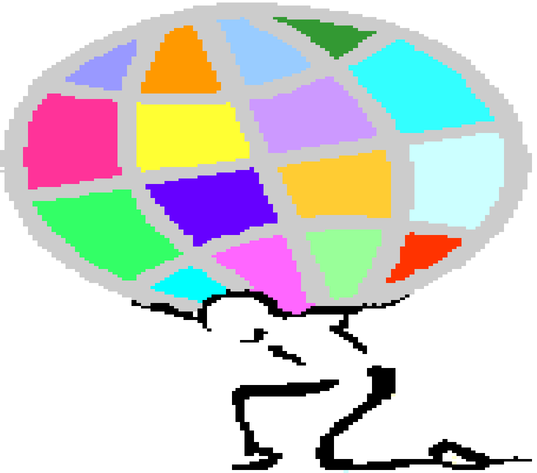
Determination of proton parton distribution functions using ATLAS data

Evelin Meoni
(Università della Calabria & INFN)
On behalf of the ATLAS Collaboration



LEPTON PHOTON 2021 MAN

30th International Symposium on Lepton Photon Interactions at High Energies

Motivations

Proton parton distribution functions (PDFs) are key ingredients to provide theoretical predictions at the LHC:

- dominant uncertainty for high precision measurements (W mass, weak-mixing angle)

- important limiting factor for Higgs production cross-sections

- discovery of new physics relies on precise knowledge of PDFs (i.e. high mass searches)

→ Frontier high-energy physics at colliders requires percent-level accuracy for PDFs

How LHC experiments may contribute to this effort?

- LHC measurements are used in global PDF fits providing important complementary information to datasets from DIS and fixed target experiments in corners of the phase space previously unexplored

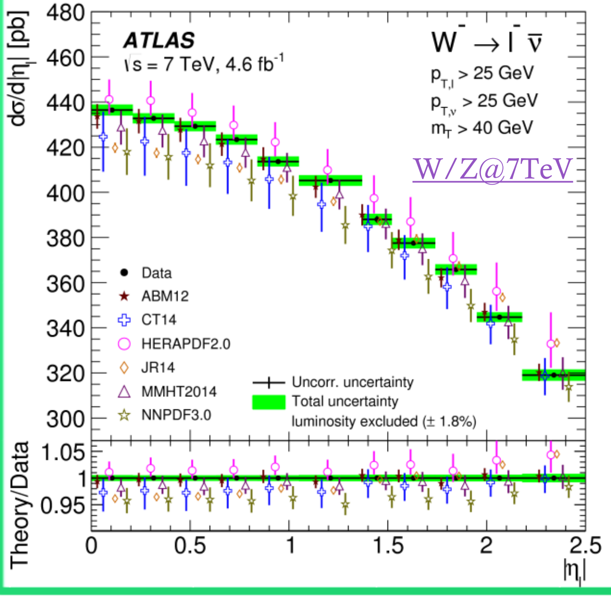
- PDFs fits from LHC collaborations allow a first direct test of the PDF-sensitivity of the measurements and represent a precious "vademecum" for global PDF fitting collaborations

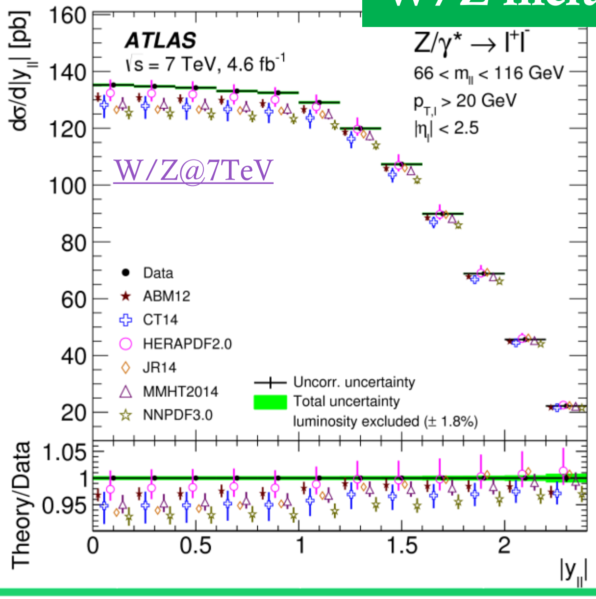
ATLAS PDFs fits

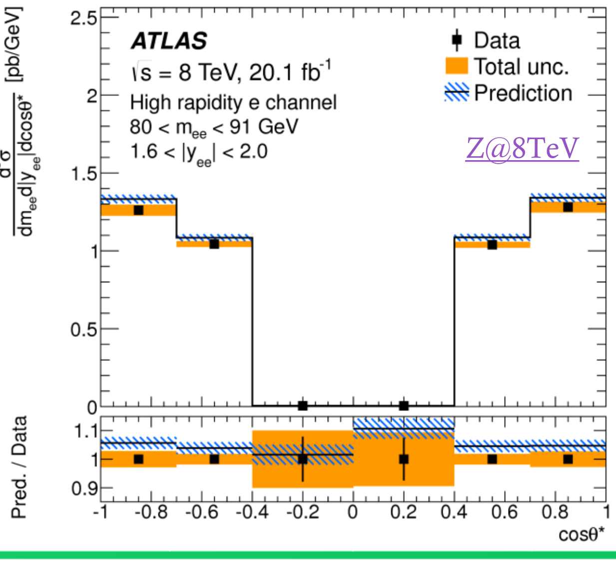
ATLAS PDF set (Publication)	Data	Sensitivity of ATLAS data
ATLASepWZ16 Eur.Phys.J.C 77 (2017) 367	HERA data + ATLAS inclusive W/Z cross-sections@ 7 TeV	Quark flavour separation
ATLASepWZVjet20 (JHEP 07 (2021) 223)	ATLASepWZ16 dataset + ATLAS W/Z+jets cross-sections@8 TeV	Quark flavour separation at higher x
ATLASepWZtop18 (ATL-PHYS-PUB-2018-017)	ATLASepWZ16 dataset + ATLAS $t\bar{t}$ cross-sections@ 8 TeV	Medium- and high-x gluon
ATLASpdf21 (arXiv:2112.11266 [hep-ex])	*Topic of my talk today!*	NEW

