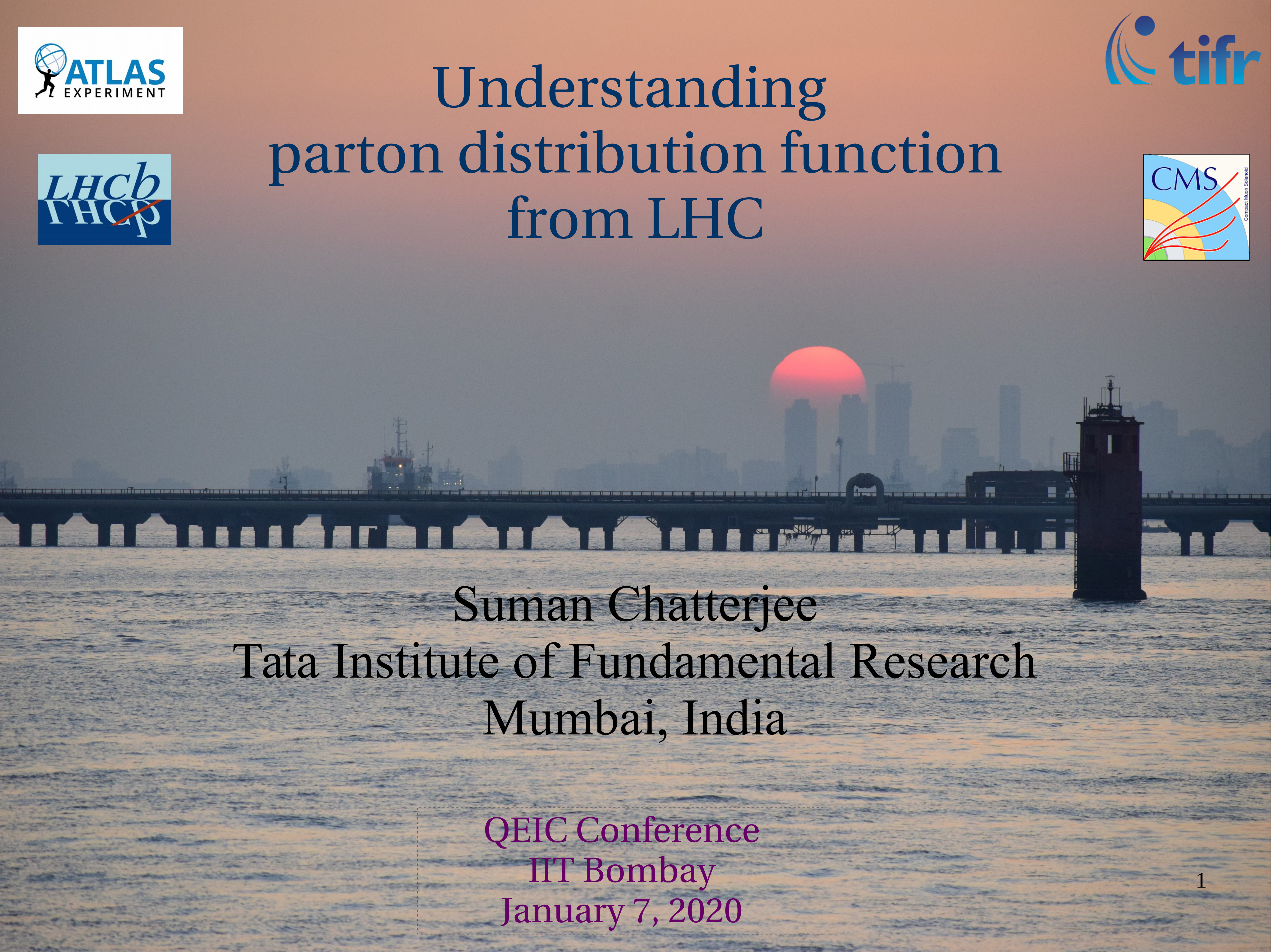


# Understanding parton distribution function from LHC

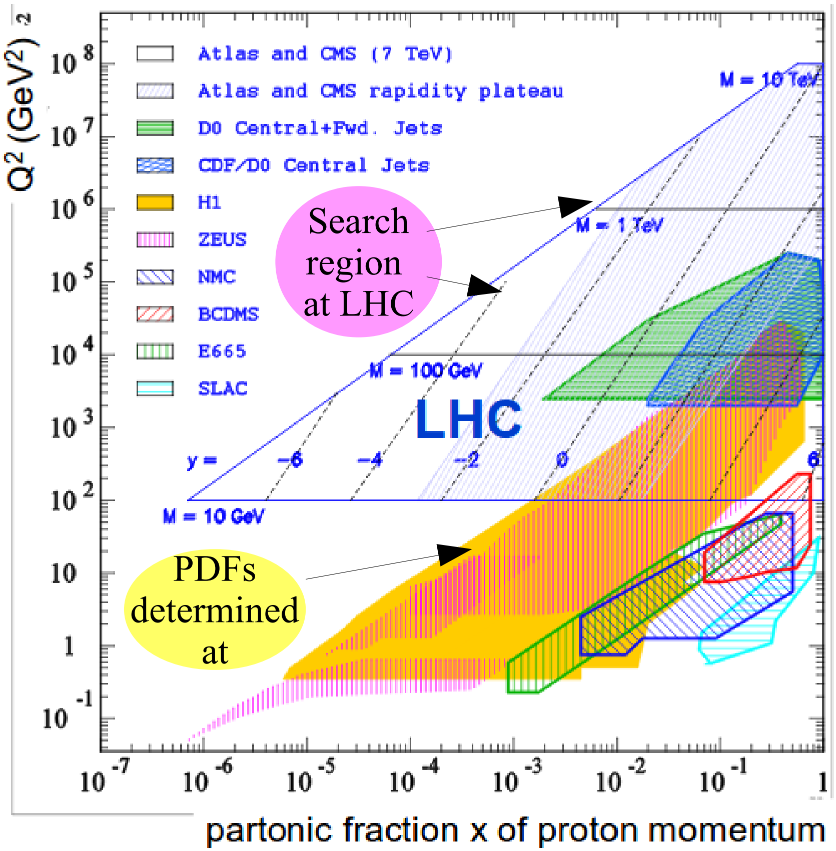
Suman Chatterjee
Tata Institute of Fundamental Research
Mumbai, India

QEIC Conference
IIT Bombay
January 7, 2020

# Parton Distribution Function

$$\underbrace{\sigma(x,Q^2)}_{hadronic\,x-sec.} = \sum_{a,b}\int_0^1 dx_1\, dx_2 \underbrace{f_a(x_1 Q^2) f_b(x_2 Q^2)}_{PDFs} \times \underbrace{\hat{\sigma}(x_1,x_2,Q^2)}_{partonic\,x-sec}$$

x = momentum fraction, Q = factorization scale




- Essential ingredient to compute cross section of any process at hadron collider
- Precision on PDF determines the accuracy of current knowledge of SM & sensitivity to beyond SM physics
- $m_W$ , $\sin\theta_W$, $m_{top}$(measured indirectly), H production by ggF <= still limited by PDF
- Can not be computed from first priciples in QCD, but evolution is governed by DGLAP equation => Needs form of PDF at some initial scale ($Q_0$)
- Experimental data helps to determine the boundary condition
- PDFs are mostly obtained fiting DIS data from HERA, neutrino scattering experiment => Mainly constrains PDF of valence quarks, total sea
- LHC offers to disentangle flavor compositions in sea PDF, determine gluon PDF, improve the PDF of valence quarks

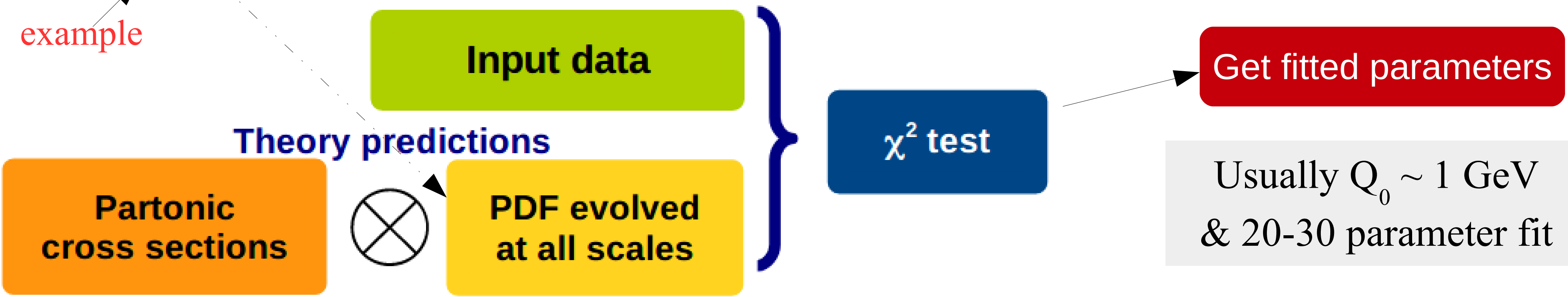
# Prescription for PDF Determination

- Parameterize by power law, and polynomial functions at an initial scale $Q_0$

$$Ax^B (1-x)^C (1+Dx+Ex^2)$$

Behavior governed by B at low x, C at high x, D, E offer additional flexibility

example




Usually $Q_0 \sim 1$ GeV & 20-30 parameter fit

- Different PDF sets available in market (CTEQ now CT, NNPDF, MMHT, HERAPDF, ABMP,...)
  Difference in =>

i) Parametric form
ii) Input data (H1, ZEUS, neutrino scattering, Tevatron, LHC)
iii) $m_b$, $m_c$, $m_{top}$, $\alpha_S$
iv) Starting scale $Q_0$
v) Flavor number scheme : fixed flavor number (FFNS) or variable flavor number (VFNS)
vi) DGLAP at LO / NLO / NNLO
....

Constraints:

$$\int (u(x) - \bar{u}(x))\, dx = 2$$
$$\int (d(x) - \bar{d}(x))\, dx = 1$$

# of active flavors : 3 (charm is massive) / 4 (bottom is massive)

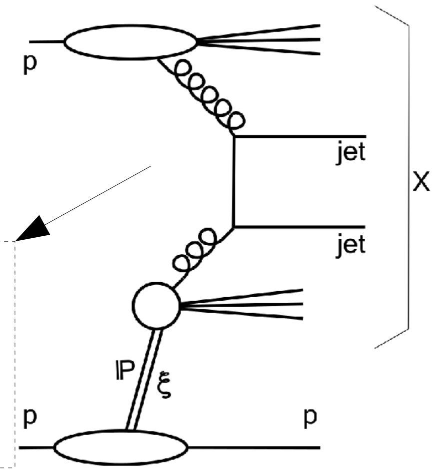
Charm (bottom) is massless at $Q_0$
Transition to FFNS at $m_c$ ( $m_b$ )

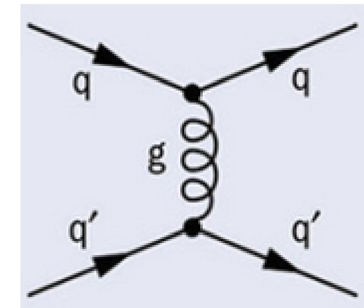
# Physics processes @ LHC & their impacts on PDF

Covered by ATLAS & CMS

**Not a complete set!!**

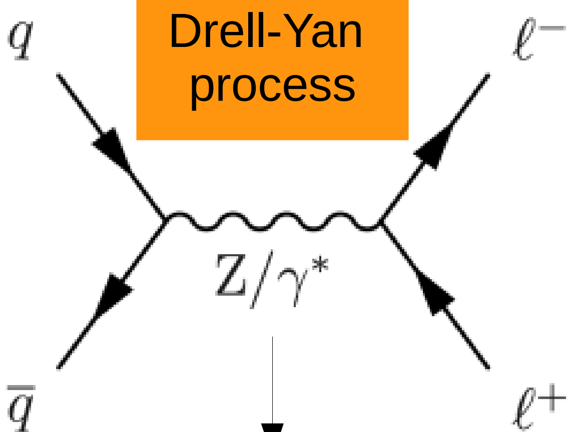
Inclusive jets



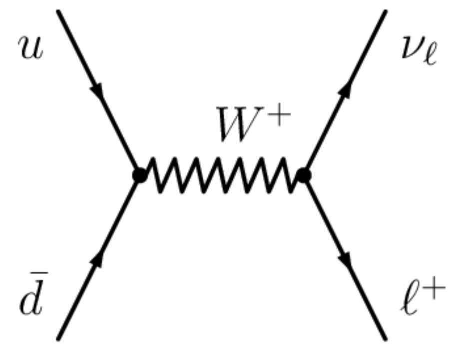




Valence quark PDF at high x
gluon PDF at medium-high x

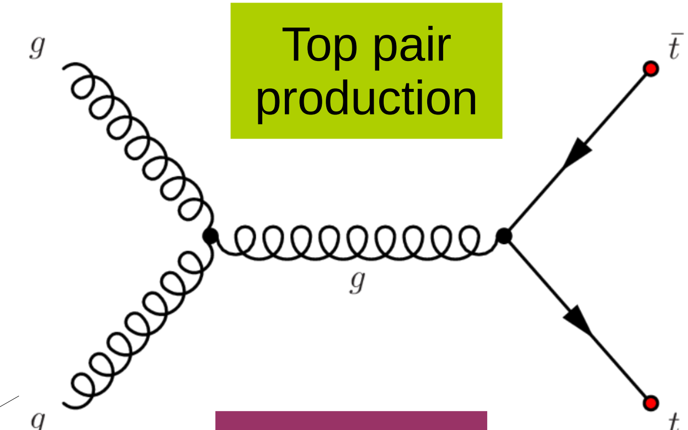
Drell-Yan process




Constraints on valence quark PDF

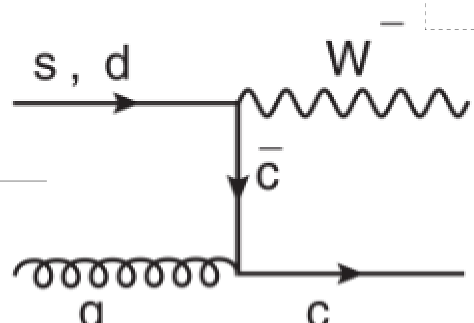



Top pair production



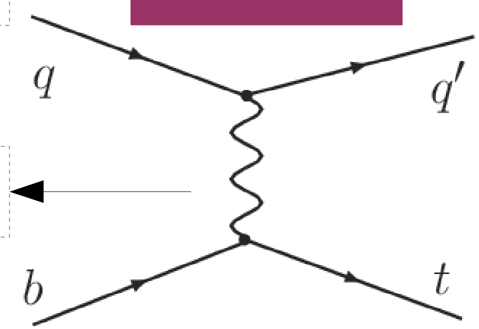

Gluon PDF at high x

W + c production




Strange quark PDF medium x

Single top production



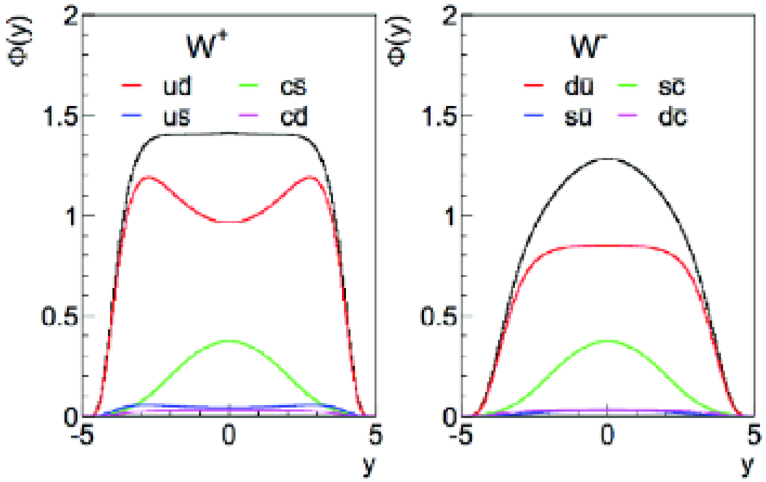

u, d, b PDF

# W charge asymmetry @ √s = 8 TeV

- p p ⟶ W ⟶ $\mu\ \nu_\mu$ (Charge of muon = charge of W)

