PDFs@LHC

BadHoneff October2023

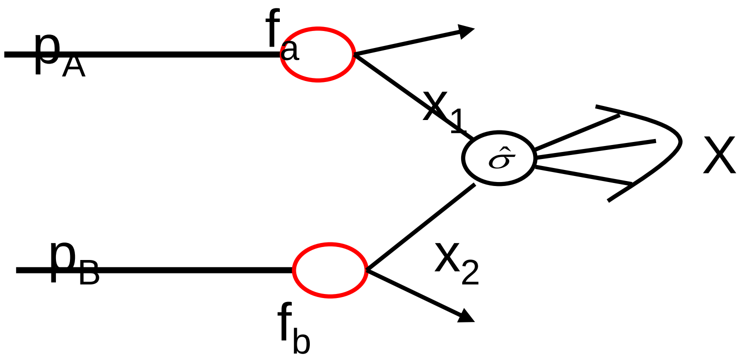
A M Cooper-Sarkar

PDFs from ATLAS
PDFs from CMS
PDFs using LHC data
from CT, MSHT, NNPDF, ABMP, PDF4LHC

PDFs at the LHC

LHC cross sections are calculated as follows

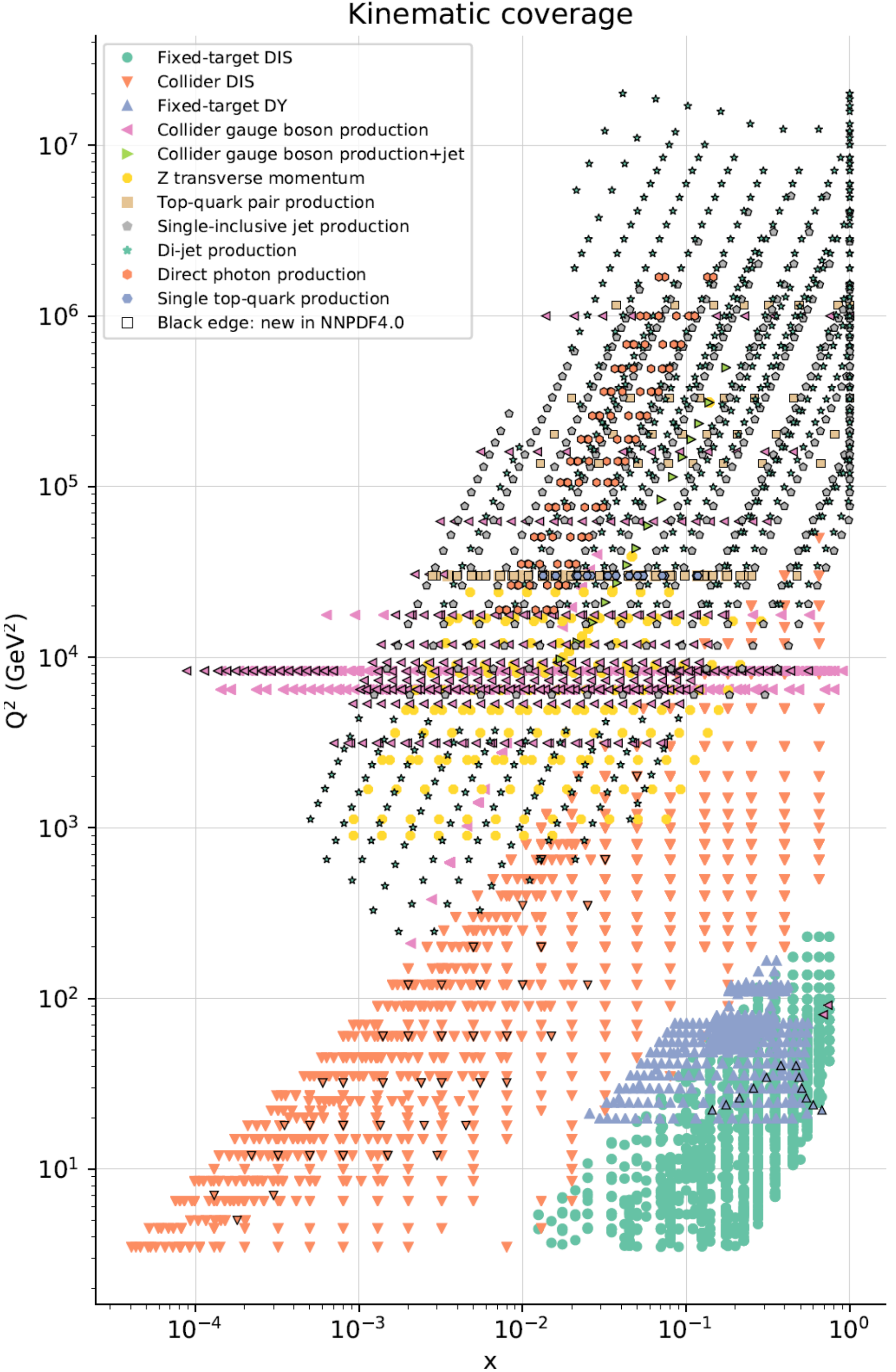
$$\sigma_X = \sum_{a,b} \int_0^1 dx_1 dx_2\, f_a(x_1, \mu_F^2)\, f_b(x_2, \mu_F^2)$$
$$\times\ \hat{\sigma}_{ab\to X}\left(x_1, x_2, \{p_i^\mu\}; \alpha_S(\mu_R^2), \alpha(\mu_R^2), \frac{Q^2}{\mu_R^2}, \frac{Q^2}{\mu_F^2}\right)$$