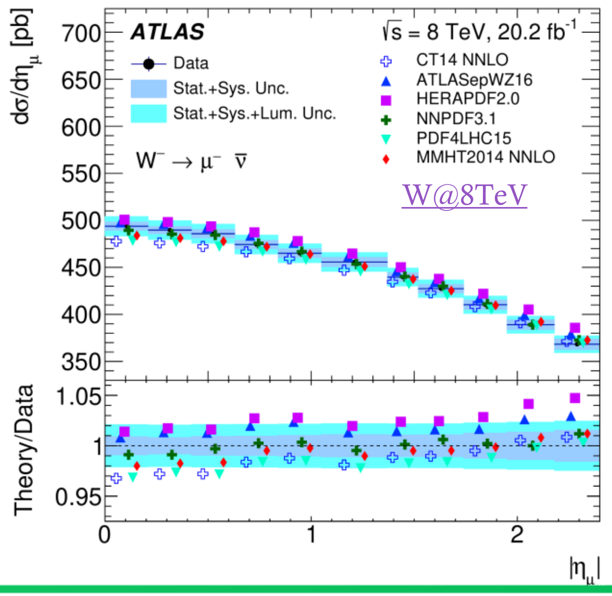
Data sets

W/Z inclusive@7,8 TeV

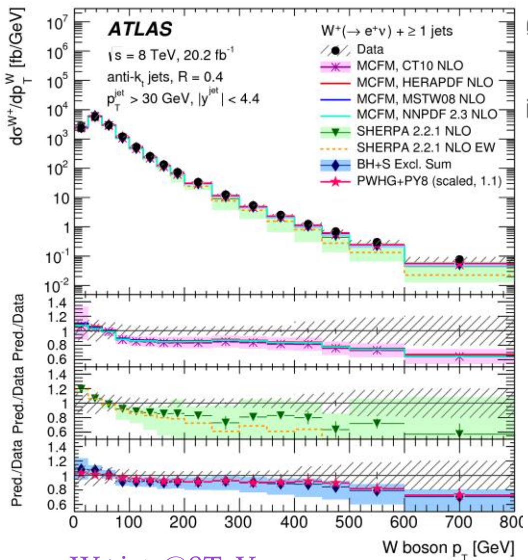




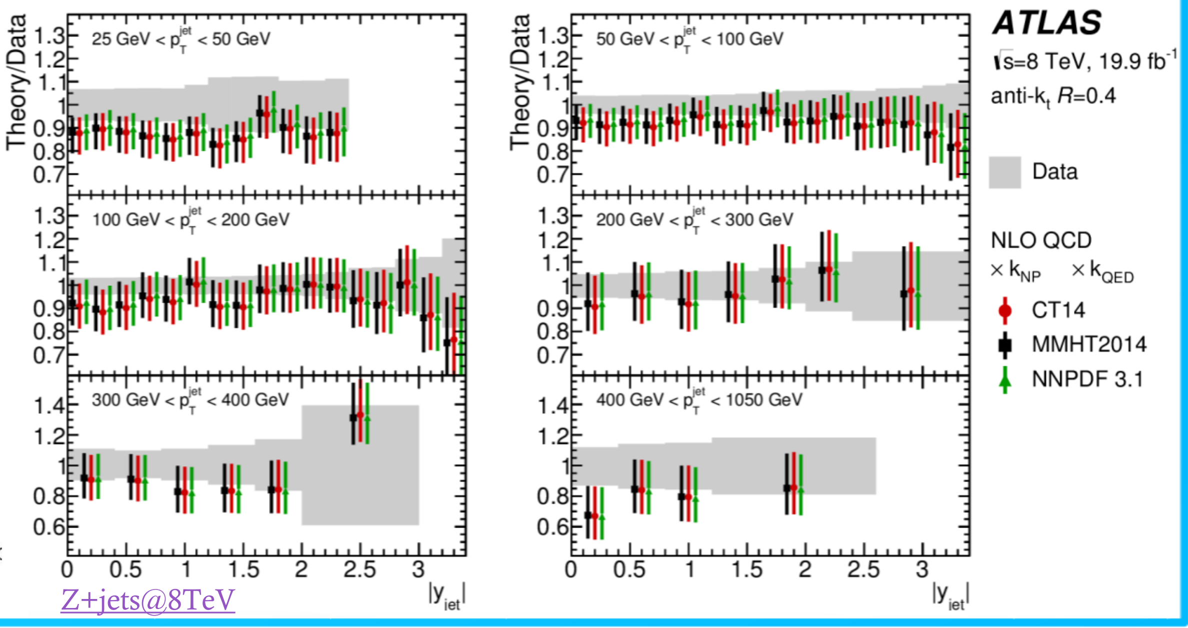




W/Z+jets @8 TeV

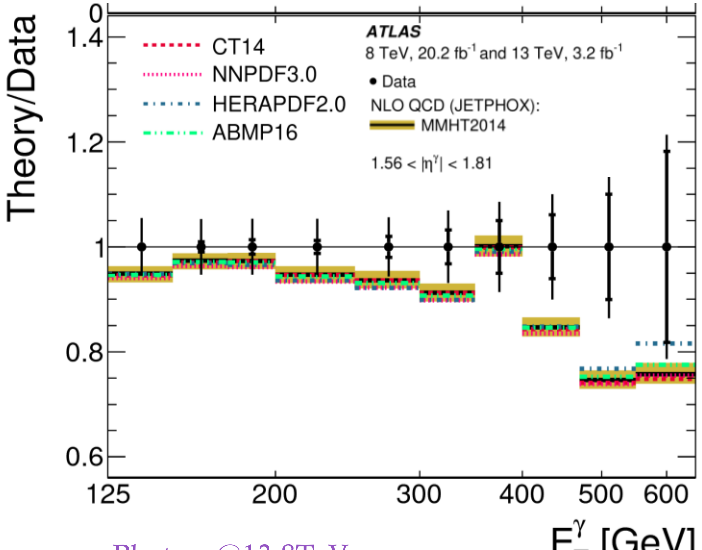


W+jets@8TeV



Z+jets@8TeV

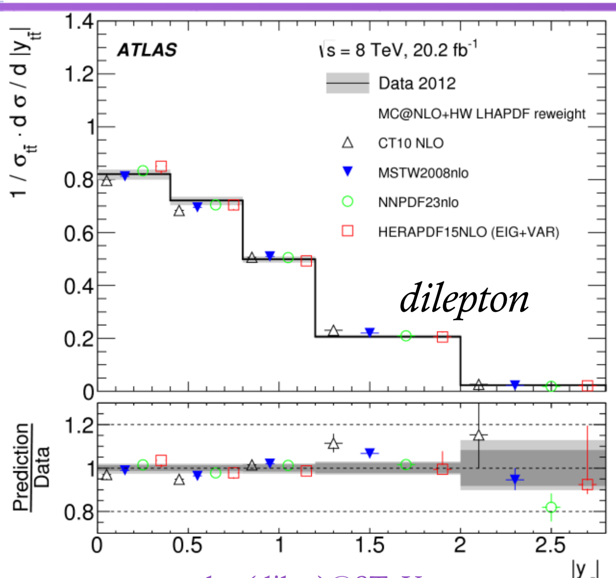
Inclusive Photons@8,13 TeV



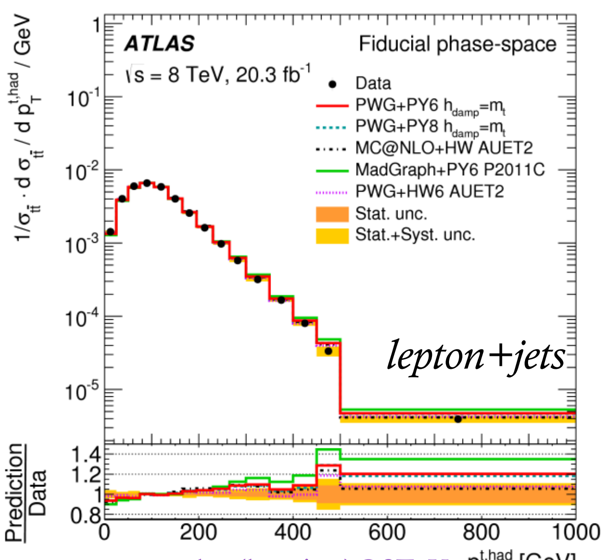
Photons@13,8TeV

Data sets

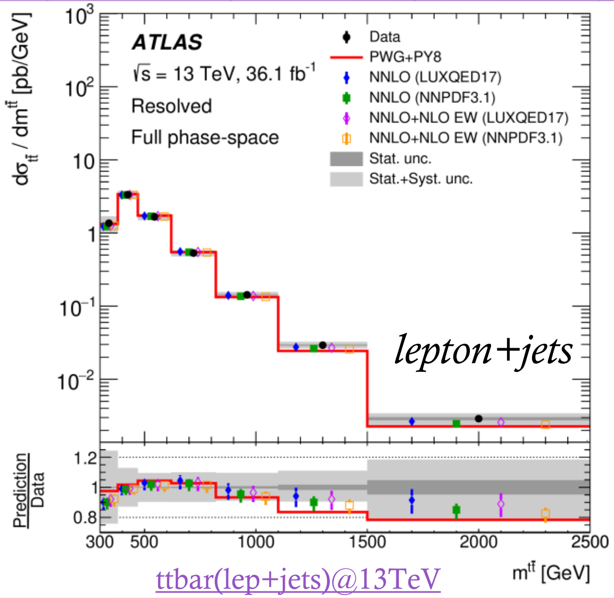
$t\bar{t}$ @8,13 TeV



[ttbar(dilep)@8TeV]



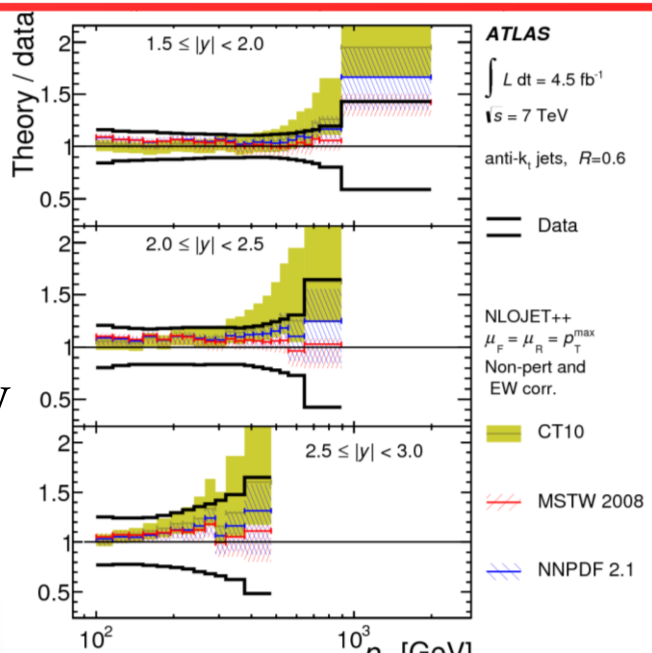
[ttbar(lep+jets)@8TeV]



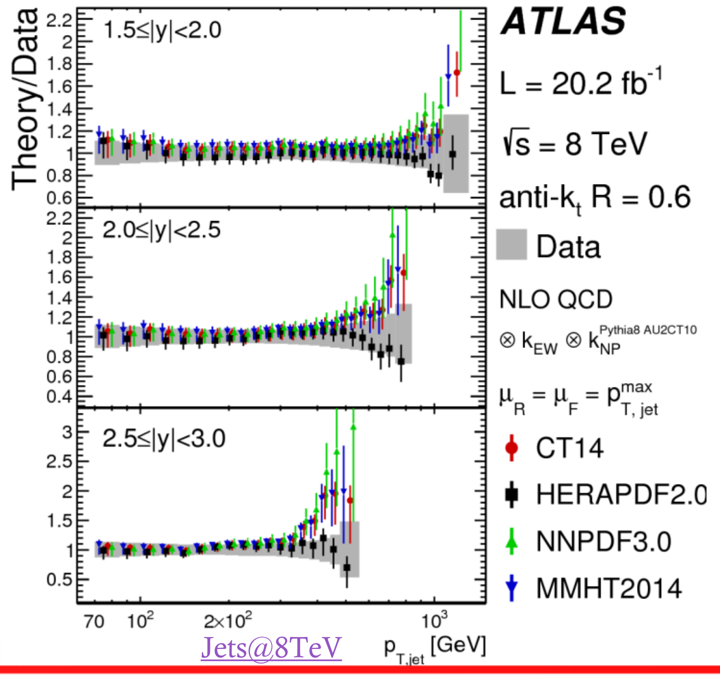
[ttbar(lep+jets)@13TeV]

Inclusive jets @7,8,13 TeV

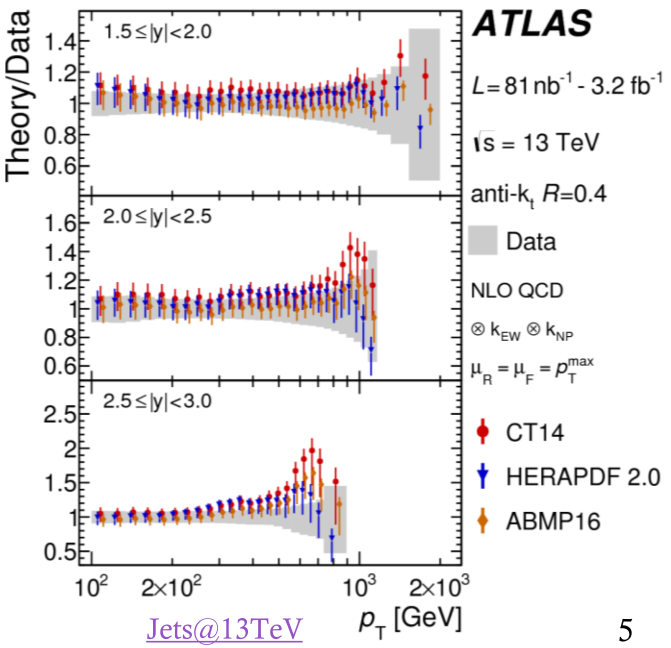
- Data at 8 TeV in the nominal fit
- Data at 7 and 13 TeV used for cross-check