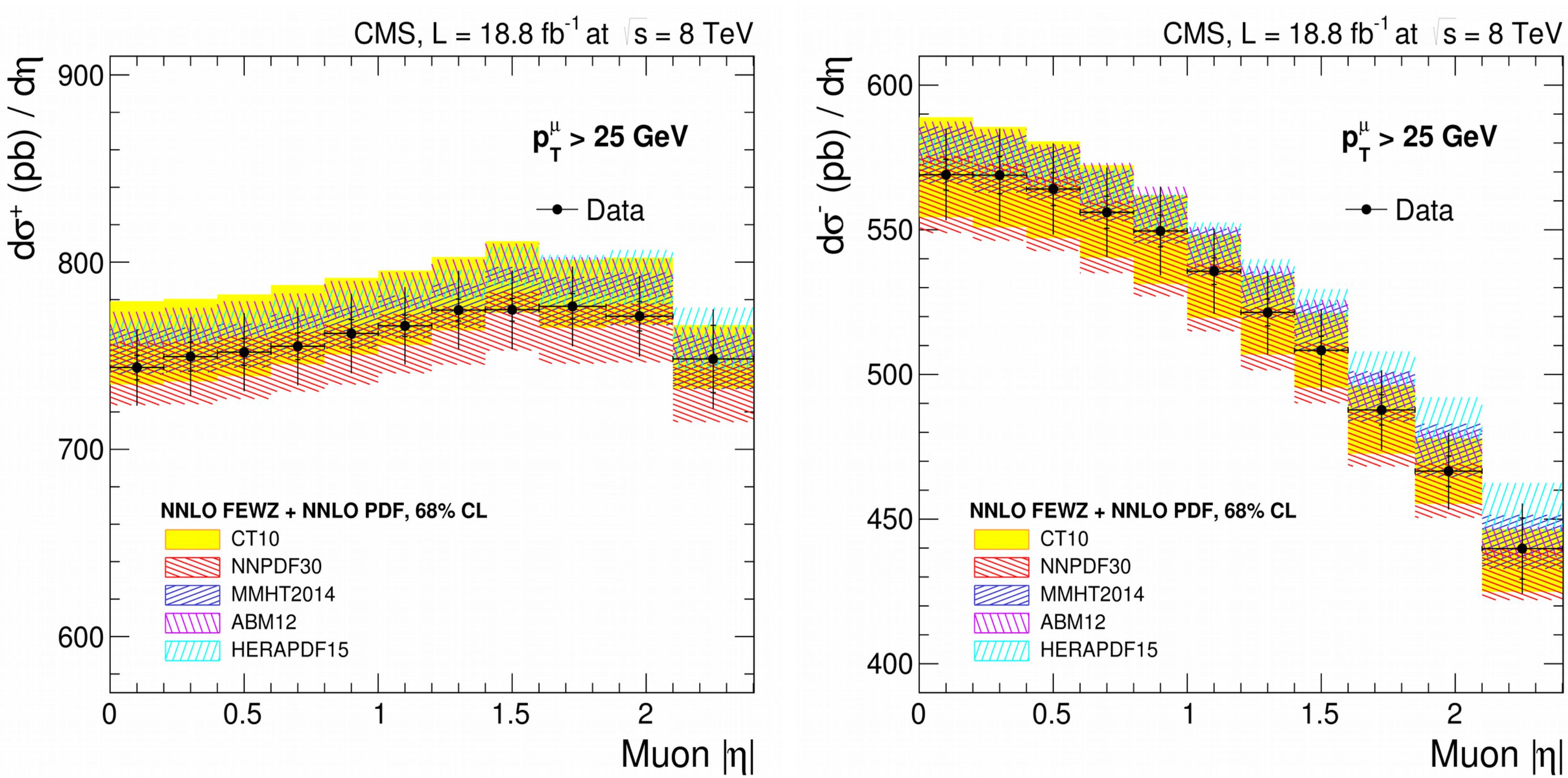
  Dominating contributions from
  $u\bar{d} \to W^+$ & $d\bar{u} \to W^-$

  $x_q = (M_W / \sqrt{s}) \exp(+y_W)$
  $x_{qbar} = (M_W / \sqrt{s}) \exp(-y_W)$




- Measurements are compared to NNLO predictions

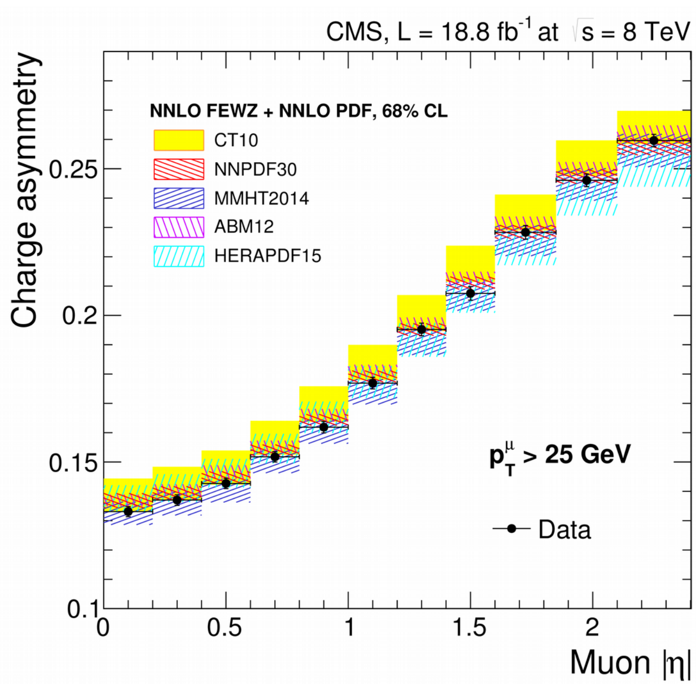
CMS-SMP-14-022




NNLO calculations describe the diffenetial distributions well

# W charge asymmetry @ $\sqrt{s}$ = 8 TeV

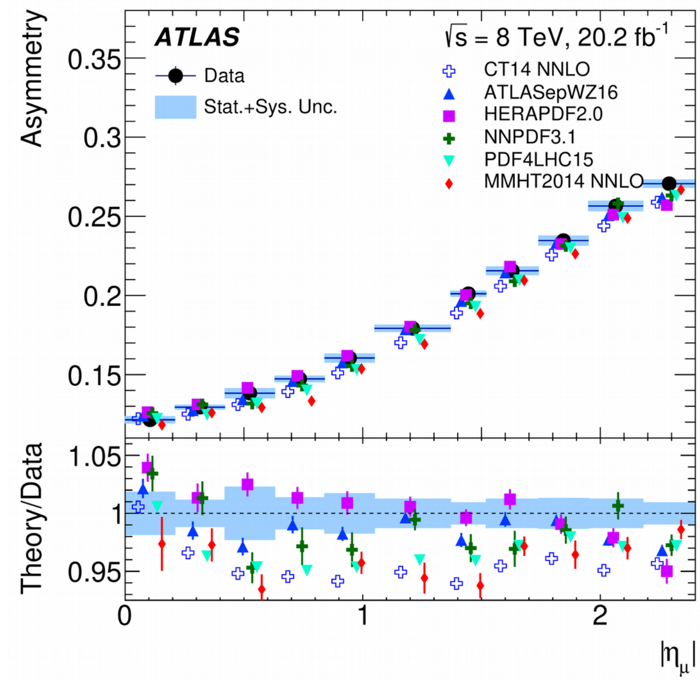
- Asymmetry = $(\sigma^{W+} - \sigma^{W-}) / (\sigma^{W+} + \sigma^{W-}) \sim (u_v - d_v)/(u_v + d_v + sea)$
- Measurements are compared to NNLO predictions

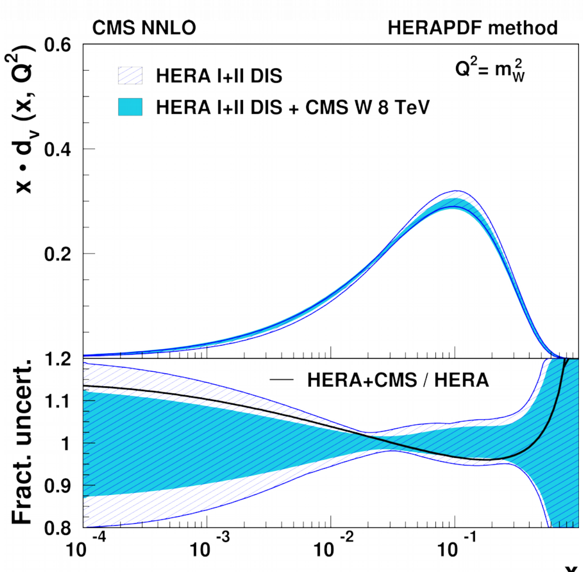


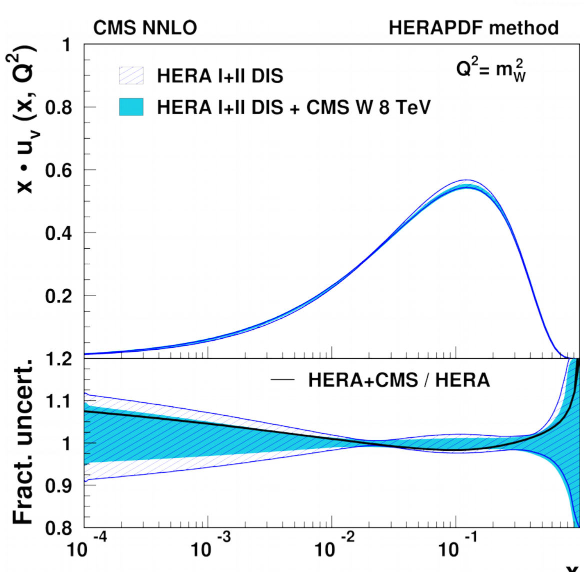

Different cross section calculators used in ATLAS and CMS

In CMS, predictions from different PDF sets are within uncertainty from data

Significant difference between prediction from different PDF sets in ATLAS



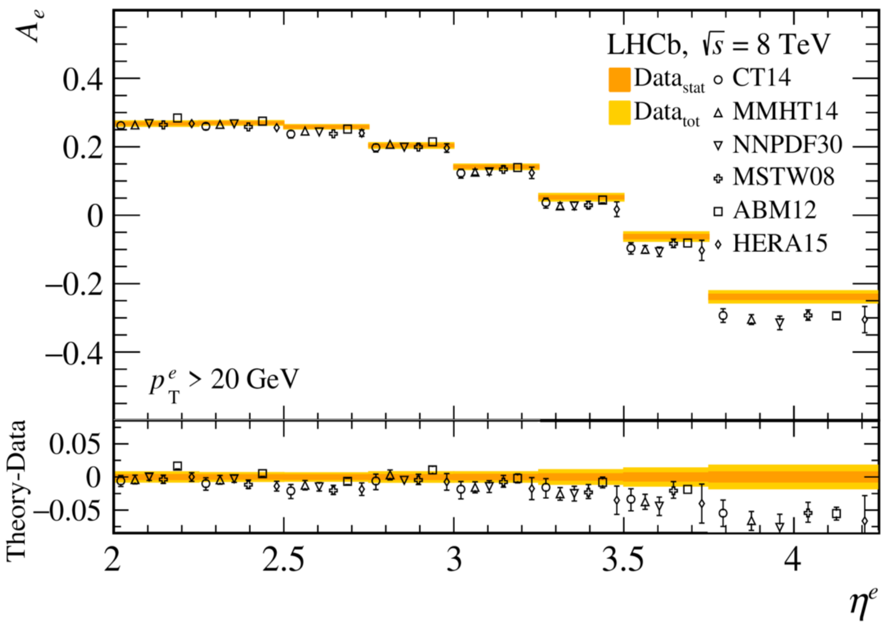

arXiv 1904.05631

Significant reduction of PDF uncertainty for valence quarks at low x

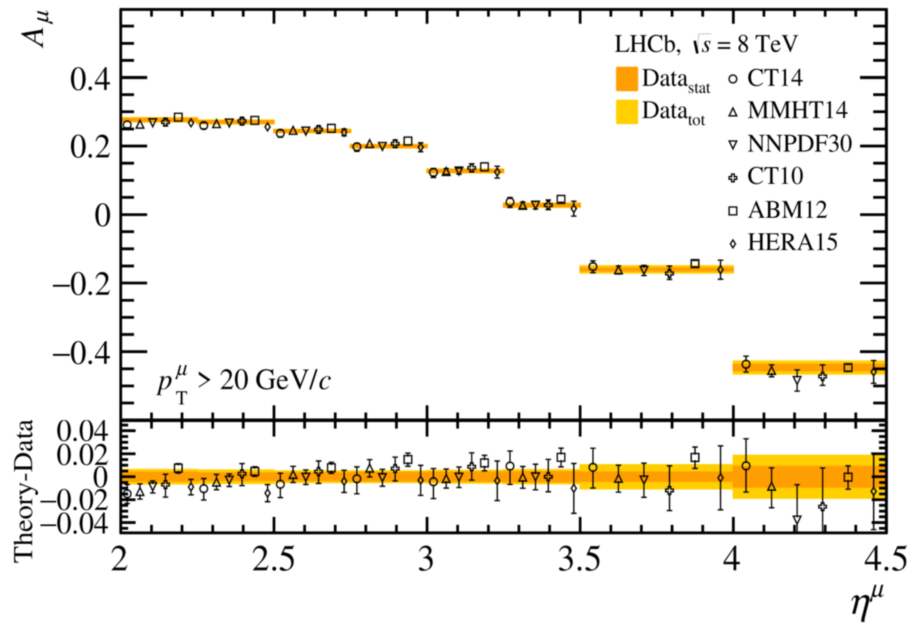
CMS-SMP-14-022

# W charge asymmetry @ $\sqrt{s}$ = 8 TeV

- Complementary measurements by LHCb in forward region (very high x)
- Measurements are compared to NNLO predictions




LHCb-PAPER-2016-024

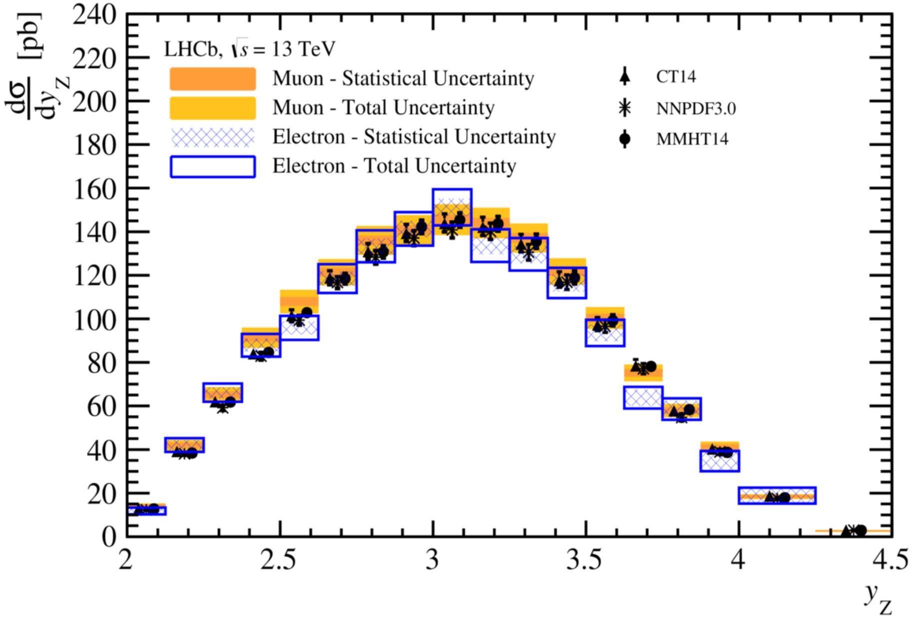



LHCb-PAPER-2015-049

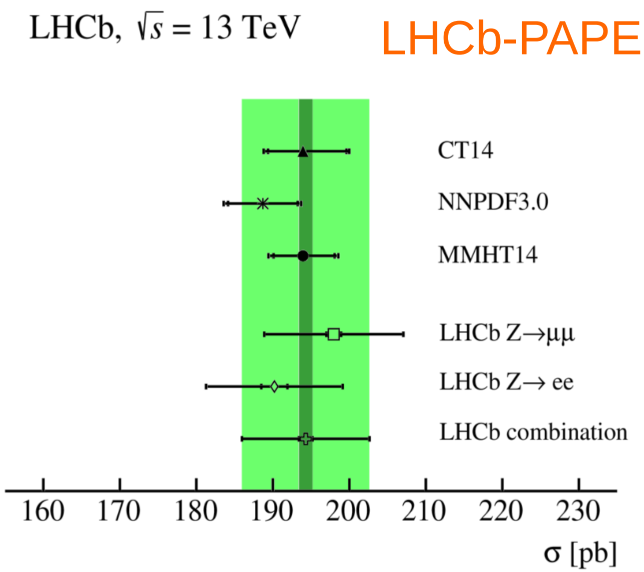
- With increase in η, prediction becomes gradually smaller than data
- Difference between theory to data comparisons between electron and muon channels
- Also can be used to check isospin symmetry for sea quarks ( $\bar{u} - \bar{d}$ )

# Z production @ $\sqrt{s}$ = 8 TeV

- Prediction for Z production cross section using different PDFs are consistent with data



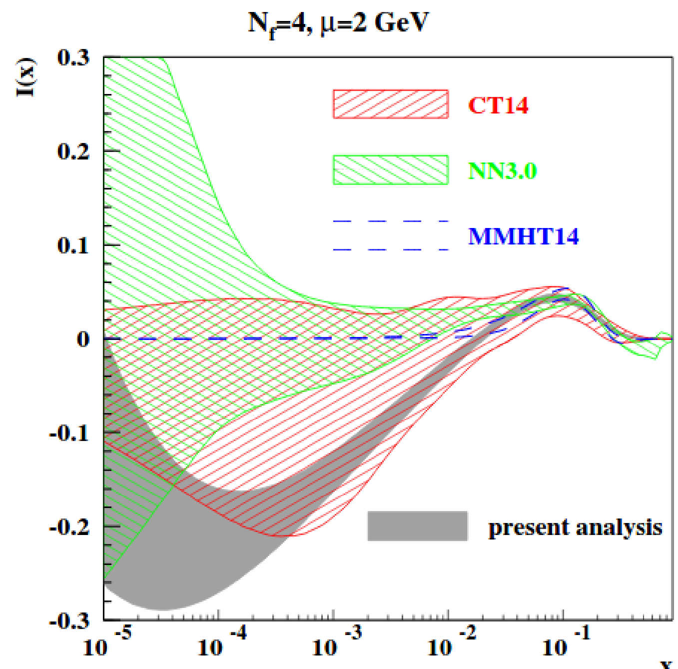

LHCb-PAPER-2016-021




- Analysis by ABMP group using LHCb data for Z, W production in PDF fits finds more $\bar{u}$ than $\bar{d}$, specially at low x in proton

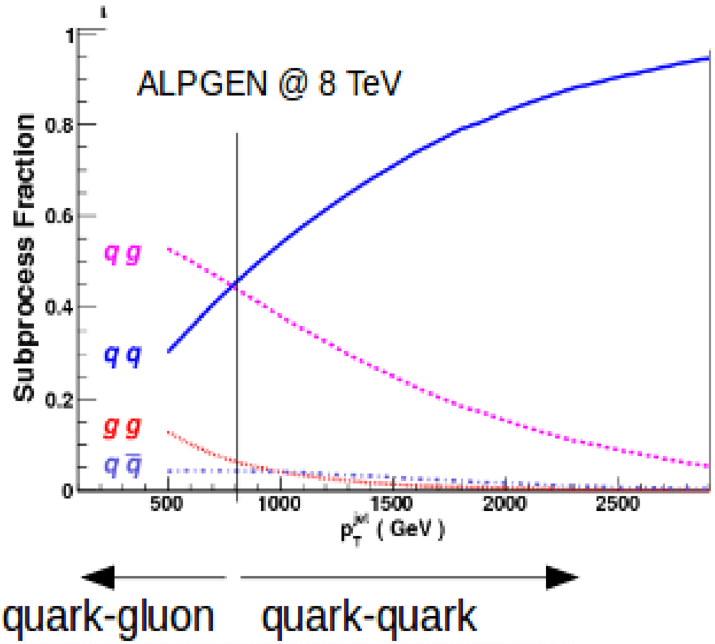
- Breaking of isospin symmetry in sea?