where X could be a Standard Model process like Drell-Yan, W, Z, production etc or could be a new physics process.

If it is an SM process then the cross sections can be used to improve our knowledge of PDFs

If it is a BSM process then the uncertainties on Parton Distribution Functions (PDFs) will impact how accurately we measure the new cross section



The HERA data are the 'backbone of all PDF fits BUT what could HERA not do?

High-x gluon and sea flavour detail s,c

What other data can we use?

- Drell-Yan data from fixed target DIS and the Tevatron and LHC
- W,Z rapidity spectra from Tevatron and LHC
- Jet pT spectra from Tevatron and LHC
- Top-anti-top differential cross-sections from LHC
- W and Z +jet spectra, or Z pt spectra from LHC
- W and Z +heavy flavours from LHC
- Beware: IS the factorisation theorem proven?-only for DY!
- Beware: there may be new physics at high scale that we 'fit away'
- Further warning, this additional information comes from many different groups– often there is no clarity on the correlations of experimental systematic uncertainties between differing LHC measurements

ATLASpdf21(arXIV:2112.11266)

ATLASpdf21 is a PDF fit to HERA combined data and to diverse PDF sensitive ATLAS data sets which have full information on correlated systematic uncertainties and for which NNLO QCD +NLO EW predictions are available

Key Features are:

- Experimental systematic correlations are applied **not only within data sets but also between data sets** with common systematic sources and information is made public
- Scale uncertainties are considered
- Flexible parametrisation We extend our parametrisation from 14 to 21 parameters since the ATLAS data allow us to release some constraints, in particular we allow ubar ≠ dbar as $x \to 0$ and sbar, dbar, ubar can all have differing shapes as $x \to 0$. We also allow more freedom at mid-x in valence and gluon. This achieves a **good description of data not in the fit** such as Tevatron W,Z data and pD/pp Drell-Yan data from E866 and E906– but we have much more control over correlated systematic uncertainties
- We consider enhanced tolerance for the final estimate of experimental uncertainties. This becomes necessary with many diverse data sets. We consider how far the PDF parameters can change before any data set is pushed outside its 68%CL, this results in $\Delta\chi^2 = T^2$, T=3. As with HERApdf we also include model/parametrisation uncertainties
- There is fair agreement with modern PDF sets CT, MSHT, NNPDF and a better fit to ATLAS data

ATLASpdf21 is a fit to many different sorts of ATLAS data:

- W,Z boson production
- W,Z boson +jets
- T-tbar data
- Direct photon production
- Inclusive jet production

These are the processes which are known to have PDF sensitivity

It is a fit to NNLO in QCD and NLO in EW achieved either by direct NNLO grids or by NLO grids + k-factor corrections

The HERA data provided the backbone of the fit (as with ALL modern PDF fits) and 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 come from ATLAS- in addition to further constraints on all PDFs x> 0.001