[Jets@7TeV]



[Jets@8TeV]



[Jets@13TeV]

Data sets and theory

Summary of ATLAS data sets

Data set	$\sqrt{s}$ [TeV]	Luminosity [fb^{-1}]	Decay channel	Observables entering the fit
Inclusive $W, Z/\gamma^*$	7	4.6	e, μ combined	η_ℓ (W), y_Z (Z)
Inclusive Z/γ^*	8	20.2	e, μ combined	$\cos\theta^*$ in bins of $y_{\ell\ell}, m_{\ell\ell}$
Inclusive W	8	20.2	μ	η_μ
$W^\pm$ + jets	8	20.2	e	p_T^W
Z + jets	8	20.2	e	p_T^{jet} in bins of $\lvert y^{\text{jet}}\rvert$
$t\bar{t}$	8	20.2	lepton + jets, dilepton	$m_{t\bar{t}}, p_T^t, y_{t\bar{t}}$
$t\bar{t}$	13	36	lepton + jets	$m_{t\bar{t}}, p_T^t, y_t, y_{t\bar{t}}^{b}$
Inclusive isolated γ	8, 13	20.2, 3.2	-	E_T^γ in bins of η^γ
Inclusive jets	7, 8, 13	4.5, 20.2, 3.2	-	p_T^{jet} in bins of $\lvert y^{\text{jet}}\rvert$

A total of 621 ATLAS data points

W,Z@7TeV; W@8TeV; Z@8TeV; W+jets@8TeV ; Z+jets@8TeV ; ttbar(dilep)@8TeV ; ttbar(lep+jets)@8TeV; ttbar(lep+jets)@13TeV; Photons@13,8TeV ; Jets@7TeV; Jets@8TeV ; Jets@13TeV.

HERA data - backbone of the fit – provide good constraints from ~ $10^{-4} < x < 0.1$-0.4 (depending on which PDF) but they cannot provide:

- Flavour separation in the sea
- High-x gluon

these mainly come from ATLAS data

Fit to **NNLO in QCD and NLO in EW** achieved either by direct NNLO grids or by k-factor corrections
- current 'state of the art' – using xFitter framework

Fit Model

•**Parametrisation** of $xu_v, xd_v, x\bar{u}, x\bar{d}, x\bar{s}$ and xg

$$xf_i(x) = A_i x^{B_i} (1-x)^{C_i} P_i(x) \qquad \text{with } P_i(x) = (1 + D_i x + E_i x^2 + F_i x^3)$$

(with an extra negative term for the gluon $-A'_g x^{B'_g} (1-x)^{C'_g}$)

- Normalisation (A_i) and slope parameters (B_i) of $x\bar{u}, x\bar{d}, x\bar{s}$ all independent of each other (strategy only recently adopted by some global PDFs fitters)
- Assumed $xs = x\bar{s}$
- A_i parameters for the valence quarks and the gluon set by the sum-rules
- D_i, E_i, F_i parameters are added until there is no further significant improvement in χ^2 → 21 free parameters
- C'_g set to 25 to avoid negative contribution at high x

Fit Model

- Same χ^2 form as standard in HERA and ATLAS PDF fits

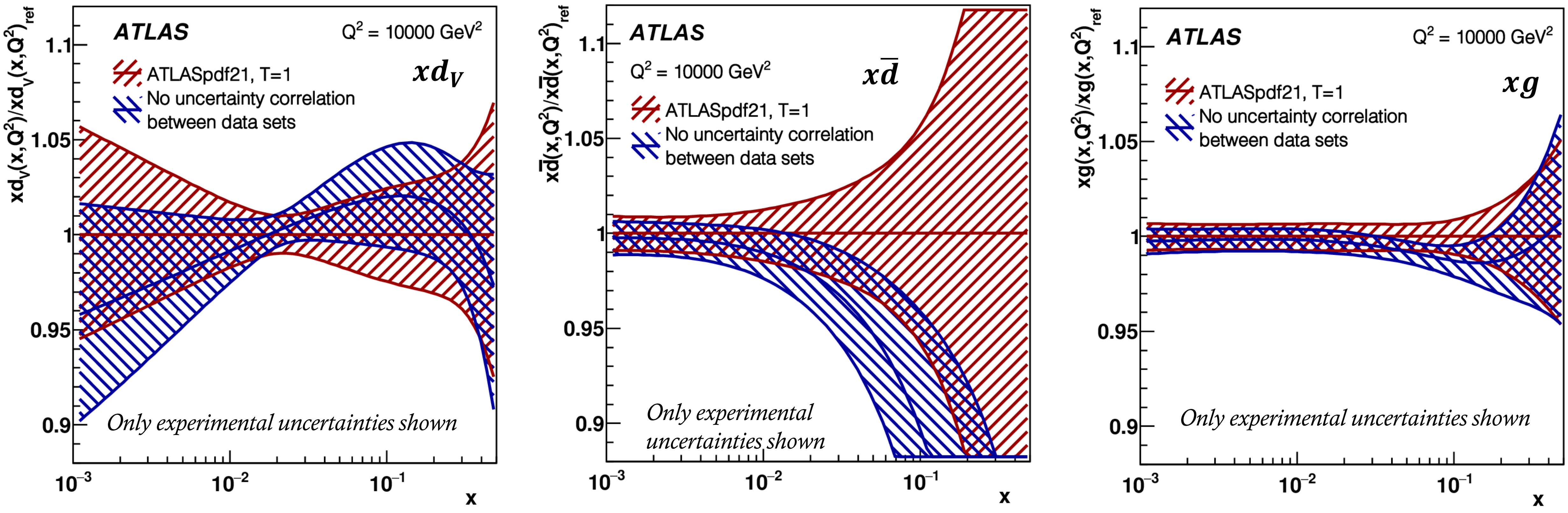
$$\chi^2 = \sum_{ik}\left(D_i - T_i(1-\sum_j \gamma_{ij} b_j)\right) C^{-1}_{\mathrm{stat,uncor},ik}(D_i, D_k)\left(D_k - T_k(1-\sum_j \gamma_{kj} b_j)\right) + \sum_i \log \frac{\delta^2_{i,\mathrm{uncor}} T_i^2 + \delta^2_{i,\mathrm{stat}} D_i T_i}{\delta^2_{i,\mathrm{uncor}} D_i^2 + \delta^2_{i,\mathrm{stat}} D_i^2} + \sum_j b_j^2$$

- Considered correlations of statistical and systematic uncertainties within and between data: Jet Energy Scale correlations play the dominant role

Systematic uncertainties		Measurements: W+jets@8TeV	Z+jets@8TeV	$t\bar{t}$@8TeV	$t\bar{t}$@13TeV	jets@8TeV
JETs	Jet flavour response	JetScaleFlav2	Flavor Response	flavres-jes	JET29NP JET Flavour Response	syst JES Flavour Response*
JETs	Jet flavour composition	JetScaleFlav1Known	Flavor Comp	flavcomp-jes	JET29NP JET Flavour Composition	syst JES Flavour Comp
JETs	Jet punchthrough	JetScalepunchT	Punch Through	punch-jes	-	syst JES PunchThrough MC15
JETs	Jet scale	JetScalePileup2	PU OffsetMu	pileoffmu-jes	-	syst JES Pileup MuOffset
JETs		-	PU Rho	pileoffrho-jes	JET29NP JET Pileup RhoTopology	syst JES Pileup Rho topology*
JETs		JetScalePileup1	PU OffsetNPV	pileoffnpv-jes	JET29NP JET Pileup OffsetNPV	syst JES Pileup NPVOffset
JETs		-	PU PtTerm	pileoffpt-jes	JET29NP JET Pileup PtTerm	syst JES Pileup Pt term
JETs	Jet JVF selection	JetJVFcut	JVF	jetvxfrac	-	syst JES Zjets JVF
JETs	B-tagged jet scale	-	btag-jes	JET29NP JET BJES Response	-	-
JETs	Jet resolution	-	jeten-res	JET JER SINGLE NP	-	-
Lept	Muon scale	-	-	mup-scale	MUON SCALE	-
Lept	Muon resolution	-	-	muonms-res	MUON MS	-
Lept	Muon identification	-	-	muid-res	MUON ID	-
Bkg	Diboson cross section	-	-	dibos-xsec	Diboson xsec	-
Bkg	Z + jets cross section	-	-	zjet-xsec	Zjets xsec	-
Bkg	Single-t cross section	-	-	singletop-xsec	st xsec	-

Correlations

Study the effect of the correlations at a scale relevant for LHC physics and focus on the middling x range where W,Z and Higgs are produced



Differences in PDFs not large (mostly in the d-type sector) and PDFs uncertainties are similar (larger on the high-x $\bar{d}$ and *gluon* PDFs if correlations are not accounted for), but correlation can be important when the aim is percent level accuracy on PDFs