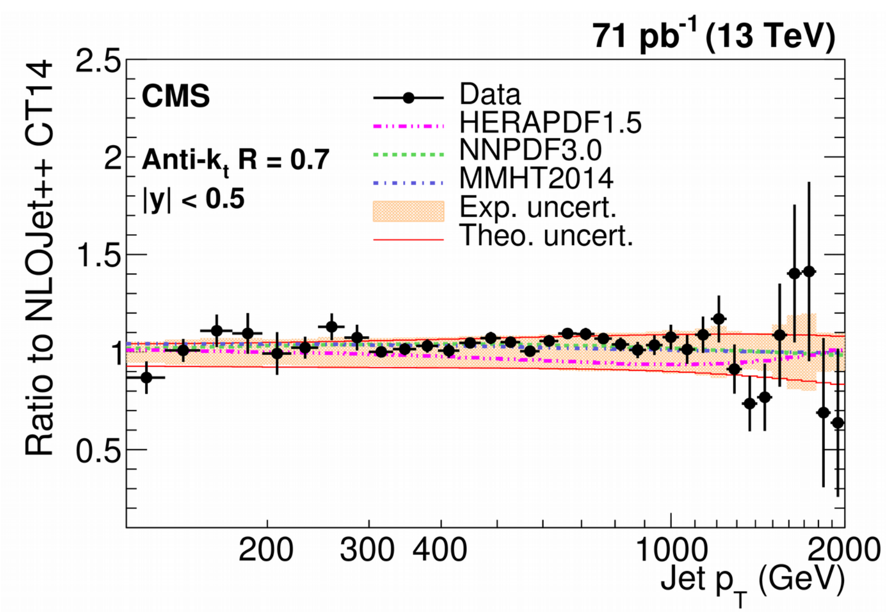
$I(x) = x\,(\bar{d} - \bar{u})$

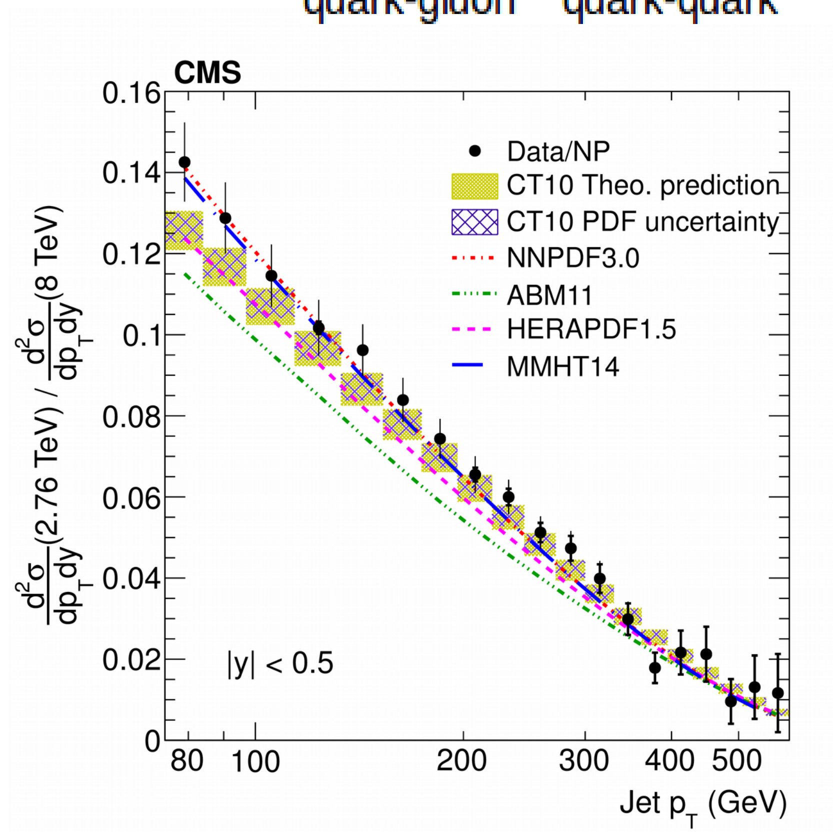


# Inclusive Jets in CMS @ $\sqrt{s}$ = 13TeV

- p p ► jet + X
- Sensitive to : gluon PDF at small and medium x as dominating sub-processes are gg & qg scattering valence quark PDF at large x (via qq scattering)
- theory prediction at NLO corrected with nonperturbative & EW effects
- Ratio of inclusive jet cross section at two different $\sqrt{s}$ is also sensitive to PDF



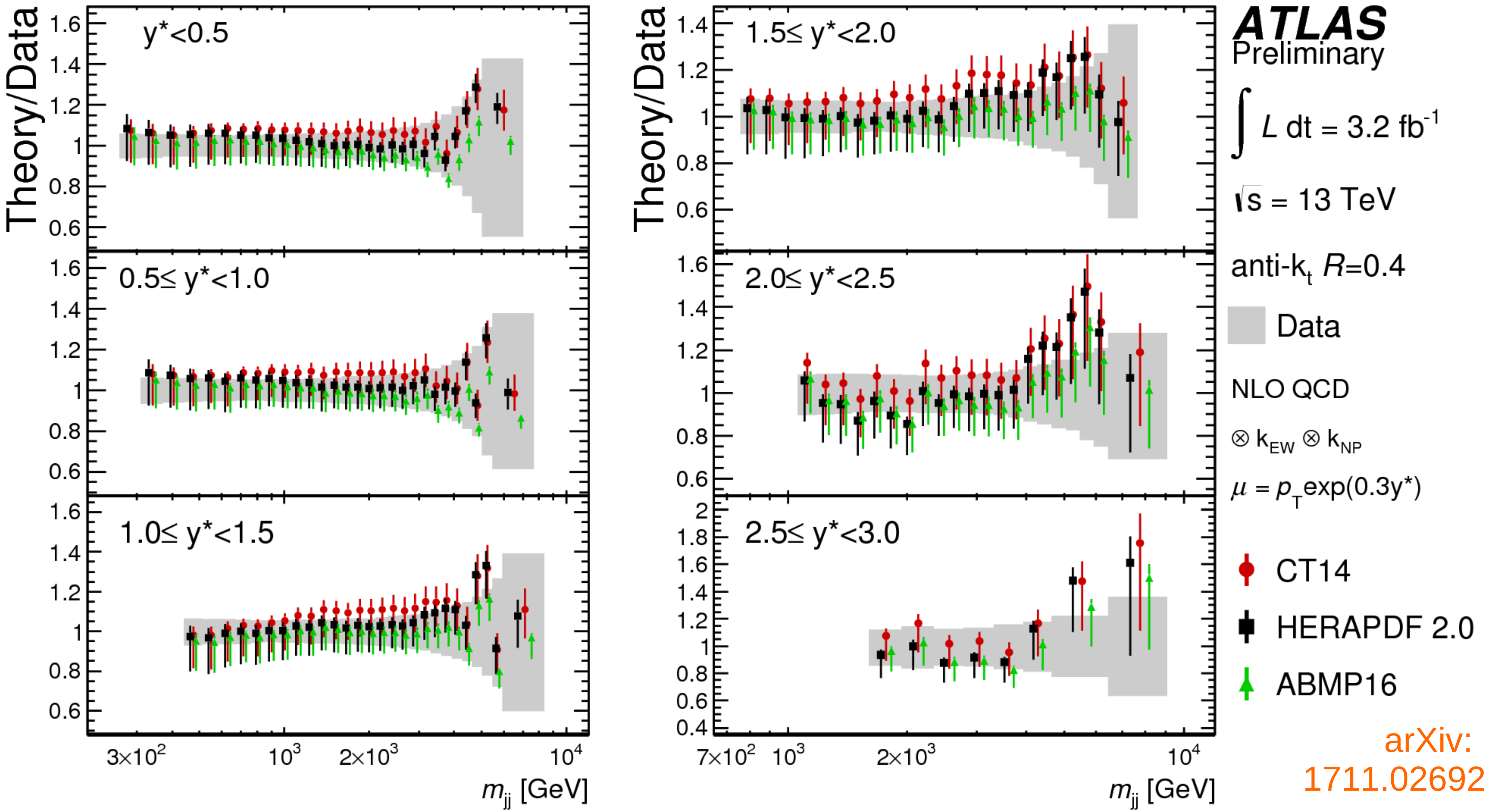

CMS-SMP-15-007

50 ns data in 2015

# DiJet Mass in ATLAS @ √s = 13TeV

- Di-AK4jet mass compared to NLO prediction (with NP and EW correction) in bins of 'half of rapidity difference of two jets' (y*)

25 ns data in 2015

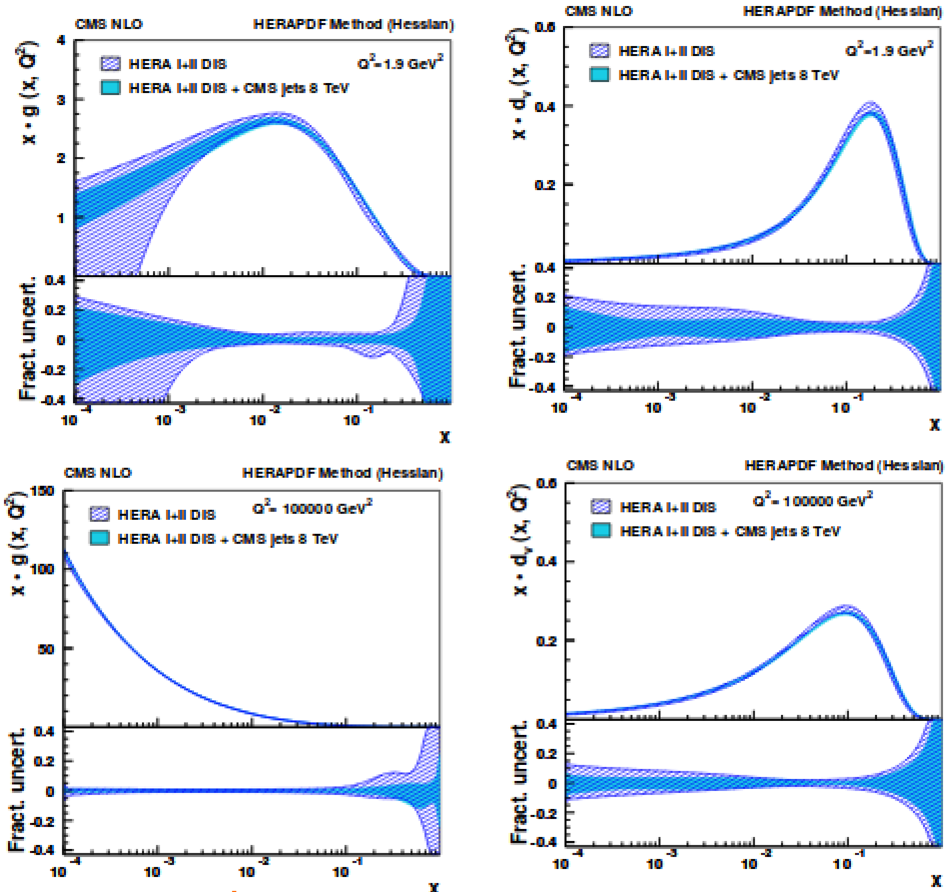


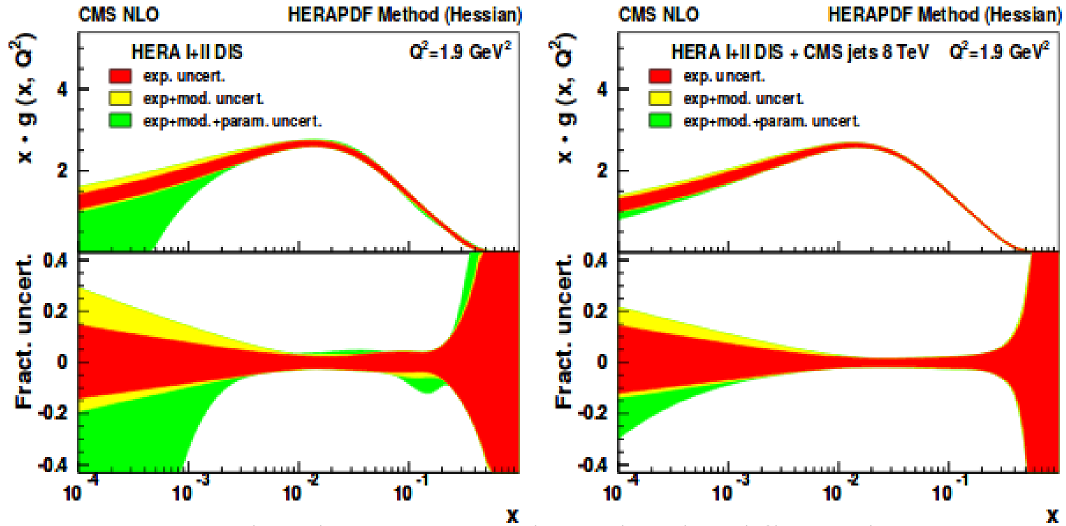

arXiv: 1711.02692



- Sensitivity to PDF for dijet mass is quite similar to inclusive jets

# Impact of Inclusive Jet Cross-Section on PDF

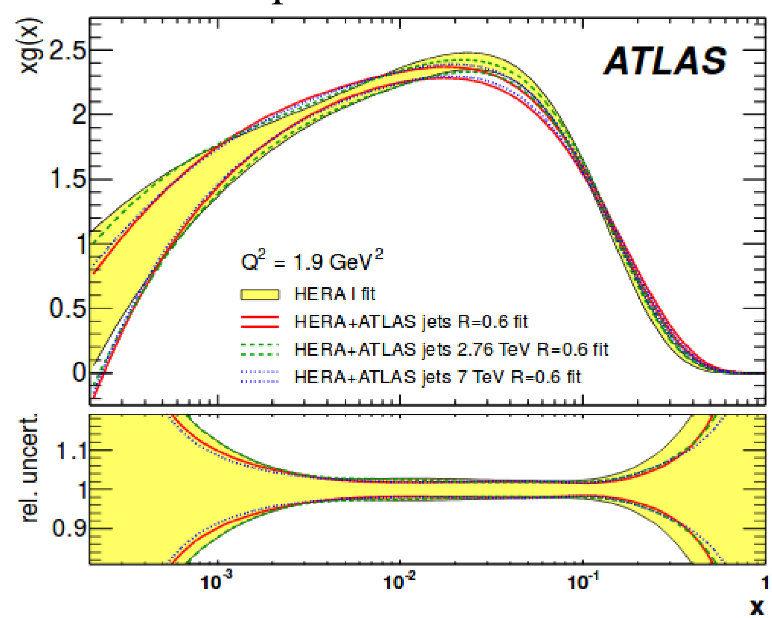



Parameterisation uncertainty is significantly reduced by the inclusion of CMS jet data

Improvement in the description both for gluon (high x)& valence quarks (low x) within proton

arXiv: 1609.05331




arXiv: 1304.4739

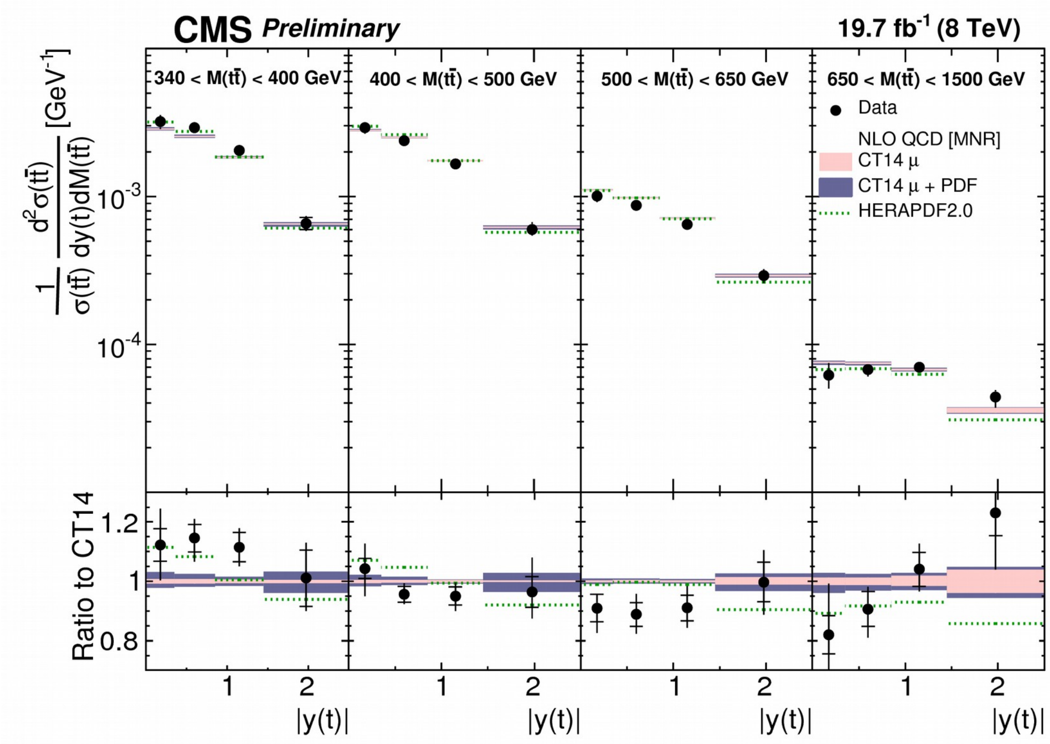
ATLAS uses the potential of inclusive jet data at both √s=2.76 TeV & √s=7 TeV (with correlation of systematics) to constrain gluon PDF

Gluon PDF is harder after being constrained by ATLAS data

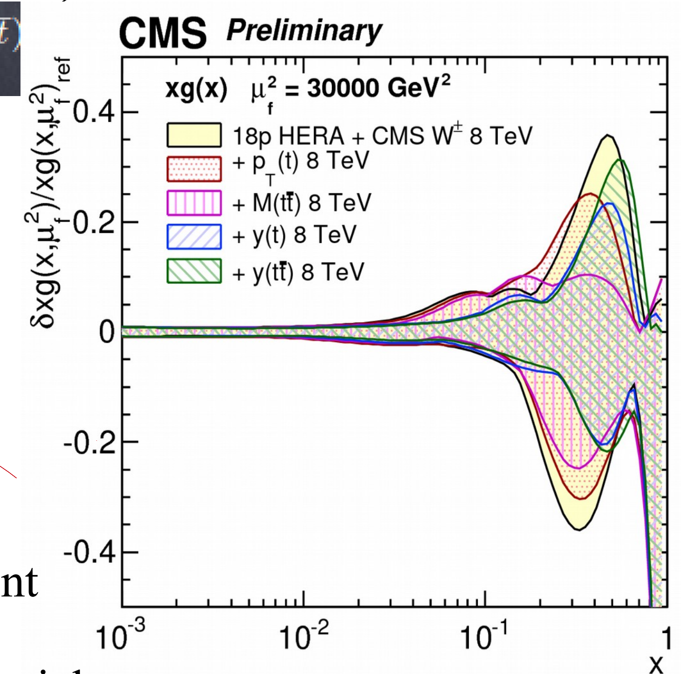
# Double Differential t-tbar Cross-Section at √s = 8 TeV

CMS-TOP-14-013

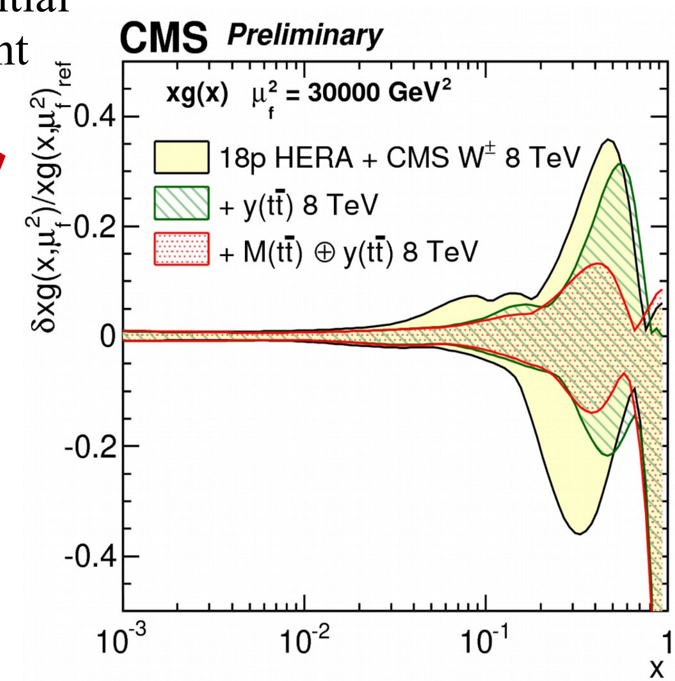
- Dominant sub-process : gg ► t tbar (t & s channel) : sensitive to gluon PDF at large x $x_{1,2} = \frac{M(t\bar{t})}{\sqrt{s}} e^{\pm y(t\bar{t})}$



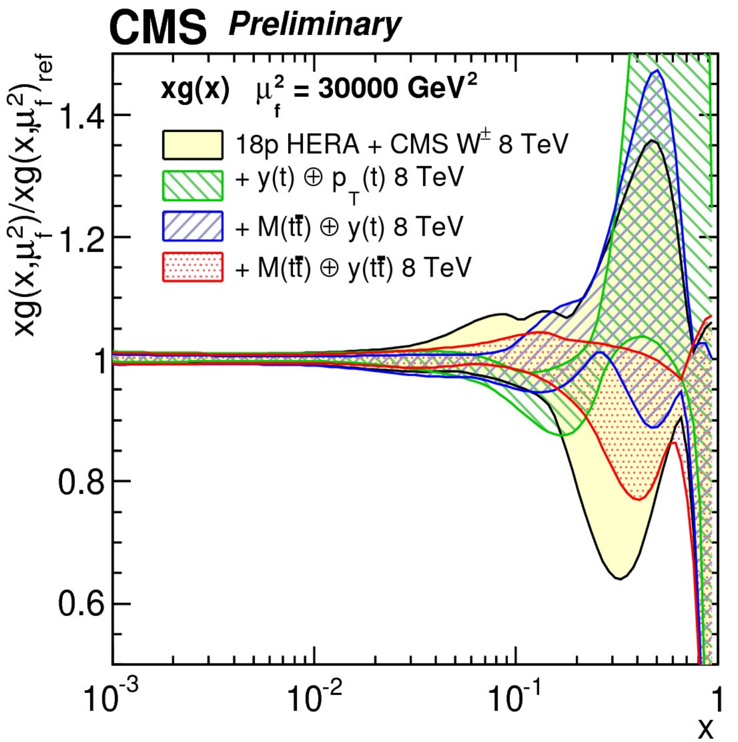

Measurement performed in dilepton channel
(e-μ+/e+μ-)
to get pure t-tbar sample




Improvement over single differential measurement



# Double Differential t-tbar Cross-Section at √s = 8 TeV

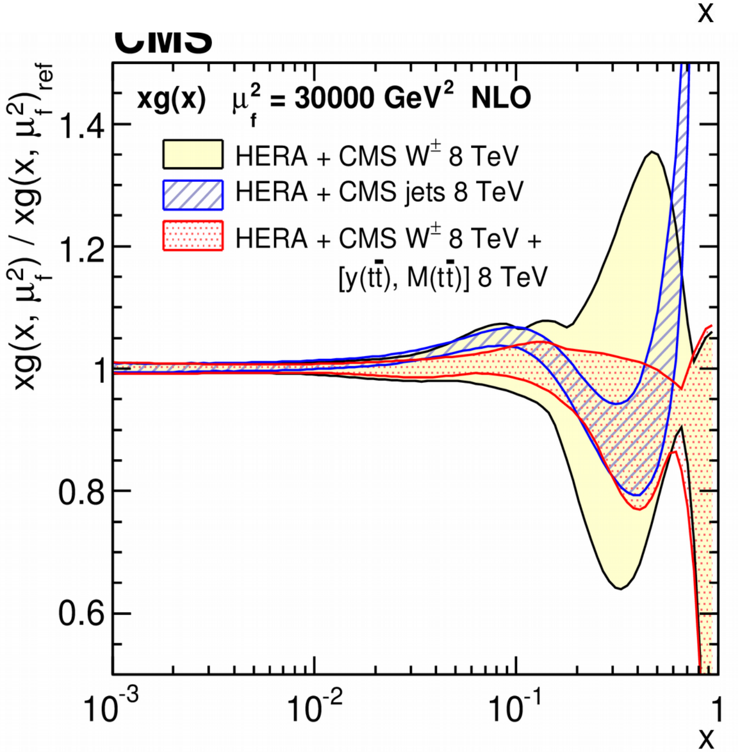



PDF fit is performed using xFitter with NLO calculation for cross section

M(tt̄),y (tt̄) combination imposes the most stringent constraint on gluon PDF

Total $\chi^2$/ndf (fitting HERA data, CMS W charge asymmetry data, CMS t-tbar) is very similar among 3 pairs of variables from t-tbar

Significant mprovement in gluon PDF at high x !

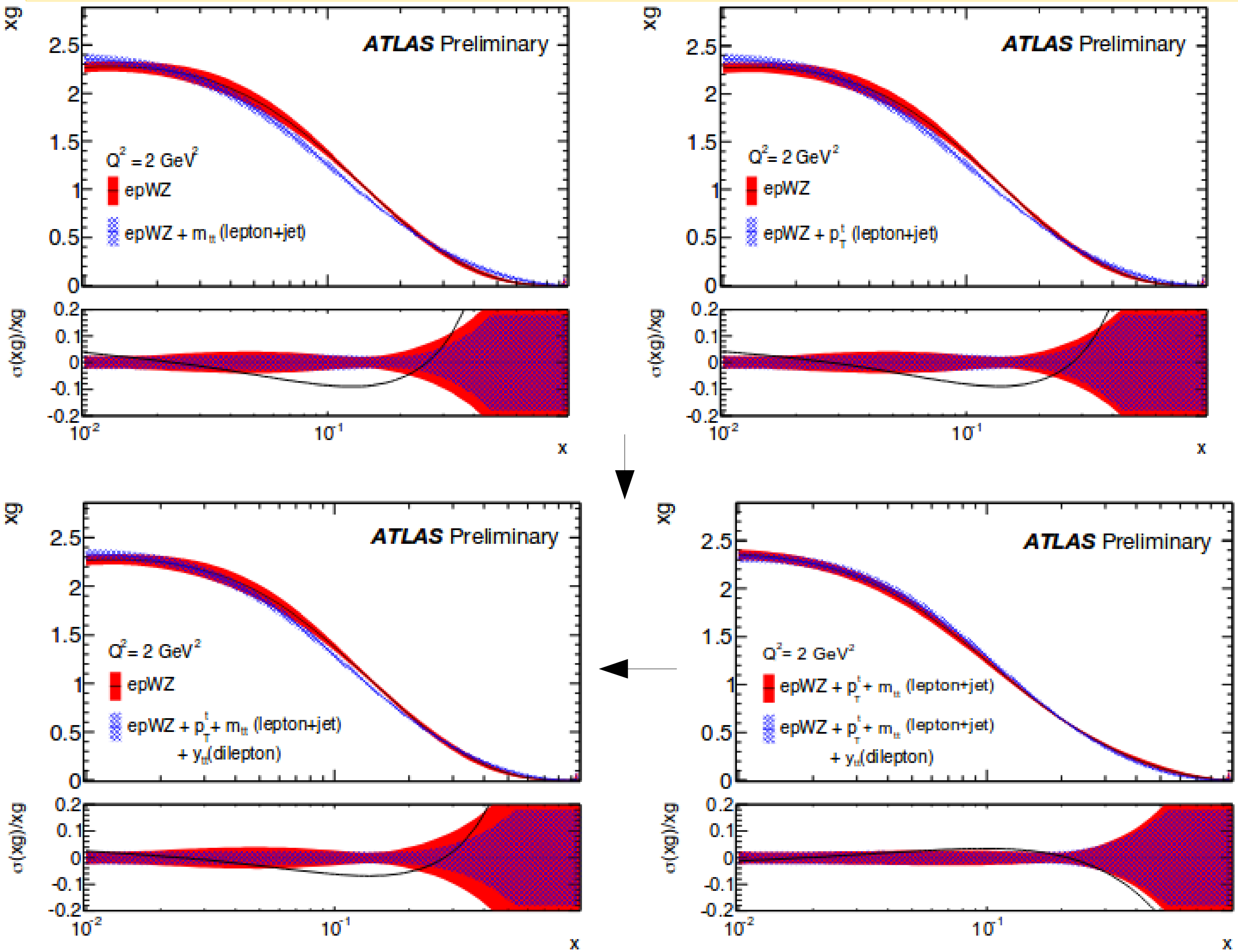



Sensitivity of gluon PDF to W asymmetry + ttbar data is very similar to the sensitvity to inclusive jets data

Only at very high x, inclusive jet data predicts much harder gluon PDF than ttbar data

# Impact Double Differential t-tbar Cross-Section (ATLAS Data) on PDF

$m_{tt}$ and $p_T^t$ have been fitted from semileptonic channel
For dileptonic channel, $y_{tt}$ is chosen to add rapidity info ( $m_{tt}$ also gives a good fit)




Statistical correlation within each spectrum & different spectra are evaluated using Bootstrap method

Systematic uncertainties are taken fully correlated between spectra

ATL-PHYS-PUB-2018-017

Both $p_T$ and $m_{tt}$ spectra makes the gluon harder

$m_{tt}$ seems to have the biggest impact to constrain gluon PDF

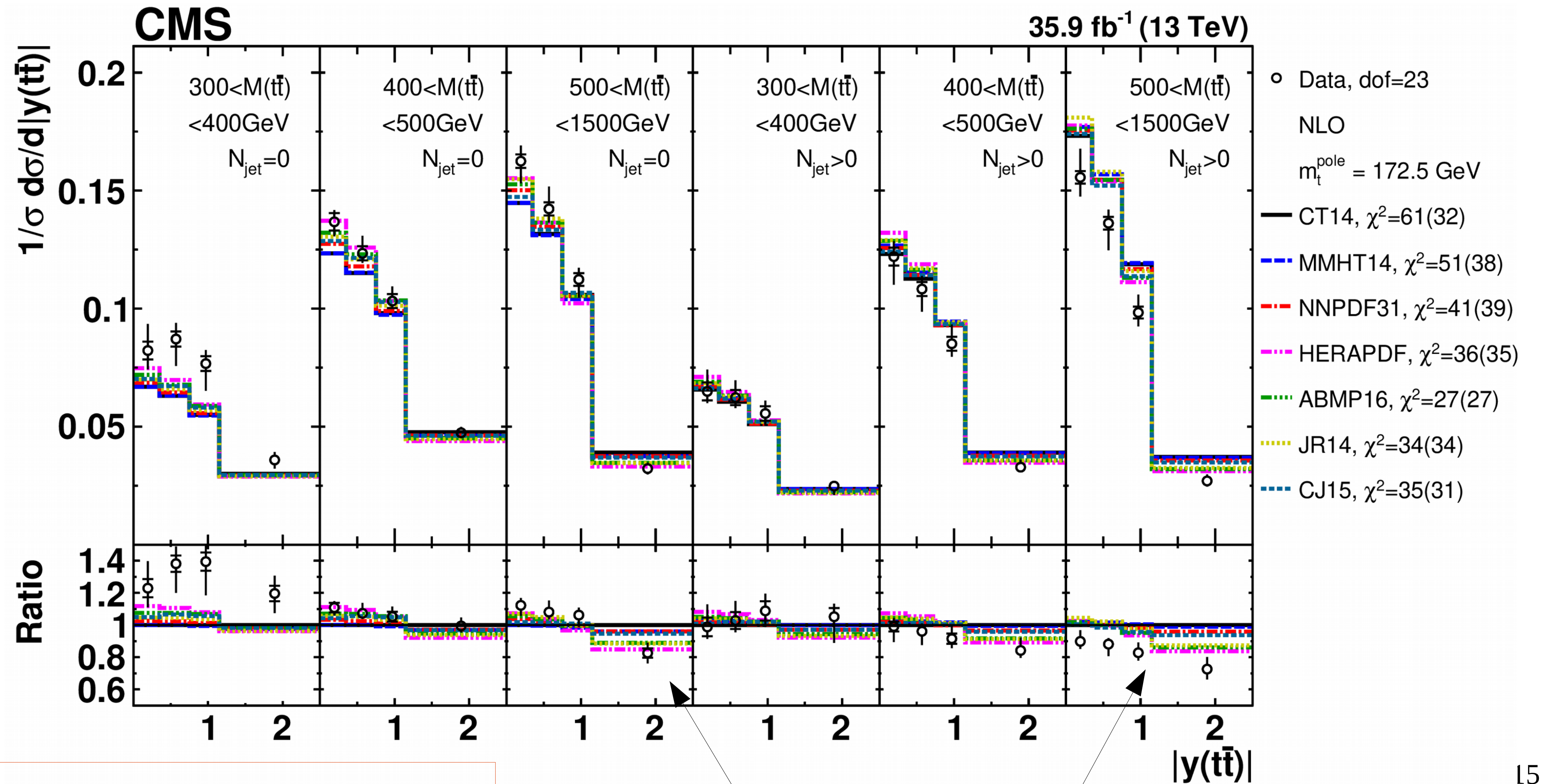
Dilepton data softens the gluon PDF w.r.t lepton+jets data



ATLAS t-tbar measurement reduces gluon PDF uncertainty significantly at large x and supports mildly harder gluon

# Double Differential t-tbar Cross-Section at √s = 13 TeV

- Differential cross sections are measured as a function of y(tt), m(tt) and # of additional jets
- Target to extract $\alpha_S$, top quark pole mass, and gluon PDF simultaneously, and also their correlations




Reminder : large y => large x!