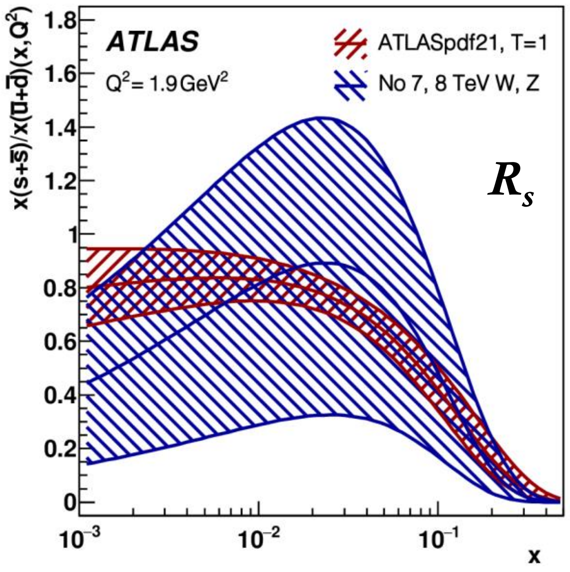
Impact of different datasets

W/Z

Without W/Z data, ratio of strange to light sea quarks very poorly determined

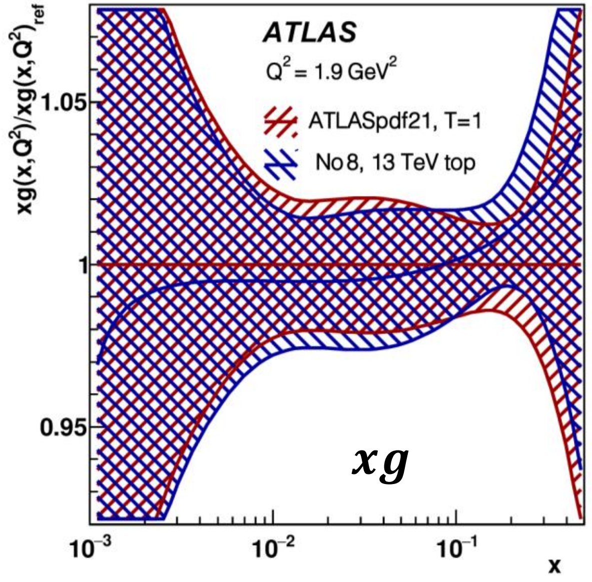
$$R_s = x(s+\bar{s})/x(\bar{u}+\bar{d})$$

Confirmed unsuppressed strange at low-x seen in previous ATLAS fits



$t\bar{t}$

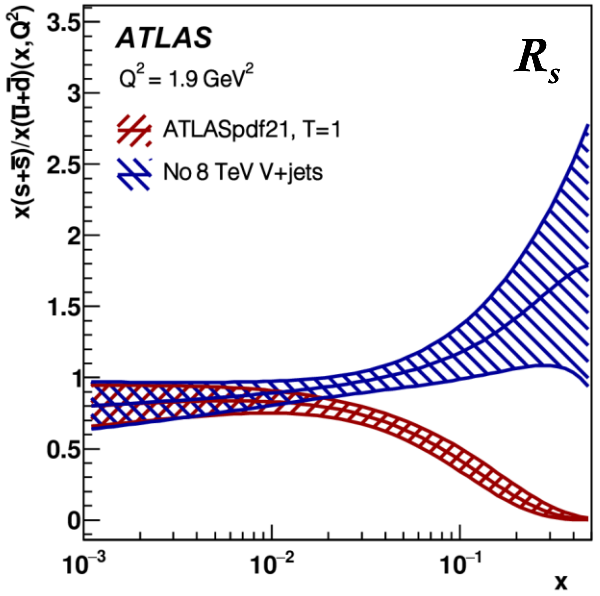
High-x gluon mildly softened when $t\bar{t}$ data added and uncertainties of high-x gluon reduced



W/Z+jets

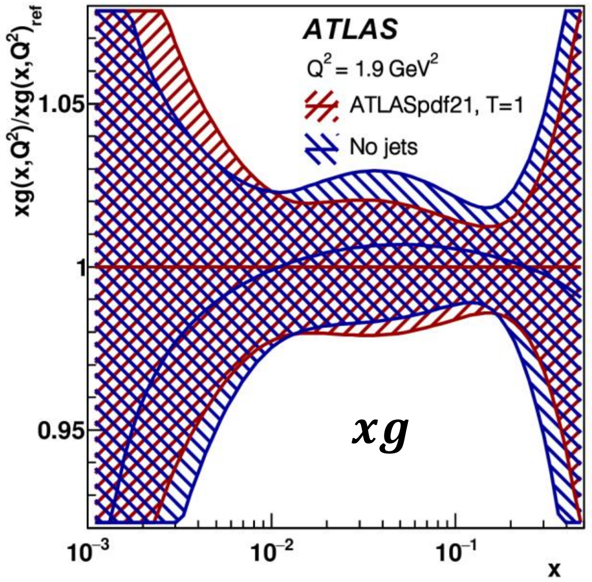
Without V+jets data little information of R_s at high x

V+jets data resolve a double minimum in the rest of data which marginally prefer $x\bar{d}\sim x\bar{s}$ (blue).V+jets data increase $\bar{d}$ and decrease $\bar{s}$ at high-x (red)



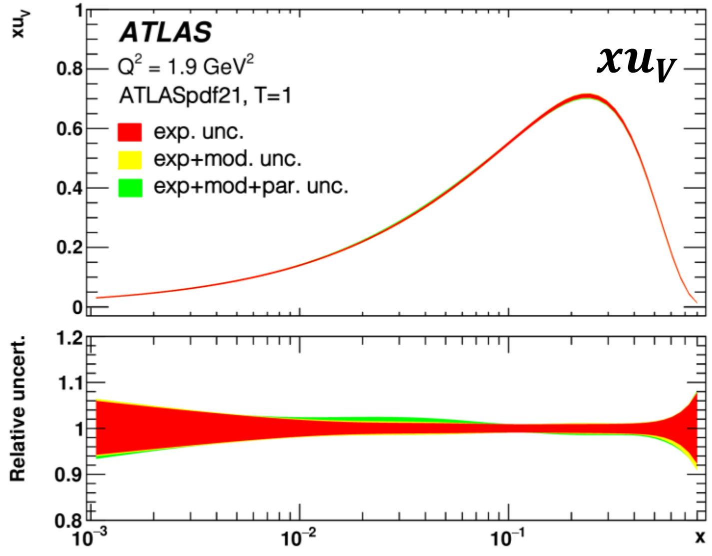
Jets

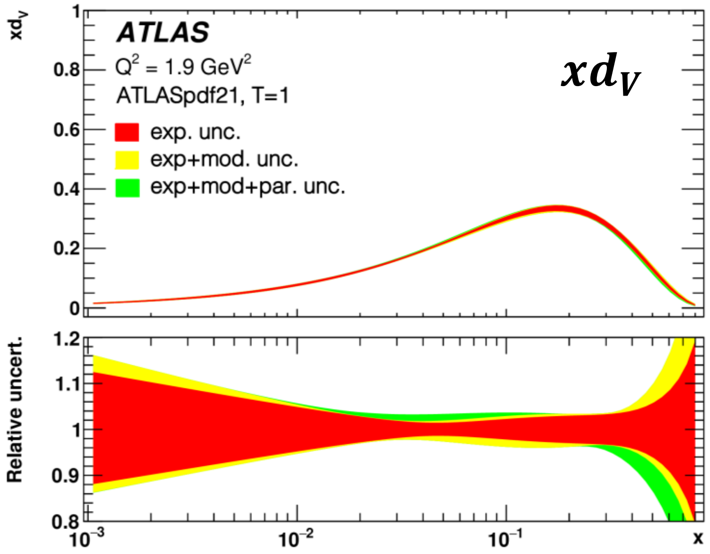
Impact of inclusive jet production (as direct photon) mostly on the high-x gluon



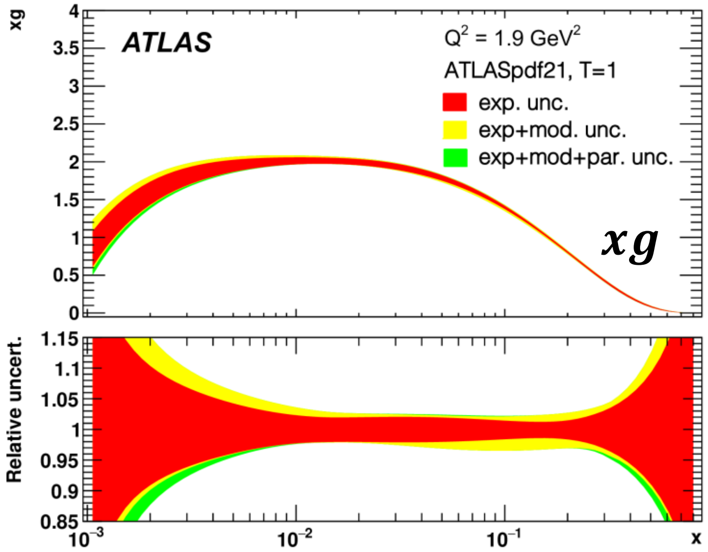
Uncertainties

Experimental uncertainties – dominant – also considered parametrisation uncertainties (extra F_{uv} and D_{dbar} parameters) and model ones (variation of Q_{min}^2 =10 GeV^2 of HERA data, starting scale Q_0^2 =1.9 GeV^2, heavy quark masses including top)





Experimental uncertainties better determined in the *u*-type sector than for *d*-type one

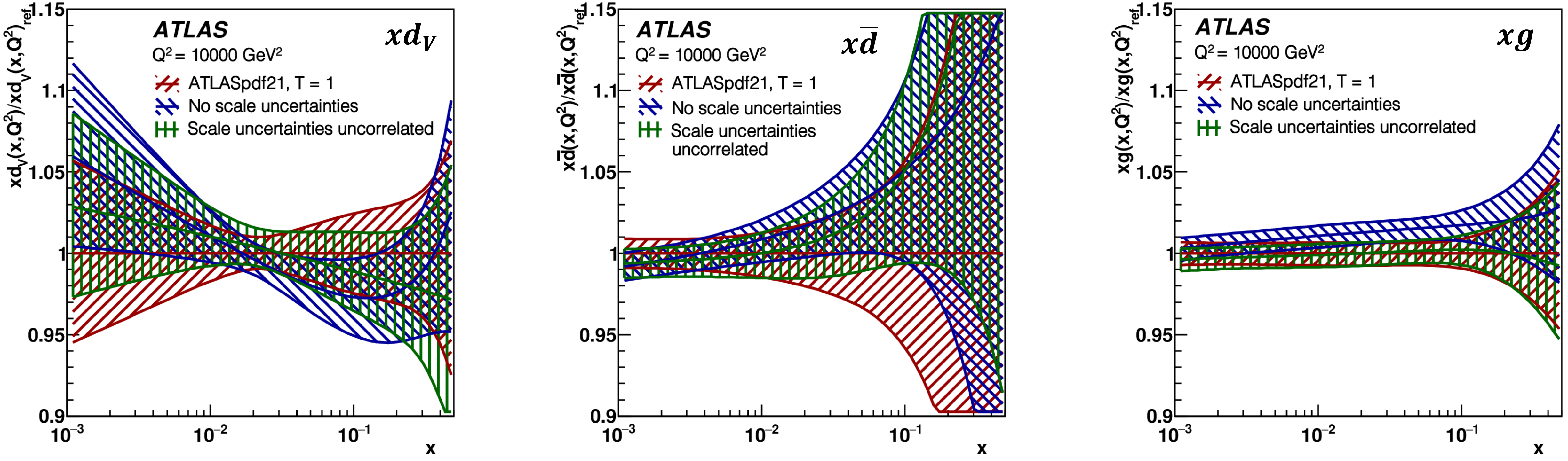


Gluon PDF well determined from $0.01 < x < 0.3$ but not at low or high-x.

Uncertainties

Also considered impact of **scale uncertainties:**

Scale uncertainties compatible with experimental ones in inclusive W/Z production (~0.5%) so they have been included in the fit



Effect of scale uncertainties not large but can be important when the aim is percent level accuracy on PDFs
Impact of scale uncertainties on other data sets also considered and found negligible

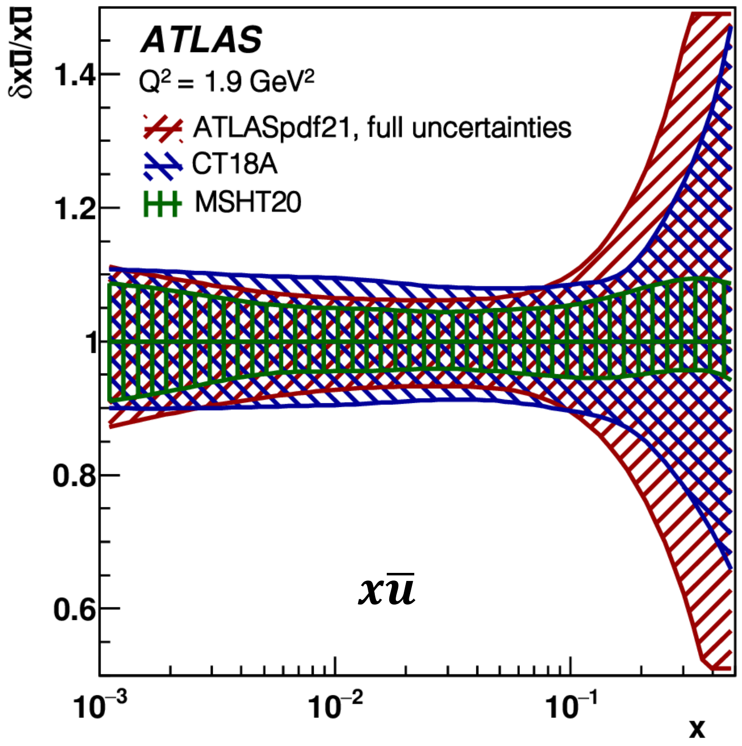
χ^2 tolerance

Experimental uncertainties in previous ATLAS+HERA fits evaluated with χ^2 tolerance $T = \sqrt{\Delta\chi^2} = 1$, but now more disparate ATLAS data

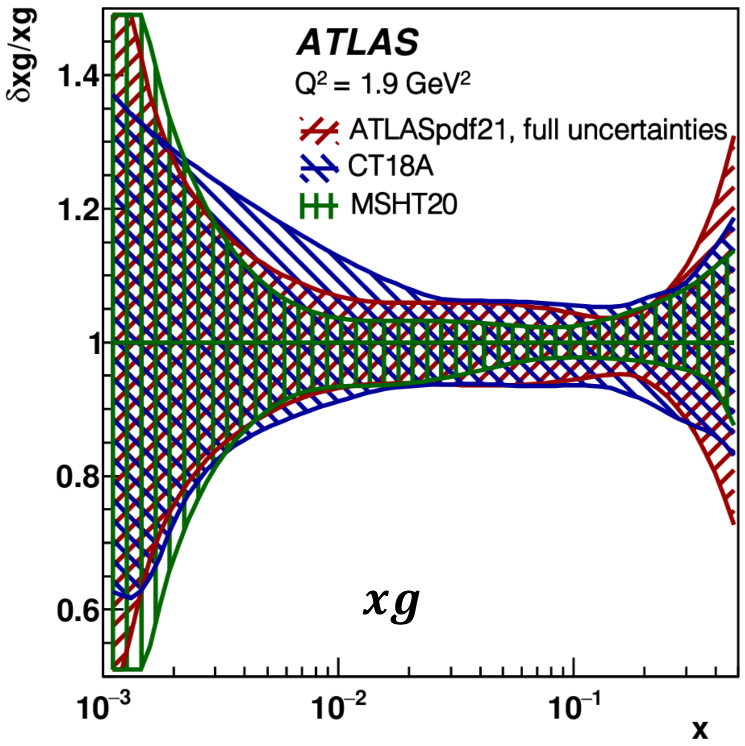
Some global PDF fitting groups (CT and MSHT) use enhanced χ^2 tolerances that are dynamic. Other groups use different methods but obtain similar uncertainties

Performed a study on an appropriate tolerance applying the procedure of the MSHT group and adopted $T = 3$

Similar uncertainties with some understood difference

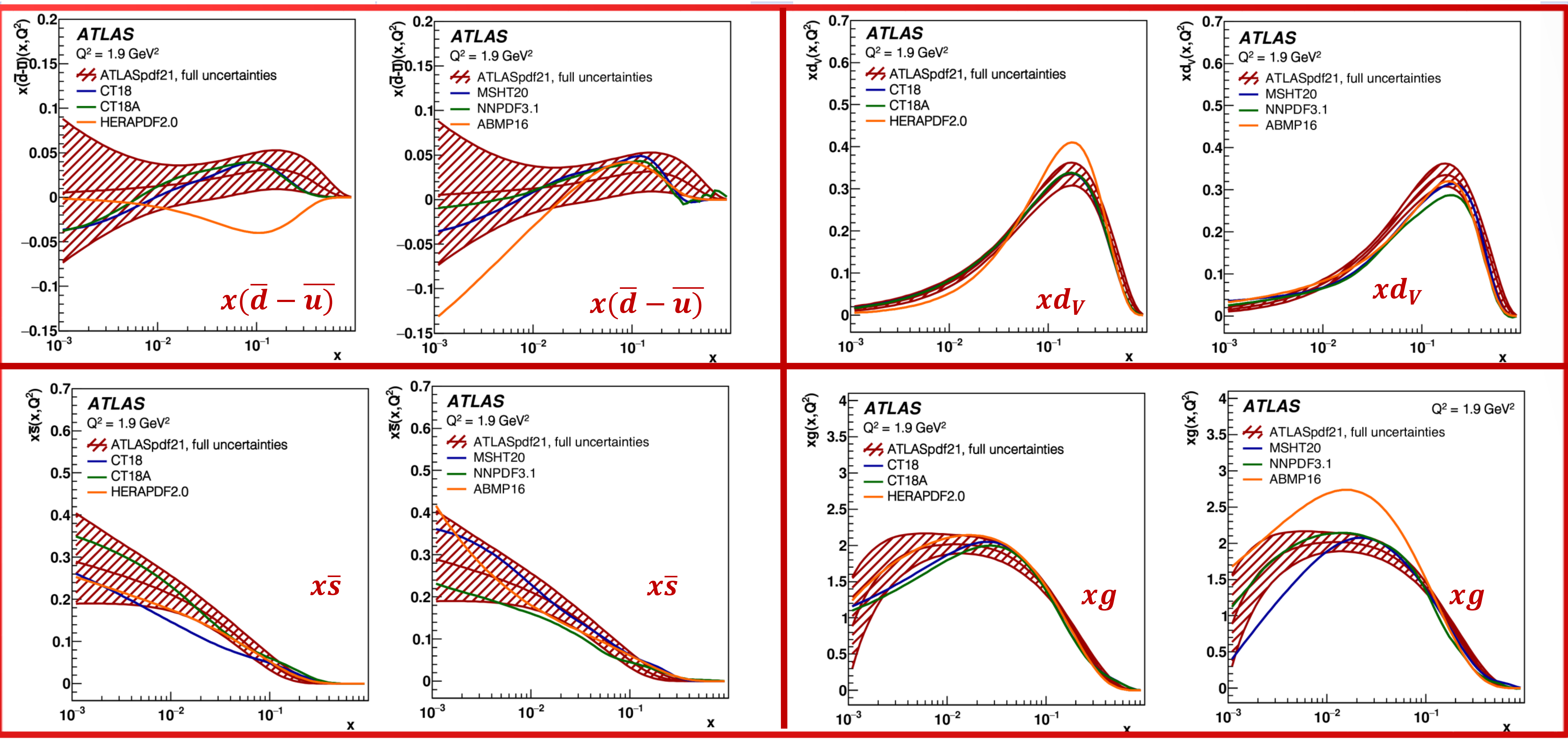


No fixed-target Drell-Yan data in ATLASpdf21



No Tevatron jet data in ATLASpdf21

Comparison with global PDFs



ATLASpdf21 fit agrees with the global fits as well as they agree with each other and ATLAS data bring the PDFs much closer to the global PDFs than to **HERAPDF2.0**

Conclusions

ATLASpdf21 fits to all PDF-sensitive ATLAS data sets that have information on correlated systematic uncertainties and NNLO QCD +NLO EW predictions:

- W/Z data allow to remove constraints on $x\bar{u}, x\bar{d}, x\bar{s}$ and confirm unsuppressed strangeness at low x
- V+jets data constraint light-sea quarks and gluon at higher x
- jets , $t\bar{t}$ and photon data constrain gluon PDF

An enhanced tolerance is used for realistic PDF uncertainty estimation and results with both choices (T=1 and T=3) made public