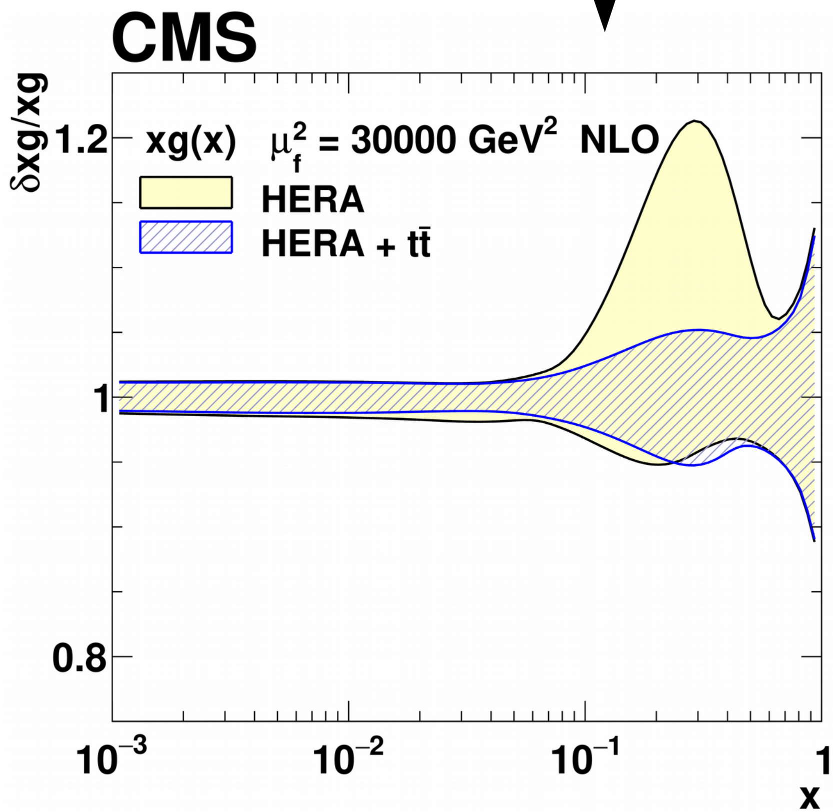
Measurement is sensitive to PDF

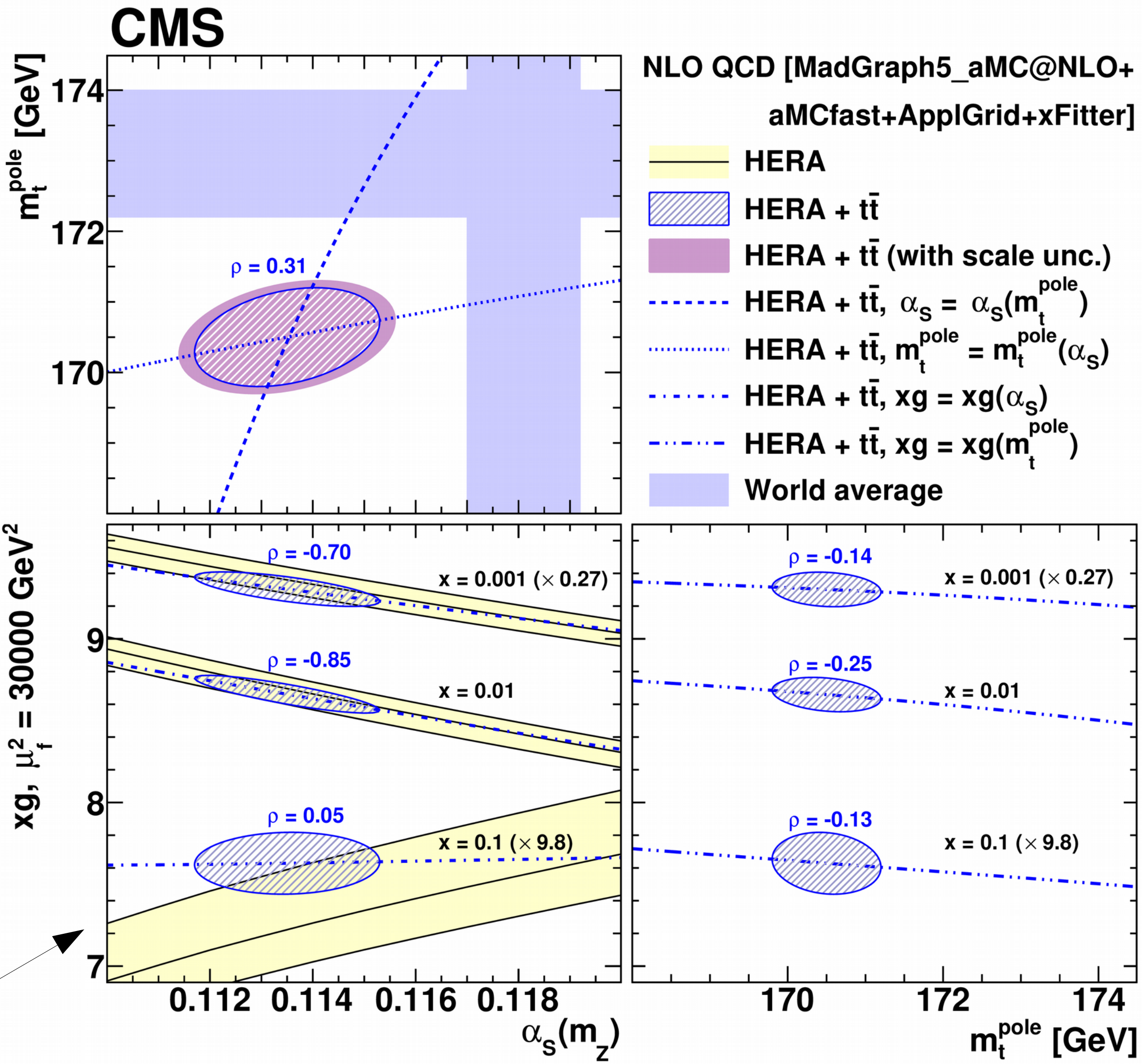
CMS-TOP-18-004

# Double Differential t-tbar Cross-Section at $\sqrt{s}$ = 13 TeV

- Extracting $\alpha_S$, top quark pole mass, and gluon PDF simultaneously, and their correlations

Reduced uncertainty in gluon PDF at large x




Correlation with $\alpha_S$ is reduced, specially at large x



CMS-TOP-18-004

# Cross-section Ratio of t-tbar to Z at $\sqrt{s}$ = 13 TeV

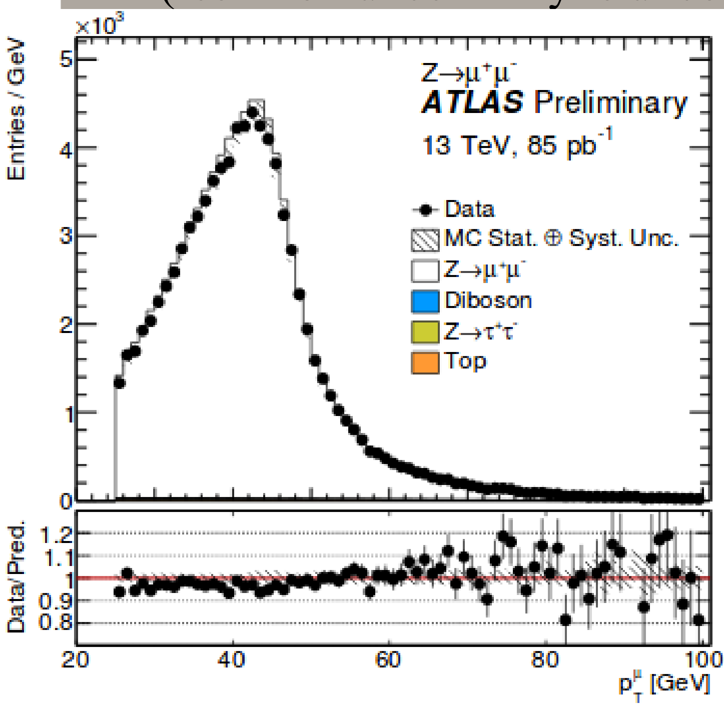
- t-tbar cross section measured with opposite sign e-μ final states (the most precise measurement in dilepton channel) ATLAS-CONF-2015-033
- Cross section of Z production and branching ratio to ee & μμ

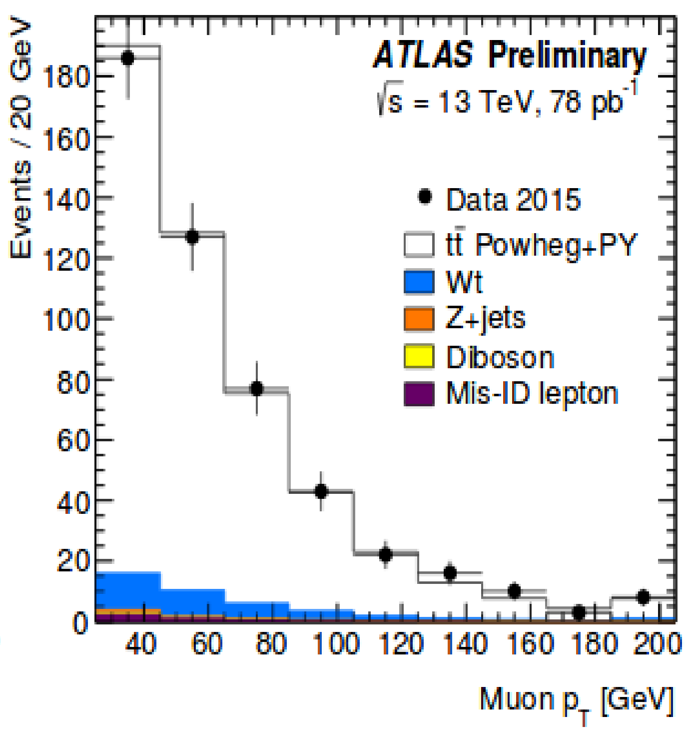
$$R_{tt/Z} = \sigma_{tt} / (0.5(\sigma_{Z\to ee} + \sigma_{Z\to\mu\mu}))$$

t-tbar & Z measurements with the same lepton identification, isolation & trigger criteria within same phase-space cuts (additional dileption mass-cut for Z cross-section)

50 ns data in 2015

Uncertainties due to lepton energy scale, resolution, id, trigger are taken fully correlated
← don't cancel completely in ratio : as spectra are different in shape
(Isolation uncertainty is uncorrelated as environment around lepton is different for two cases)

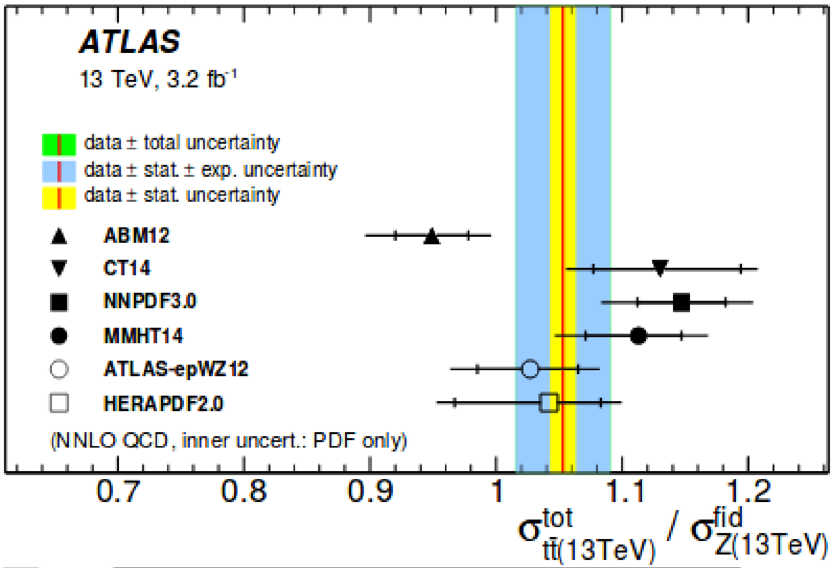


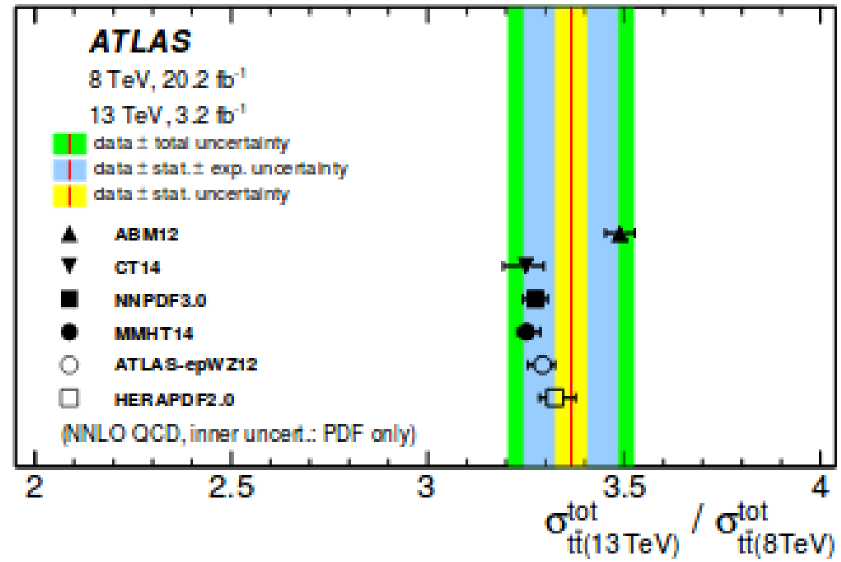




ATLAS-CONF-2015-049

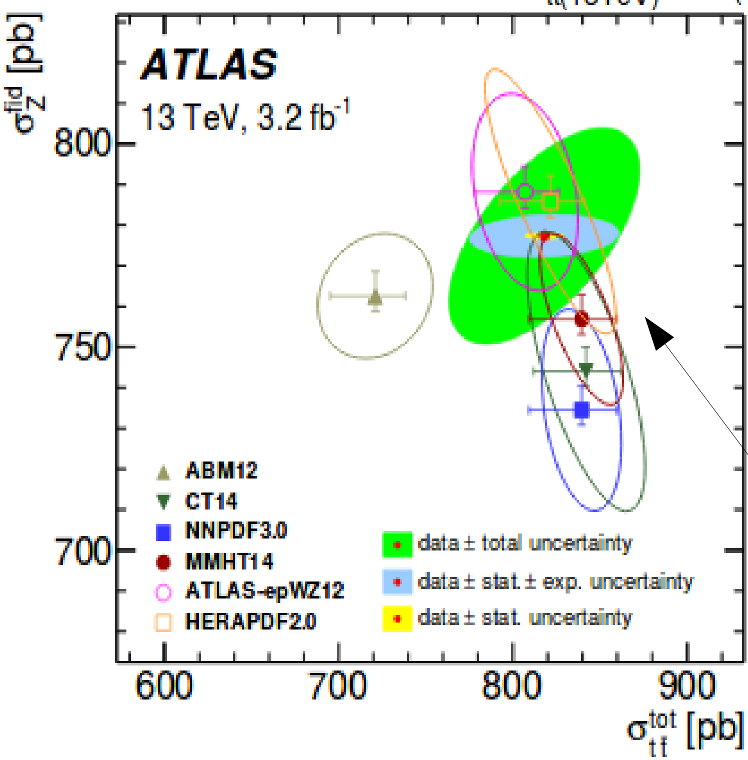
Ratio is sensitive to both valence quark & gluon PDF

# Cross-section Ratio of t-tbar to Z at $\sqrt{s}$ = 13 TeV




ATLAS-epWZ12 determined from HERA data + W->lv +Z->ll data From ATLAS (7 TeV)




ABM12, ATLAS-epWZ12 & HERAPDF2.0 does not use collider jet data => softer gluon density at large x

ABM12 uses lower value for $\alpha_s$

For σ(t-tbar)/σ(Z) ratio, experimental uncertainty competes with theoretical uncertainty (with full run 2 data exp systematics will go down)

25 ns data in 2015

Correlation between measured cross-sections is in the opposite direction to predictions from different PDF sets
: Data can favour one PDF over others!

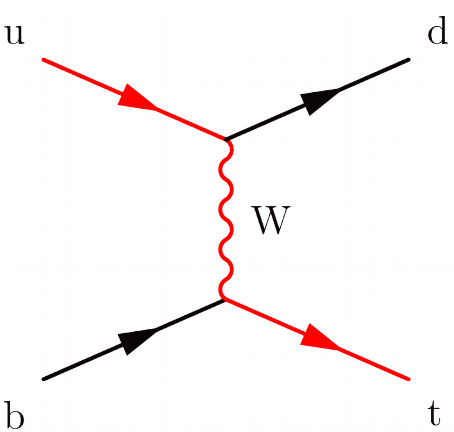
arXiv: 1612.03636

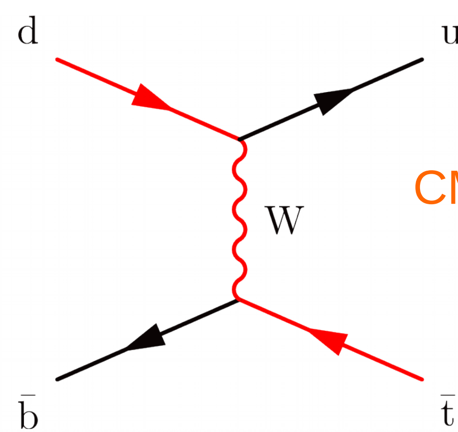
# Single Top Production in t channel at $\sqrt{s}$ = 13 TeV

- Probes light quark via top quark charge (t/tbar)

Lepton + jets channel

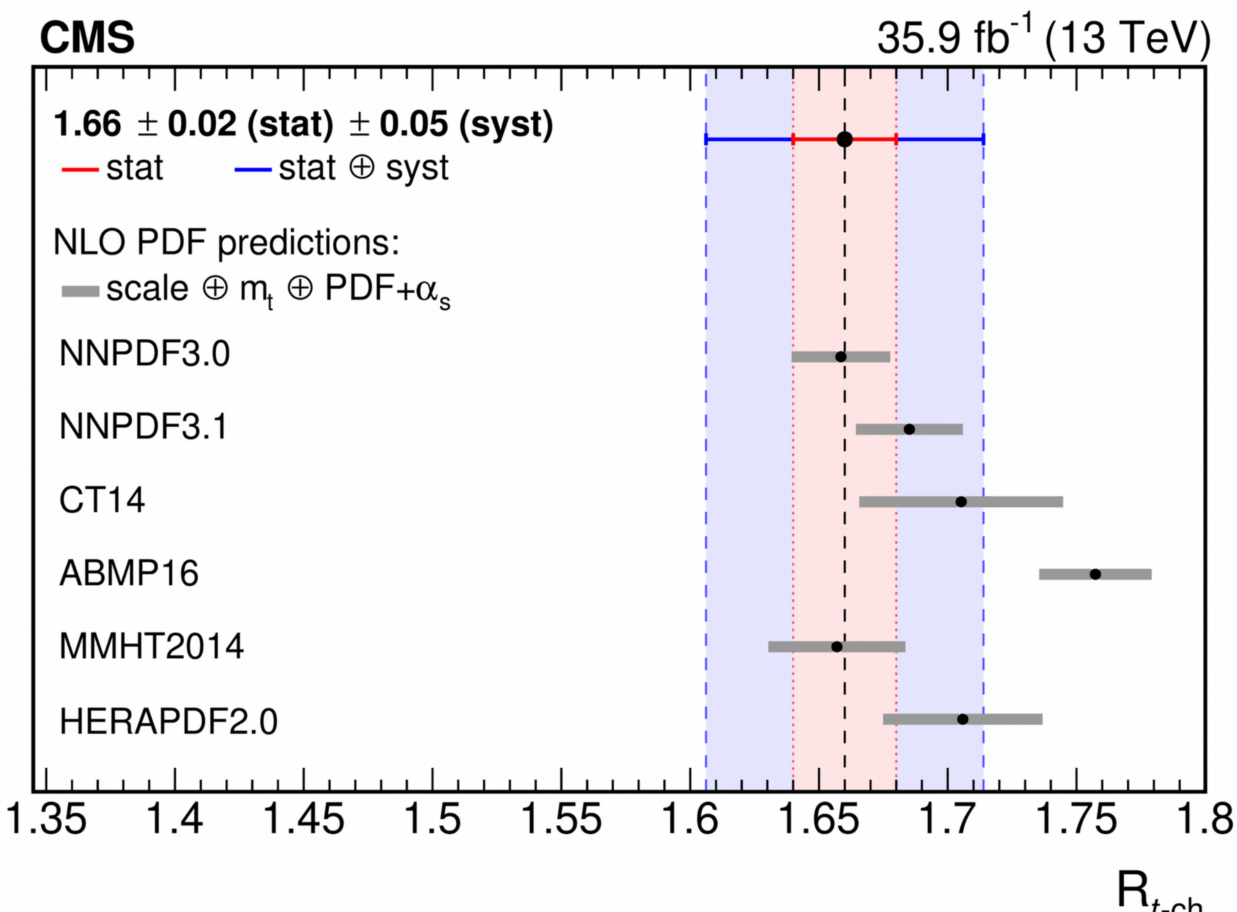
$\sigma_{t\text{-ch},t} / \sigma_{t\text{-ch, tbar}}$ : senstive to u/d PDF