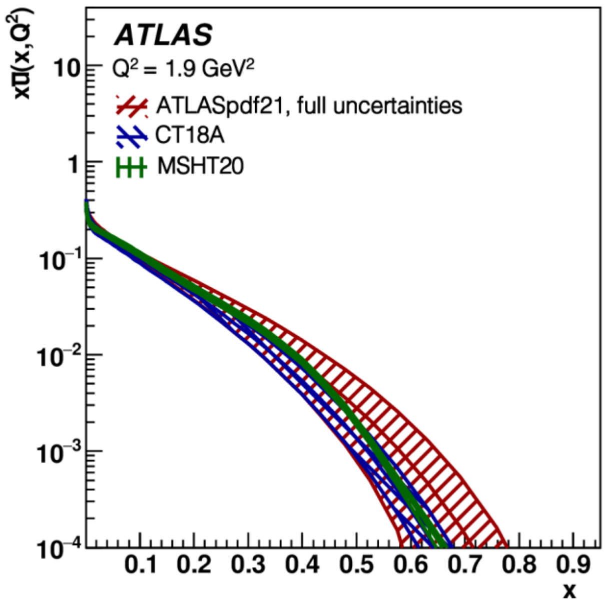
There is reasonable agreement with modern PDF sets

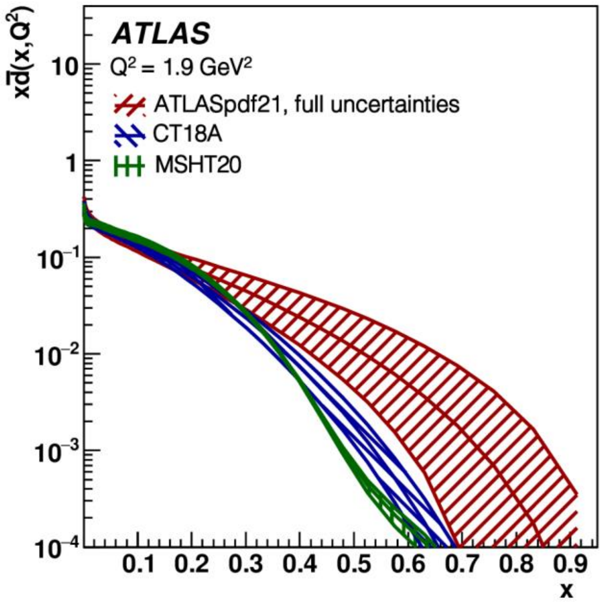
Correlations between data sets are included and information made public for global PDF fitters
Scale uncertainties considered and included when comparable to experimental uncertainties:
→Both aspects relatively small but important for future high precision measurements as m_W and $\sin^2\theta_W$

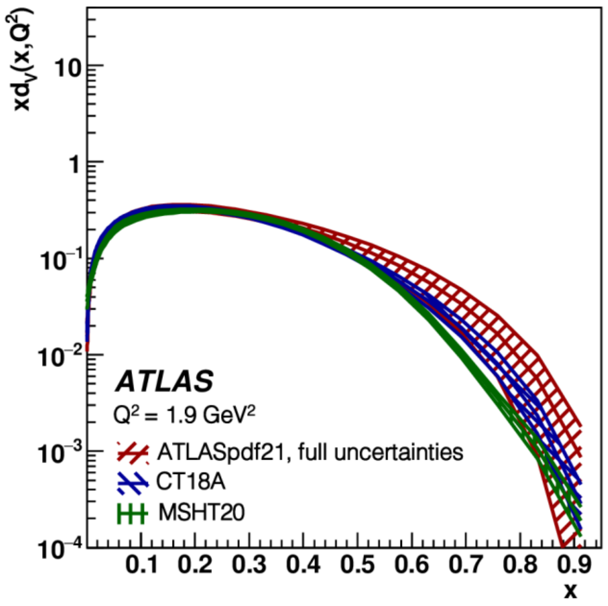
ATLASpdf21 (arXiv:2112.11266 [hep-ex]) represents a precious "vademecum" on the treatment of ATLAS data for global PDF fitting collaborations

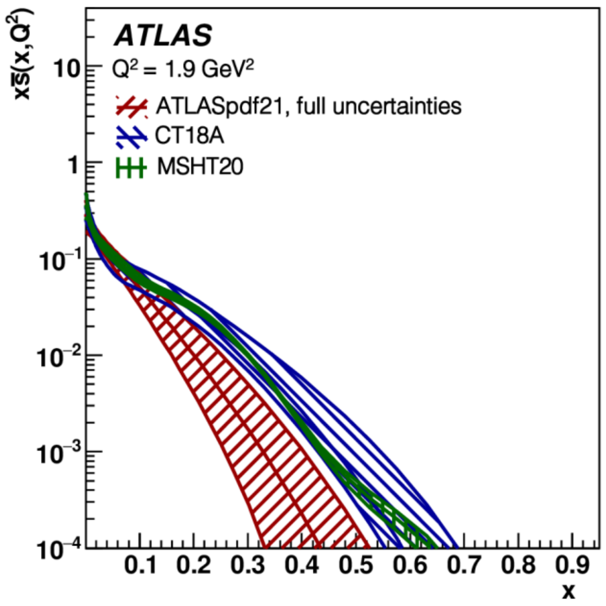
BACKUP

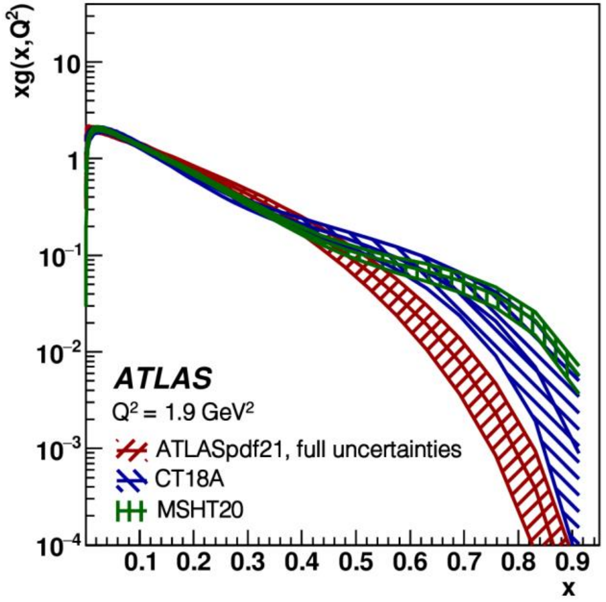
Comparison with global PDFs at high x









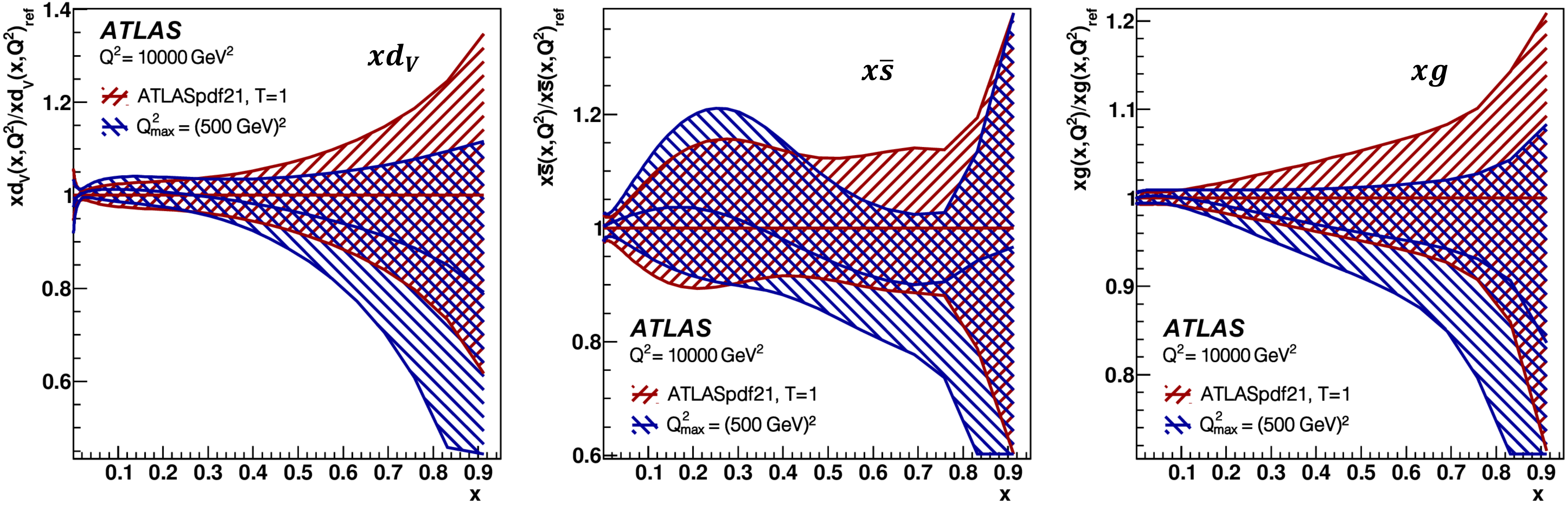


Some discrepancies for $x \geq 0.7$, where no data available to determine any of the PDFs. Nevertheless agreement within $\sim 2\sigma$.

High Q^2 data

Subtle effects of BSM physics may be present in high-scale data. If these are included in a PDF fit, then the estimates of background in searches could be distorted.