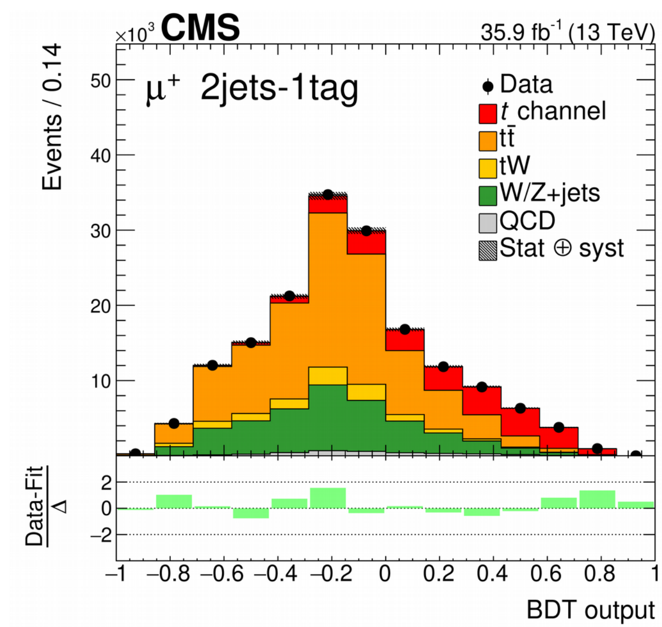




CMS-TOP-17-011

Signal extraction from
2-jet-1-tag signal region &
3-jet-1-tag & 3-jet-2-tag control reg



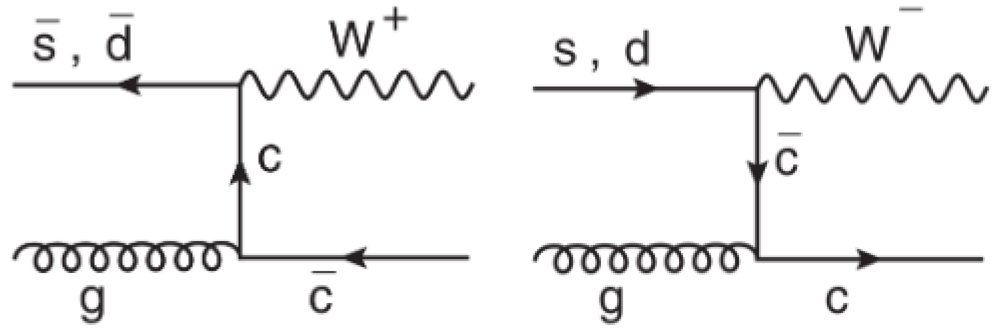

<u>Signal extraction using MVA</u>




$\sigma_{t\text{-ch},t}$ = 136.3 pb +- 1% (stat) +-15% (sys)
$\sigma_{t\text{-ch, tbar}}$ = 82.7 pb +- 1% (stat) +-16% (sys)
$\sigma_{t\text{-ch}}$ = 219.0 pb+- 1% (stat) +-15% (sys)
$\Sigma^{SM}_{t\text{-ch}}$ = 217 +7-5% (scale) +-6% (PDF)

# W+c cross-Section at $\sqrt{s}$ = 13 TeV

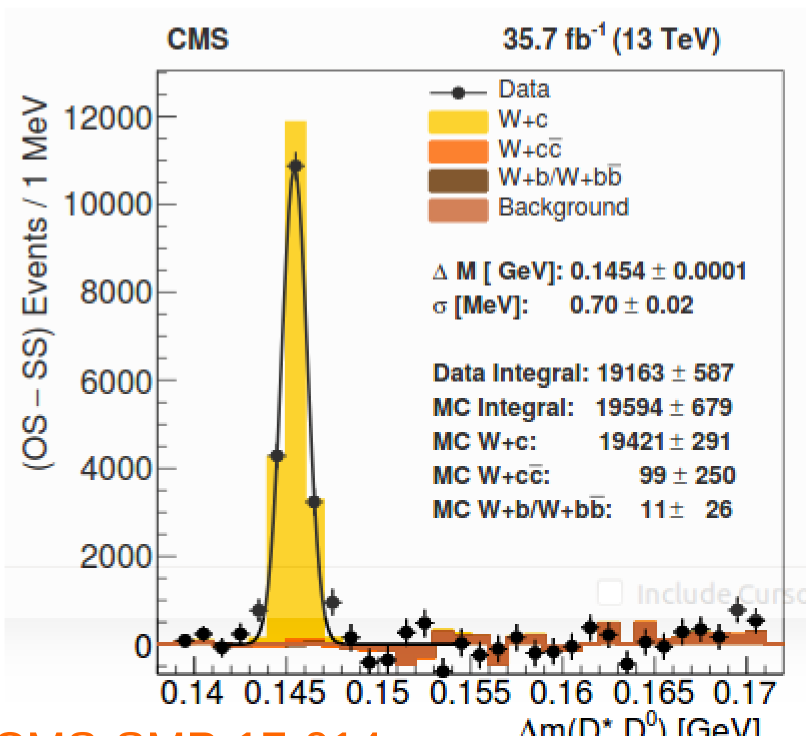
- W $\rightarrow \mu\, \nu_\mu$
- Direct sensitivity to strange quark PDF
- Charge of W (= charge of μ) is opposite to charge of c quark (= charge of D*)



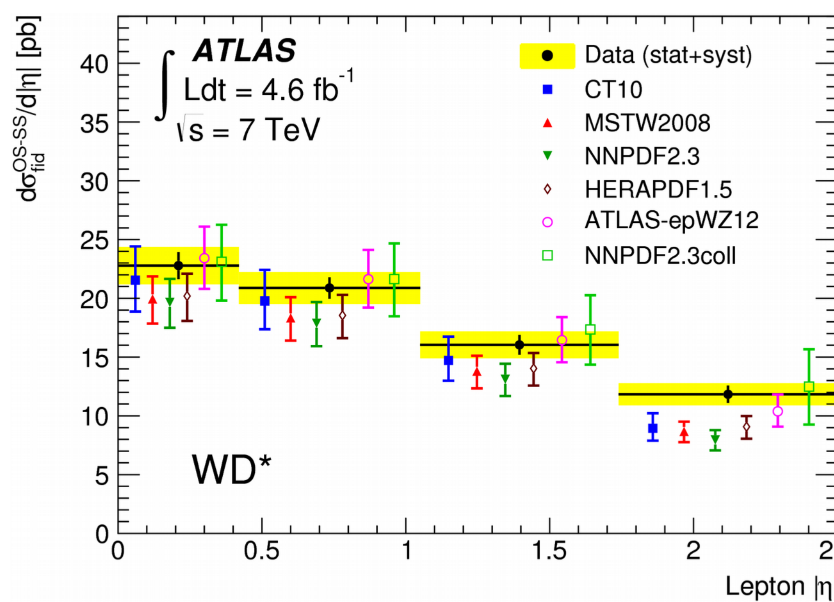

Hadronic decay of c quark hadronizing into D* Meson, benifitted from complete reconstruction using tracks ($D^* \rightarrow D^0\, \pi \rightarrow K\, \pi\, \pi$)

Data driven background subtraction

Using $D^*$-$D^0$ mass difference to identify D* => Reduces bkg from W+g, with g splitting into c-$\bar{c}$




CMS-SMP-17-014

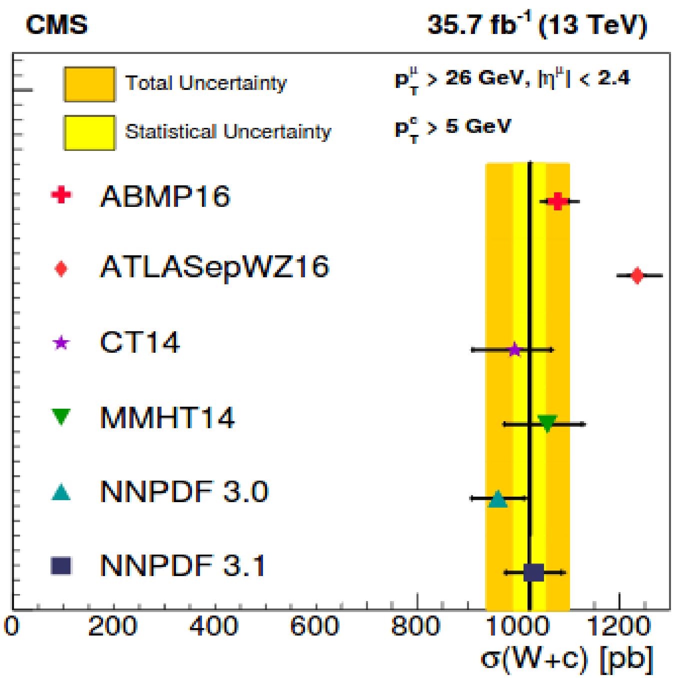


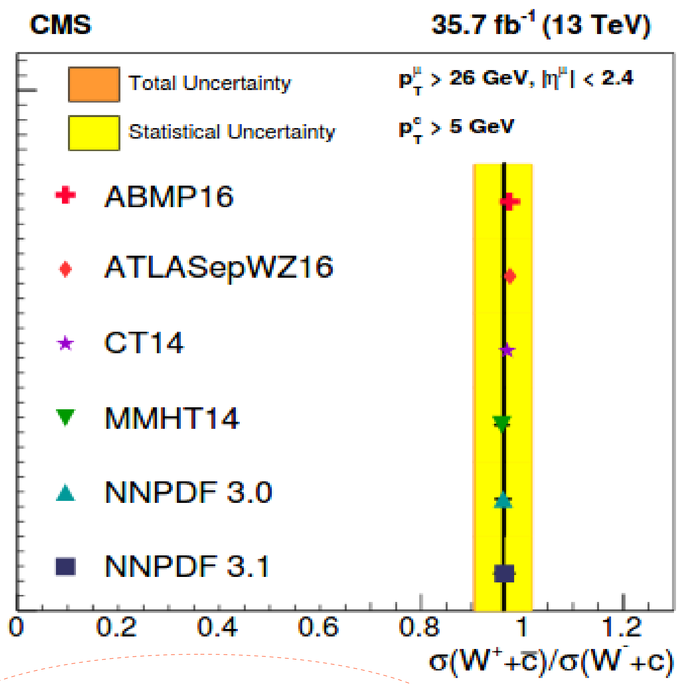

Most of PDF sets predict lower cross section than measured in data, except ATLAS-epWZ12 and NNPDF2.3 coll

NNPDF2.30 excluding neutrino scattering data (has large s fraction than $\bar{d}$ )

# W+c cross-Section at √s = 13 TeV

- Systematic uncertainty cancels to a great extent in the ratio



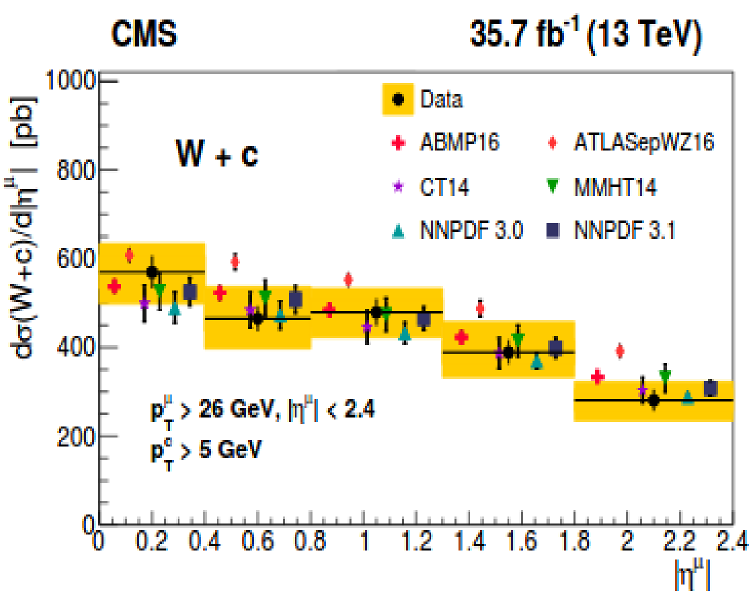




ATLASSepWZ16 (PDF fiting HERA + ATLAS W, Z data) predicts larger cross section for W+c production, although difference cancels in ratio

Differential distributions are also well described by NLO calculations with most of the PDF sets considered

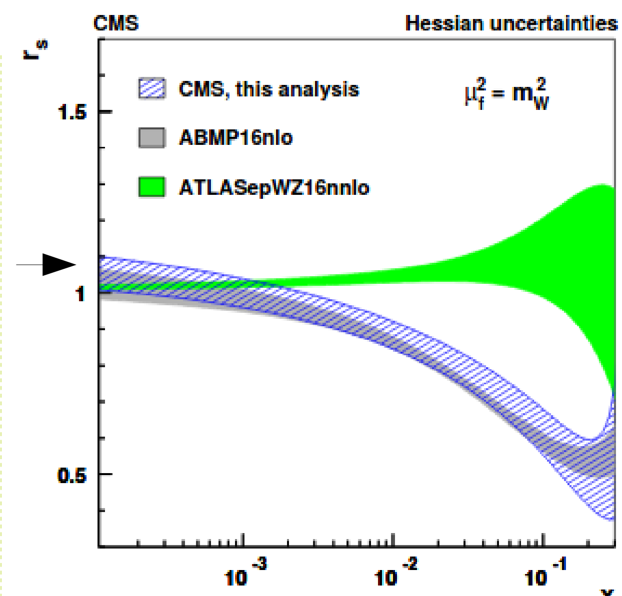
CMS-SMP-17-014

Tension between CMS & ATLAS results



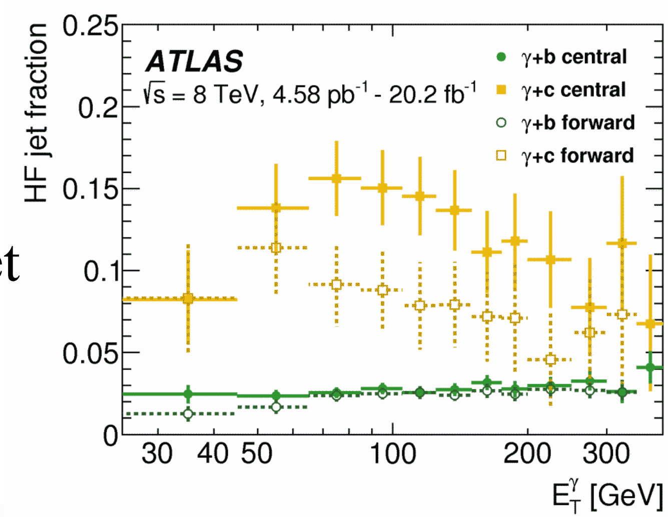

Strangenss supression in sea quark observed, consistent with ABMP16 using fitted collider data

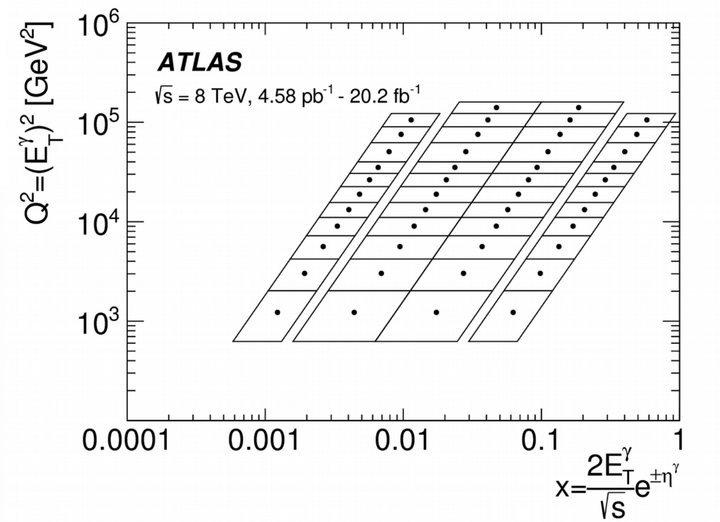
ATLAS data predicts large strange quark fraction at high x