Check if the PDFs differ when we cut out possible hidden new physics in the high scale data (Q_{MAX}^2 =(500 GeV)^2)
Mainly removed jet data

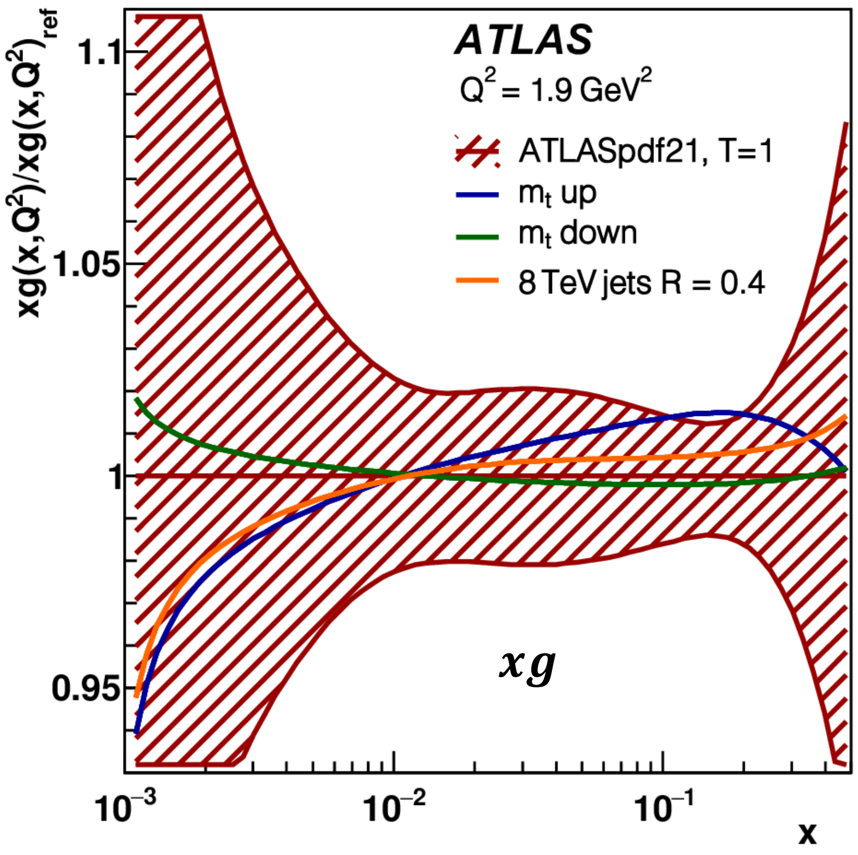


PDF shapes and uncertainties not strongly affected →maximum Q^2 cut not considered in the fit

Jet radius

Model uncertainty also includes Jet radius choice:

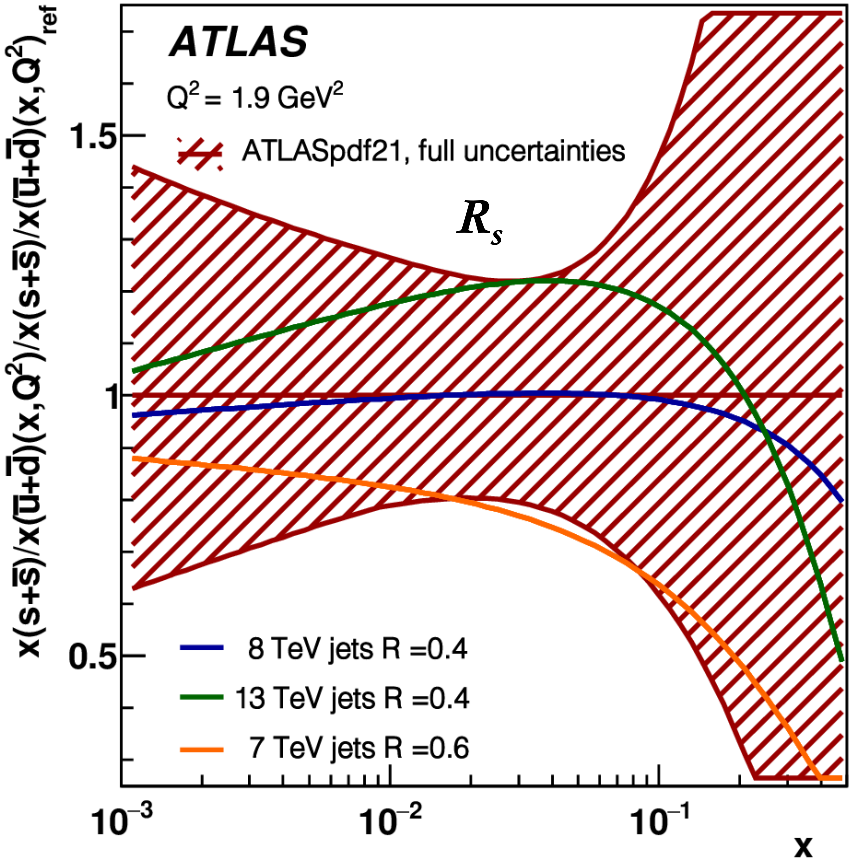
W/Z+jets and $t\bar{t}$ measurements performed with a jet radius of R=0.4 instead inclusive jets measurement performed with R=0.6. Applied a jet radius choice uncertainty considering the results of the fit obtained with R=0.4 instead of R=0.6 for inclusive jets

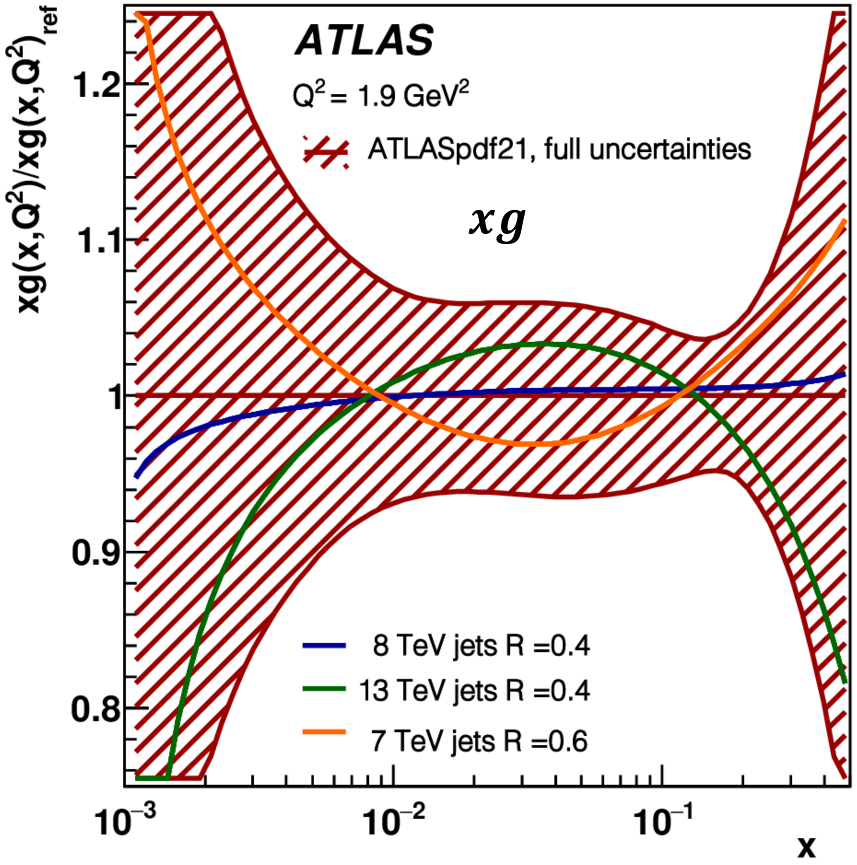


Only significant change is in the gluon PDF

Jet data

Inclusive jet data@7,8 and 13 TeV not included simultaneously in the fit since full experimental systematic uncertainty correlations between these data sets are not known





Differences in PDFs dominated by the change in centre-of-mass energy.

PDFs using 8 TeV jet data lie between those of the 7 TeV jet data and the 13 TeV jet data at 13 TeV.
Differences are not significant compared to the full PDFs uncertainties

χ^2

$$\chi^2 = \sum_{ik}\left(D_i - T_i(1-\sum_j \gamma_{ij} b_j)\right) C^{-1}_{\text{stat,uncor},ik}(D_i, D_k)\left(D_k - T_k(1-\sum_j \gamma_{kj} b_j)\right) + \sum_i \log \frac{\delta^2_{i,\text{uncor}} T_i^2 + \delta^2_{i,\text{stat}} D_i T_i}{\delta^2_{i,\text{uncor}} D_i^2 + \delta^2_{i,\text{stat}} D_i^2} + \sum_j b_j^2$$

Partial χ^2 : for each data set 1st contribution being the main, 2nd being a small bias correction

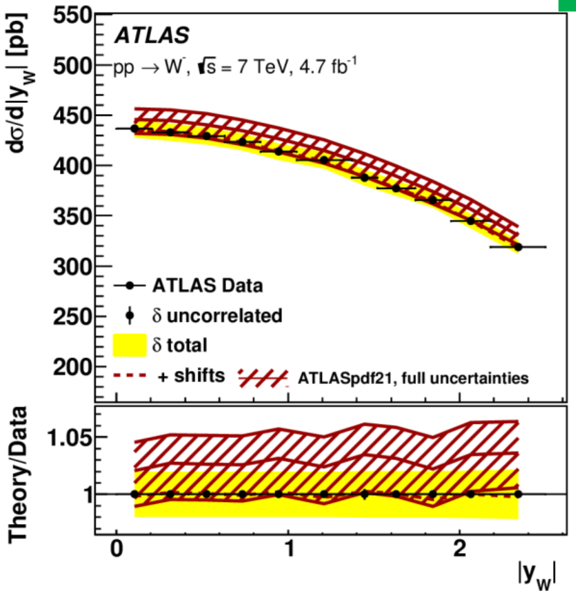
Correlated term between data sets

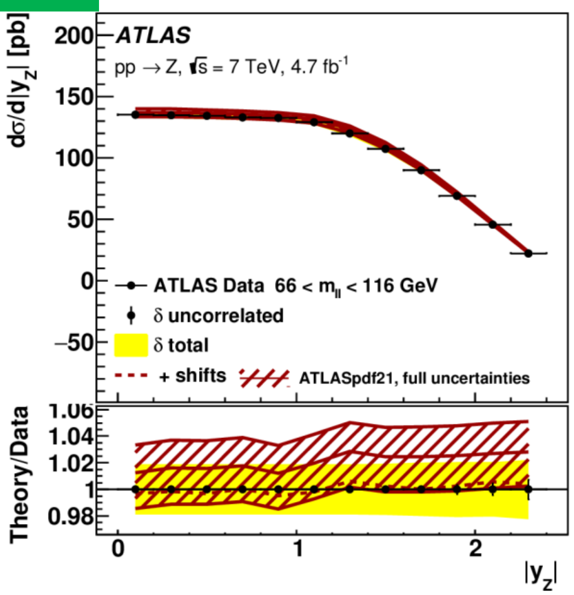
D_i = measured data; T_i= theory prediction;
γ_{ij} correlated systematics accounted for using nuisance parameters b_j
$C_{stat,uncor,ik}$ = statistical (plus uncorrelated) covariance matrix;
$\delta_{i,uncor}$=uncorrelated systematics uncertainties;
$\delta_{i,stat}$= statistical uncertainties