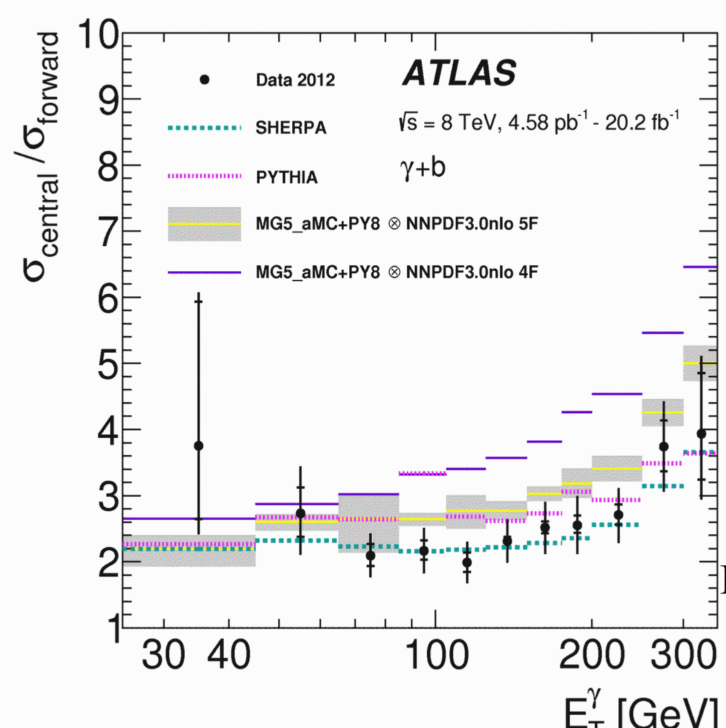
# Isolated photon + HF jets @ $\sqrt{s}$ = 8 TeV

- p p ► $\gamma$ + b/c jets (q+g ► $\gamma$+q : Compton process)
- Sensitivity to extrinicity & intrinsicity of b and c quarks in proton
- Identification of b, c jets using MVA discriminator
- Better calibration of $\gamma$ than jet
- 5 FS : massless b in proton
  4 FS : massless c in proton
  massive b from g -> $b\bar{b}$



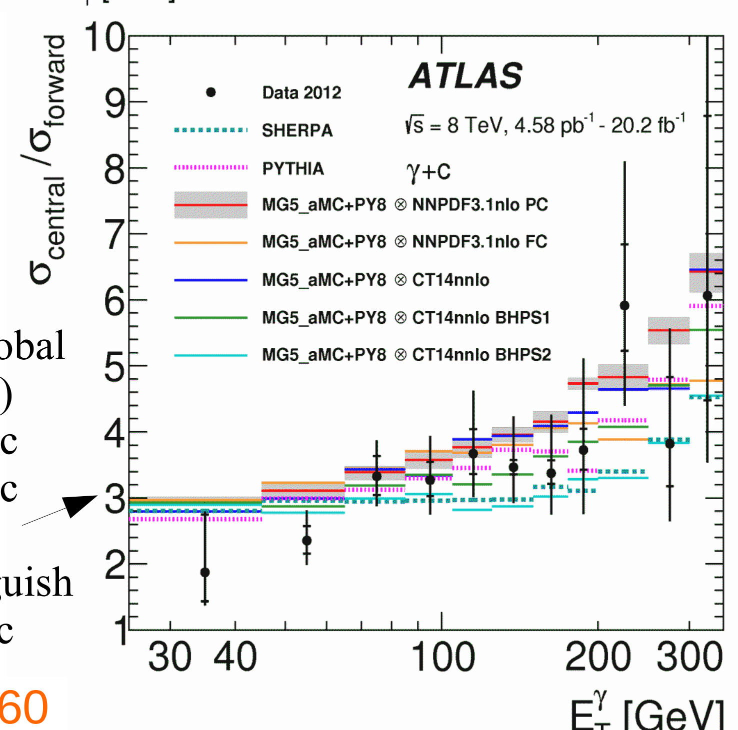

Central ($|\eta^\gamma|$<1.37) => low x
Forward (1.54<$|\eta^\gamma|$<2.37) => high x




5 FS (although at the edge of uncertainties) gives much better description of data than 4 FS

FC : fitted charm PDF in global fit ( 0.26 % instrinsic c)
BHPS1 : 0.6% instrinsic c
BHPS2 : 2.1% instrinsic c

Erros are too large to distinguish different descriptions of c




arXiv : 1710.09560

# Conclusion

- A number of potential analyses explored PDF sensitivity by ATLAS & CMS;



W charge asymmetry => Constrains u, d PDFs more at low-medium x, no major impact on gluon PDF

LHCb W+Z data => indicates difference between PDF of $\bar{u}$ & $\bar{d}$

Inclusive jets => Constrains u, d PDFs both at low & high x, gluon PDF at high x

ttbar data => Predicts harder gluon PDF, reduces uncertainty at high x, and also correlation with $\alpha_S$

Single top => Gives a measure of total sea, and asymmetry between u and d PDFs

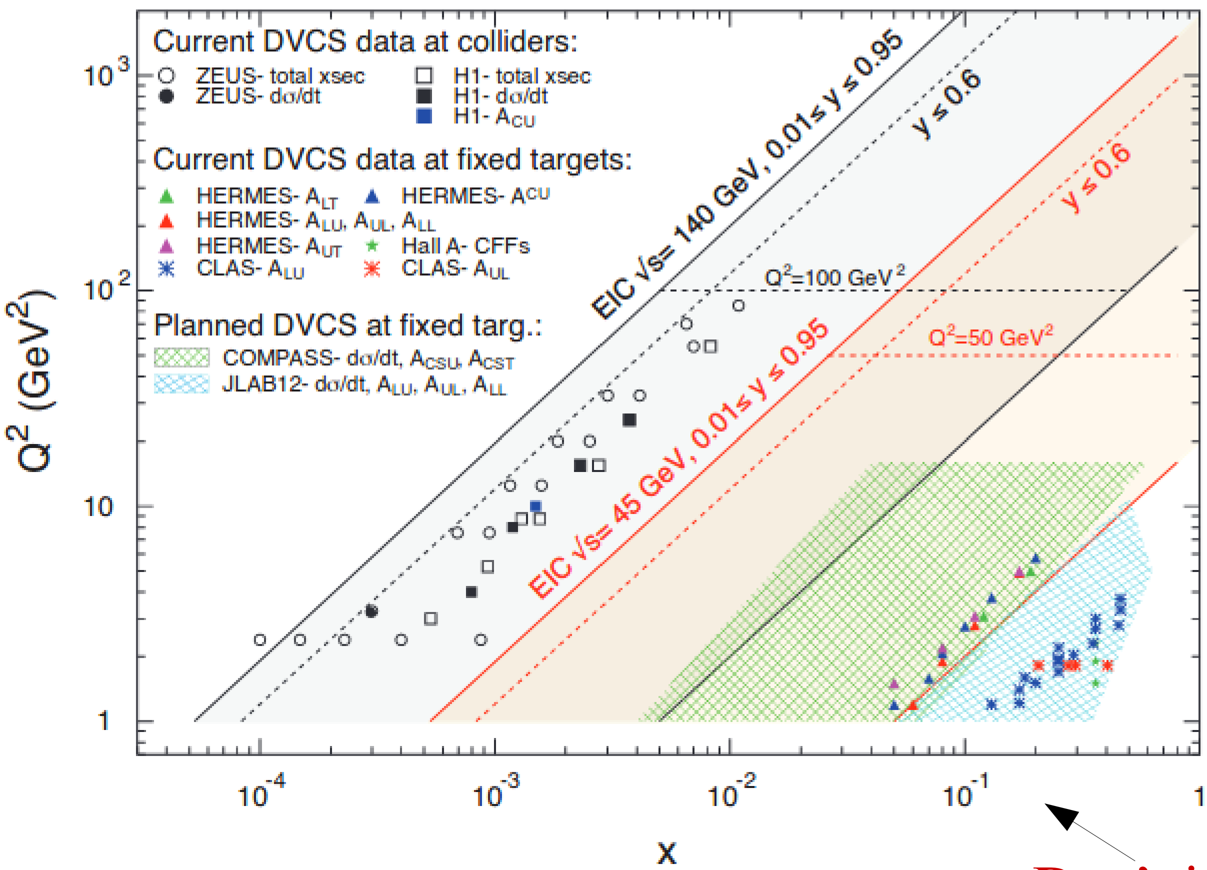
W+c => Different predictions for strangeness supression from ATLAS & CMS

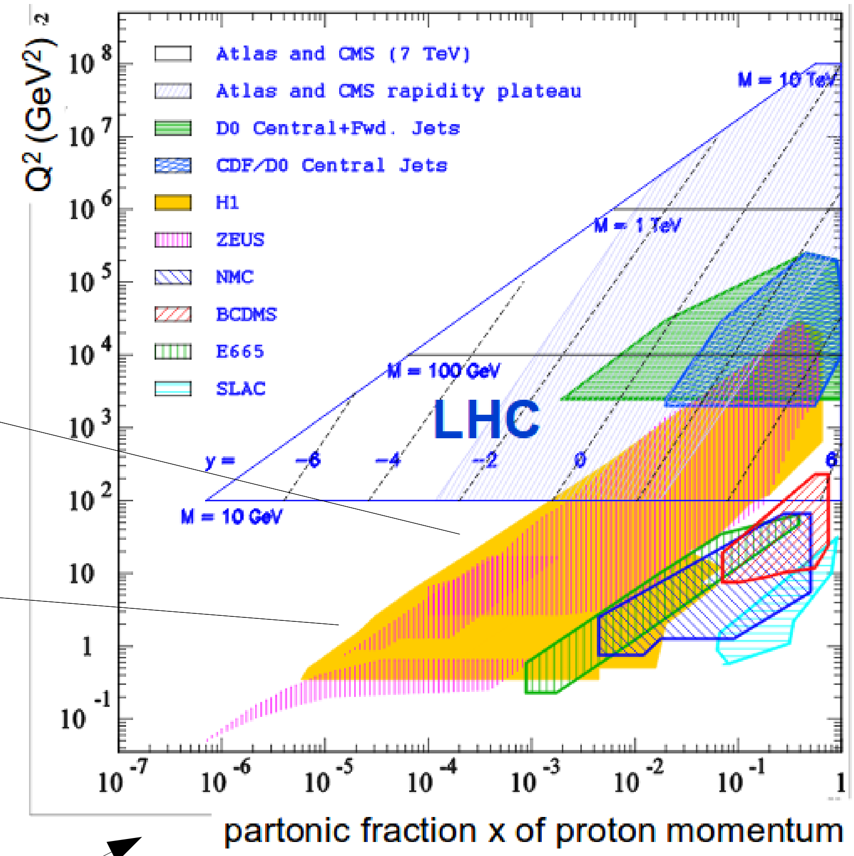
γ+jet => Probes heavy quark PDF in proton

  many more to come ..

- Still, hadron collider has nasty environment, initial states are not precisely know
  => larger uncertainty both from theoretical and experimental sides

# Conclusion



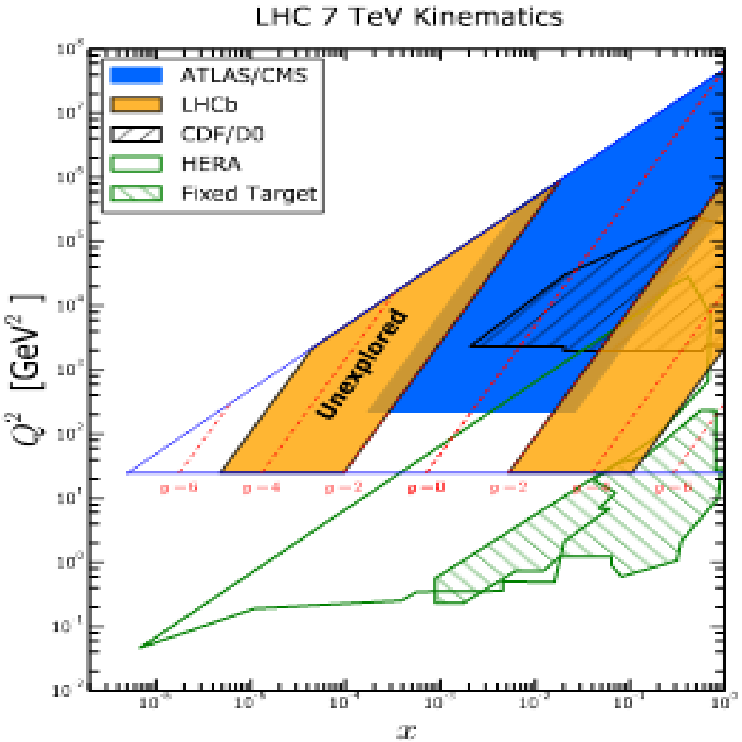

Don't judge equally!!

- Let EIC explore x-$Q^2$ plane with higher resolution, and enlighten the region not yet probed well, and exploit the progress in precision calculation!

THANK YOU!

# More Material ....

# PDF in LHCb




Forward tracking detector

- Unique capabability to explore exclusive regions in x, specially small x region (measurement reported here probes large x portion only)
- Complementary to the measurements by ATLAS & CMS
- Excellent vertex reconstruction (useful for b,c tagging), very good lepton reconstruction & identification up to 100 GeV

## Limitation:

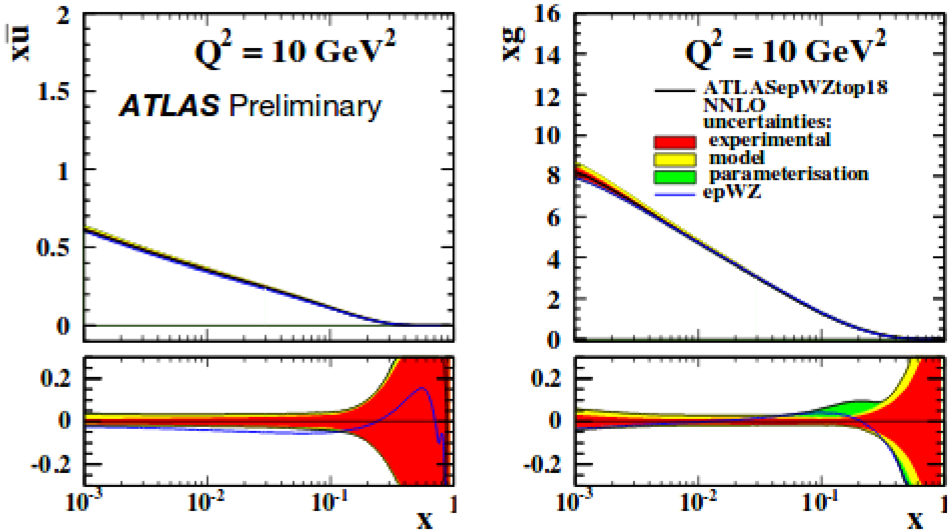
Low rate for data taking => limited statistics

Full event reconstruction is not always possible (hadronic objects suffer more)

# Impact Double Differential t-tbar Cross-Section (ATLAS Data) on PDF

ATL-PHYS-PUB-2018-017

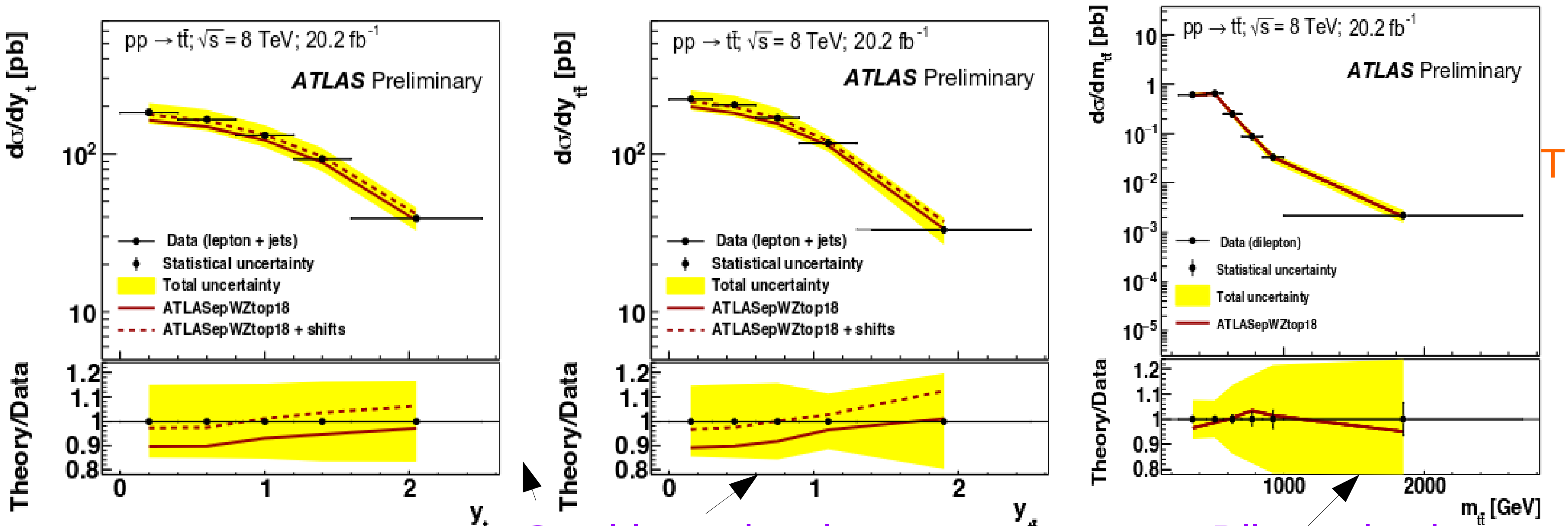
Modelling uncertainty involves : c-quark mass, b-quark mass, min $Q^2$ on HERA data, $Q^2$ used as starting point of PDF evolution, top mass, $\alpha_s$




Uncertainty from experimental sources compete with modelling / parameterisation uncertainty

ATLASSepWZtop18 will be available in LHAPDFv6!

Data & Prediction agree within uncertainty band




These variables have not been used in PDF fit

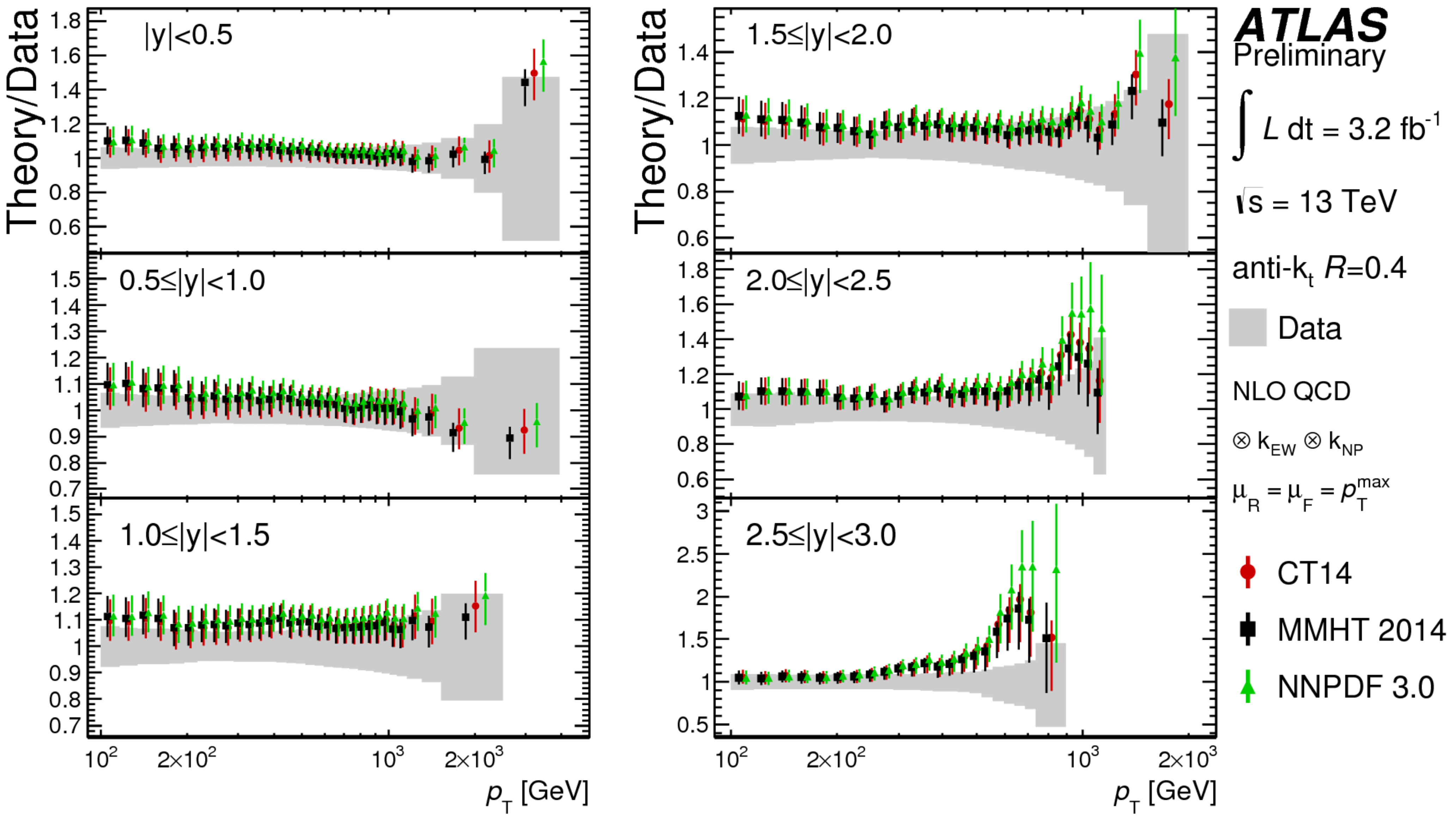
Semi-leptonic t-tbar

Dileptonic t-bar

# Inclusive Jets in ATLAS @ $\sqrt{s}$ = 13TeV

25 ns data in 2015

- Fluctuation in data is less than CMS (larger dataset used in ATLAS)



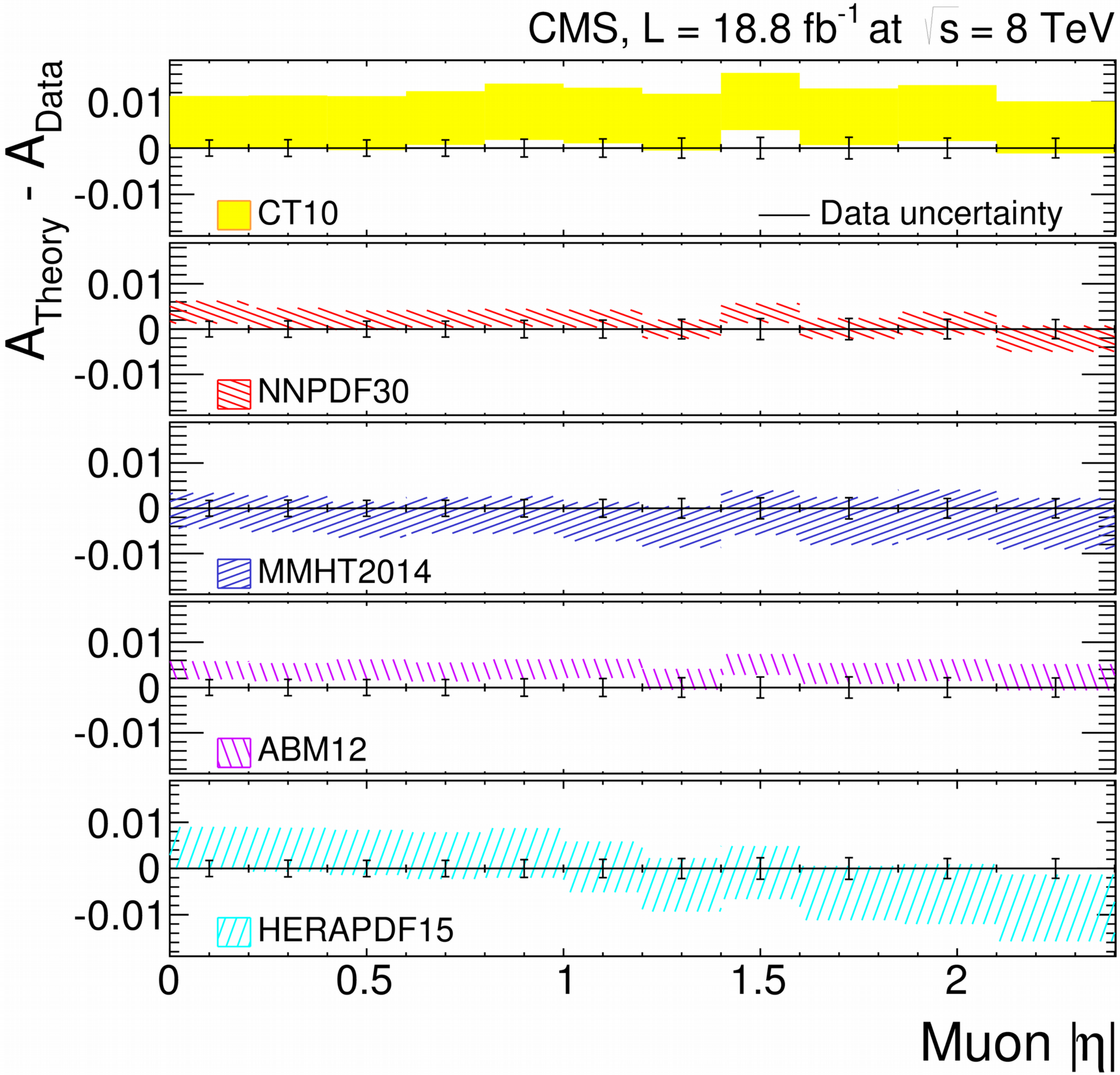

arXiv: 1711.02692



- Prediction from NNPDF, MMHT are similar to CT14

# W charge asymmetry @ √s = 8 TeV

- p p → W → $\mu\, \nu_\mu$ (Charge of muon = charge of W)




CMS-SMP-14-022

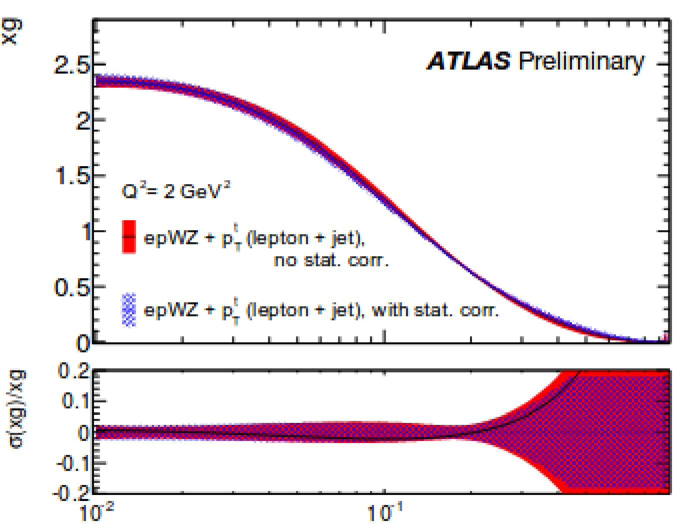
# Impact Double Differential t-tbar Cross-Section (ATLAS Data) on PDF

- ATLASepWZ : PDF obtained fitting HERA data & ATLAS inclusive W, $Z/\gamma^*$ data + t-tbar (semileptonic+dileptonic channels) => ATLASepWZtop18
- Statistical correlation within each spectrum & between different spectra is determined using Bootstrap method
- Systematic uncertainties are taken fully correlated between different spectra using nuisance parameter for each source

(parton-shower model is decorrelated between spectra for lepton+jets to maintain quality of the fit)

available in LHAPDFv6!

| | | lepton+jets spectrum | | | | dilepton spectrum | |
|---|---|---|---|---|---|---|---|
| | | $m_{tt}$ | $p_T^t$ | $y_{tt}$ | $y_t$ | $m_{tt}$ | $y_{tt}$ |
| Total $\chi^2$/NDF | | 1238.4 / 1062 | 1239.4 / 1063 | 1257.5 / 1060 | 1246.5 / 1060 | 1233.8 / 1061 | 1233.8 / 1060 |
| Partial $\chi^2$/NDP | HERA | 1153 / 1016 | 1151 / 1016 | 1149 / 1016 | 1146 / 1016 | 1152 / 1016 | 1147 / 1016 |
| Partial $\chi^2$/NDP | ATLAS $W, Z/\gamma^*$ | 82.0 / 55 | 82.1 / 55 | 86.4 / 55 | 85.0 / 55 | 79.3 / 55 | 82.8 / 55 |
| Partial $\chi^2$/NDP | ATLAS $t\bar{t}$ | 3.4 / 7 | 7.9 / 8 | 19.7 / 5 | 18.3 / 5 | 2.6 / 6 | 4.5 / 5 |




Observables have been chosen based on $\chi^2$ of the fits to the corresponding spectra and also after combined fit

ATL-PHYS-PUB-2018-017

| | | lepton+jets spectra | | | |
|---|---|---|---|---|---|
| | | $p_T^t$ and $y_t$ with statistical correlations | $p_T^t$ and $y_t$ without statistical correlations | $p_T^t$ and $m_{tt}$ with statistical correlations | $p_T^t$ and $m_{tt}$ without statistical correlations |
| Total $\chi^2$/NDF | | 1264 / 1068 | 1260 / 1068 | 1290 / 1070 | 1287 / 1070 |
| Partial $\chi^2$/NDP | HERA | 1148 / 1016 | 1147 / 1016 | 1162 / 1016 | 1162 / 1016 |
| Partial $\chi^2$/NDP | ATLAS $W, Z/\gamma^*$ | 82.7 / 55 | 83.5 / 55 | 83.2 / 55 | 83.1 / 55 |
| Partial $\chi^2$/NDP | ATLAS $t\bar{t}$ | 33 / 13 | 30 / 13 | 45 / 15 | 42 / 15 |

# Cross-Section for Isolated Photon + Jets $\sqrt{s}$ = 13 TeV

- Dominant sub-process : qg ⟶ qγ: sensitive to gluon PDF

**Requirements on photons**

$E_T^{\gamma} > 125$ GeV, $|\eta^{\gamma}| < 2.37$ (excluding $1.37 < |\eta^{\gamma}| < 1.56$)

$E_T^{iso} < 4.2 \cdot 10^{-3} \cdot E_T^{\gamma} + 10$ GeV

**Requirements on jets**

anti-$k_t$ algorithm with $R = 0.4$

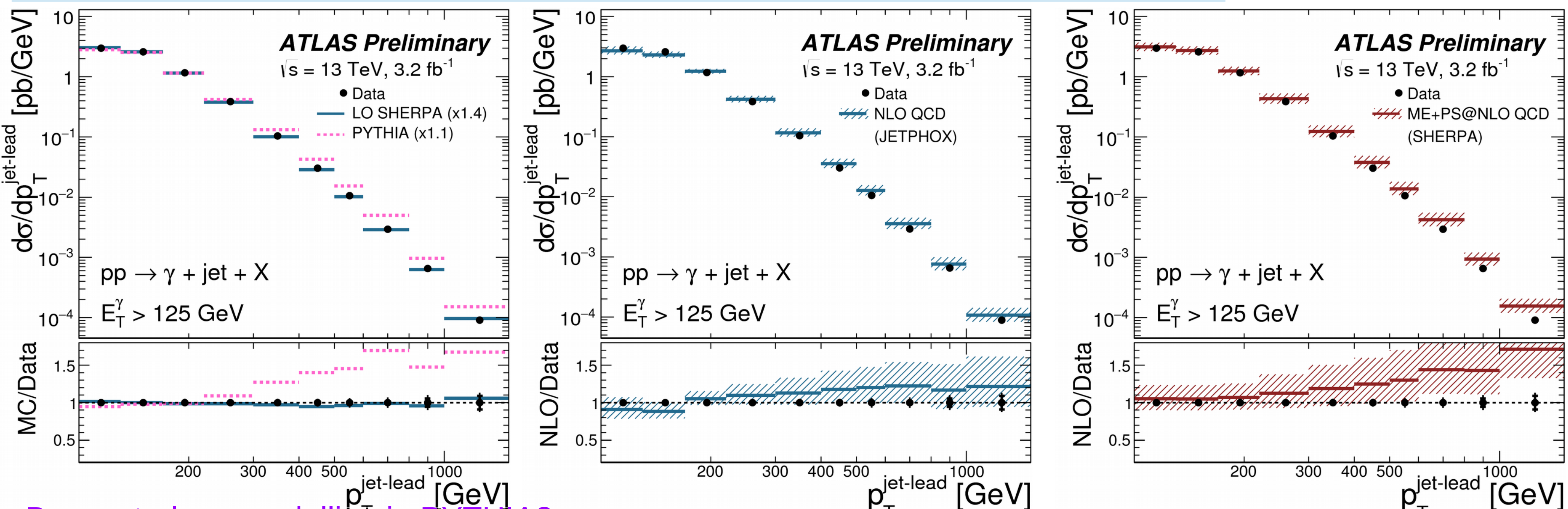
$p_T^{jet-lead} > 100$ GeV, $|y^{jet-lead}| < 2.37$

$\Delta R^{\gamma-jet} > 0.8$

MC prediction doesn't depend very much on PDFs, but sensitivity can be explored more

ATLAS-CONF-2017-059

PDFs used : CT10 NLO (in LO SHERPA ) NNPDF2.3LO(in PYTHIA8) MMHT2014 (in JETPHOX) NNPDF3.0NNLO(in NLO SHERPA)




Bremsstralung modelling in PYTHIA8 predicts more radiation than seen in data

# Impact Double Differential t-tbar Cross-Section (ATLAS Data) on PDF

| Systematic uncertainty source | lepton+jets spectrum | | | |
|---|---|---|---|---|
| | $p_T^t$ | $y_t$ | $y_{tt}$ | $m_{tt}$ |
| Hard scattering model | +0.74± 0.31 | +0.48± 0.22 | +0.92± 0.37 | -0.43± 0.20 |
| Parton shower model | -1.32± 0.43 | -0.79± 0.26 | -0.51± 0.17 | +0.39±0.13 |
| ISR/FSR model | -0.47± 0.18 | -0.87±0.30 | -1.27± 0.38 | +0.33± 0.10 |

| | | lepton+jets spectra | | |
|---|---|---|---|---|
| | | $p_T^t$ and $y_t$ decorrelate 2-point uncertainties | $p_T^t$ and $m_{tt}$ decorrelate 2-point uncertainties | $p_T^t$ and $m_{tt}$ decorrelate parton-shower model uncertainty |
| Total $\chi^2$/NDF | | 1259 / 1068 | 1247 / 1070 | 1248 / 1070 |
| Partial $\chi^2$/NDP | HERA | 1147 / 1016 | 1154 / 1016 | 1153 / 1016 |
| Partial $\chi^2$/NDP | ATLAS $W, Z/\gamma^*$ | 83.9 / 55 | 81.9 / 55 | 81.6 / 55 |
| Partial $\chi^2$/NDP | ATLAS $t\bar{t}$ | 27.8 / 13 | 11.5 / 15 | 14.1 / 15 |

ATL-PHYS-PUB-2018-017

| | | lepton+jets spectra | | | |
|---|---|---|---|---|---|
| | | $p_T^t$ and $y_t$ with statistical correlations | $p_T^t$ and $y_t$ without statistical correlations | $p_T^t$ and $m_{tt}$ with statistical correlations | $p_T^t$ and $m_{tt}$ without statistical correlations |
| Total $\chi^2$/NDF | | 1264 / 1068 | 1260 / 1068 | 1290 / 1070 | 1287 / 1070 |
| Partial $\chi^2$/NDP | HERA | 1148 / 1016 | 1147 / 1016 | 1162 / 1016 | 1162 / 1016 |
| Partial $\chi^2$/NDP | ATLAS $W, Z/\gamma^*$ | 82.7 / 55 | 83.5 / 55 | 83.2 / 55 | 83.1 / 55 |
| Partial $\chi^2$/NDP | ATLAS $t\bar{t}$ | 33 / 13 | 30 / 13 | 45 / 15 | 42 / 15 |