Total χ^2/NDF	2010/1620
HERA χ^2/NDP	1112/1016
HERA correlated term	50
ATLAS W, Z 7 TeV χ^2/NDP	68/55
ATLAS Z/γ^* 8 TeV χ^2/NDP	208/184
ATLAS W 8 TeV χ^2/NDP	31/22
ATLAS W and Z/γ^* 7 and 8 TeV correlated term	71 = (38 + 33)
ATLAS direct γ 13/8 TeV χ^2/NDP	27/47
ATLAS direct γ 13/8 TeV correlated term	6
ATLAS V+jets 8 TeV χ^2/NDP	105/93
ATLAS $t\bar{t}$ 8 TeV χ^2/NDP	13/20
ATLAS $t\bar{t}$ 13 TeV χ^2/NDP	25/29
ATLAS inclusive jets 8 TeV χ^2/NDF	207/171
ATLAS V+jets 8 TeV and $t\bar{t}$ + jets 8,13 TeV and R = 0.6 inclusive jets 8 TeV correlated term	87 = (16 + 9 + 21 + 41)

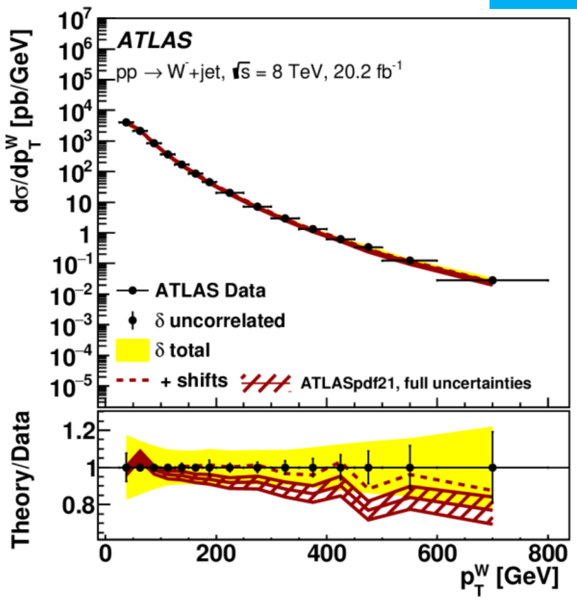
Data sets

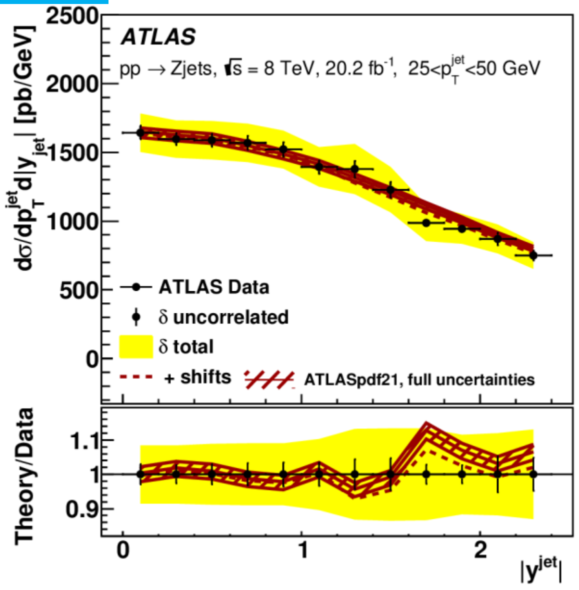
W/Z



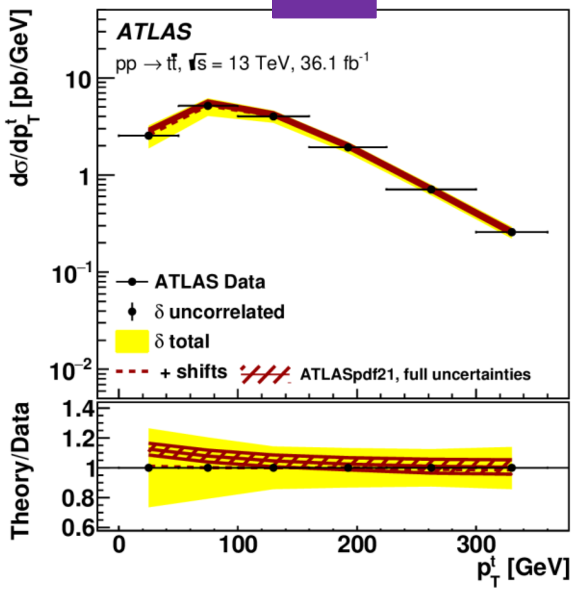


W/Z+jets

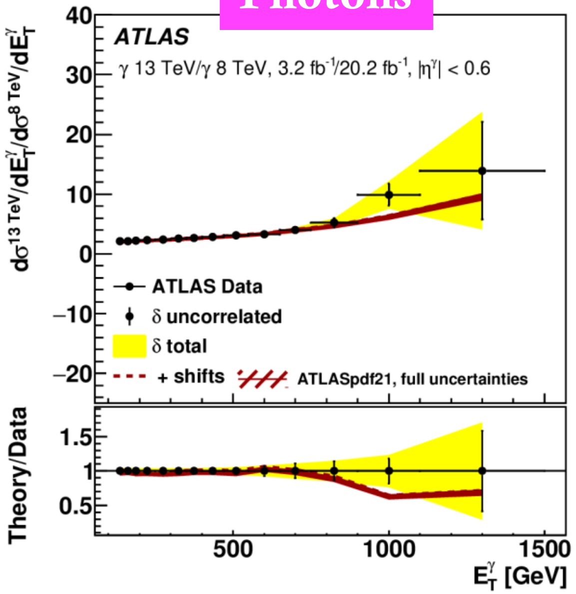




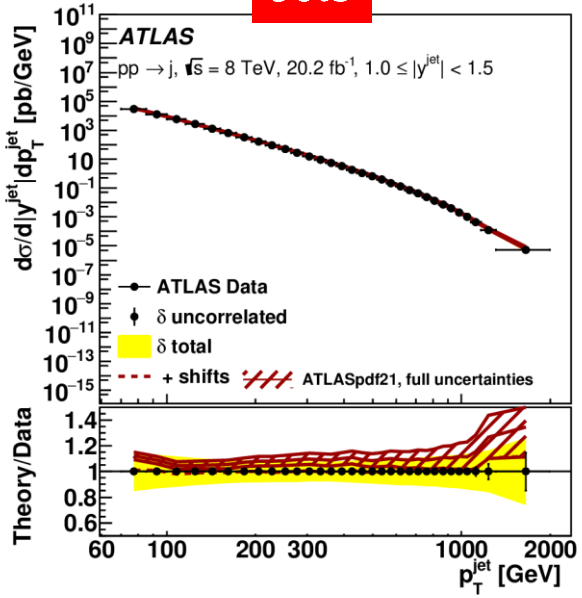
$t\bar{t}$



Photons

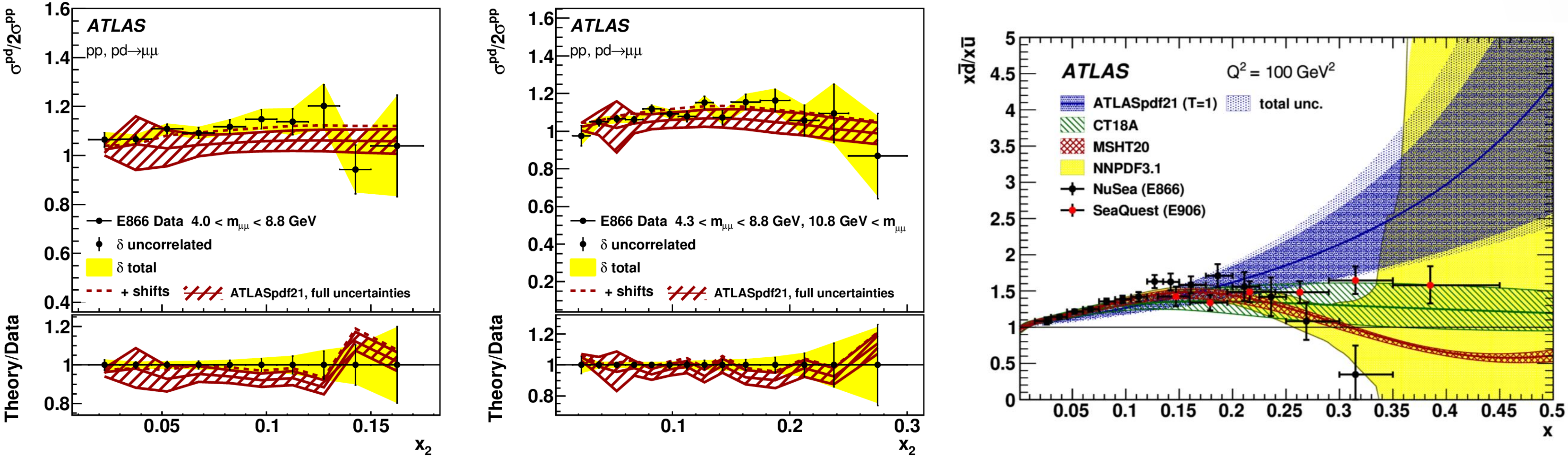


Jets



Fixed-target Drell-Yan data

Fixed-target Drell–Yan data from E866 and E906, uniquely able to constrain the difference $x(\bar{d}\text{-}\bar{u})$ at high x



Quite good description of the E866 data, even better agreement with the new data from E906