Data set	$\sqrt{s}$ [TeV]	Luminosity [fb^{-1}]	Decay channel	Observables entering the fit
Inclusive $W, Z/\gamma^*$ [9]	7	4.6	e, μ combined	η_l (W), y_Z (Z)
Inclusive Z/γ^* [13]	8	20.2	e, μ combined	$\cos\theta$ in bins of $y_{\ell\ell}, M_{\ell\ell}$
Inclusive W [12]	8	20.2	μ	η_μ
$W^\pm$ + jets [23]	8	20.2	e	p_T^W
Z + jets [24]	8	20.2	e	p_T^{jets} in bins of $\|y_{\text{jets}}\|$
$t\bar{t}$ [25, 26]	8	20.2	lepton + jets, dilepton	$m_{t\bar{t}}, p_T^t, y_{t\bar{t}}$
$t\bar{t}$ [15]	13	36	lepton + jets	$m_{t\bar{t}}, p_T^t, y_t, y_{t\bar{t}}$
Inclusive isolated γ [14]	8, 13	20.2, 3.2	-	E_T^γ in bins of η^γ
Inclusive jets [16–18]	7, 8, 13	4.5, 20.2, 3.2	-	p_T in bins of $\|y_{\text{jets}}\|$

Correlation of systematic sources between data sets

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

Entries in the same row are considered 100% correlated for the central fit

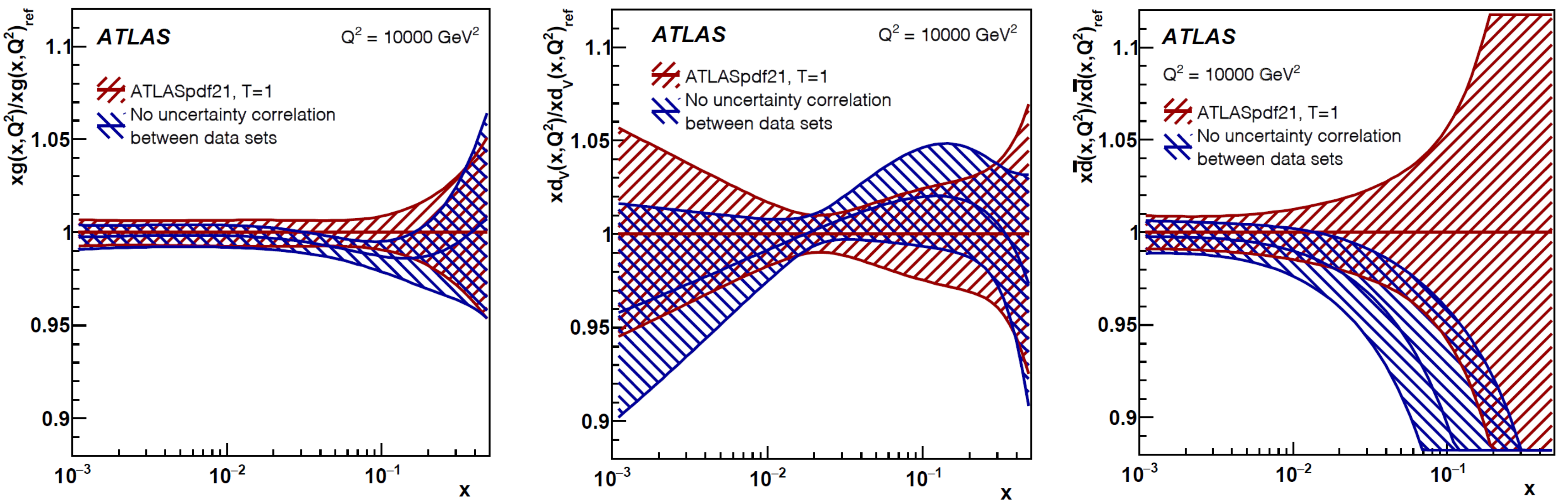
Cross checks are made of alternative degrees of correlation for inclusive jets since jet radius R=0.6 is used for these, rather than R=0.4 which is used for the other for the other data sets: V+jets and t-tbar in lepton+jets channel.

Note these mostly involve the JetEnergyScale systematic correlations, which are the largest, lepton systematics are much smaller

Effect of correlations between data sets

Lets look at a scale relevant for LHC physics and focus on the middling x range where W,Z and Higgs are produced

gluon ratio **$d_{valence}$ ratio** **$\bar{d}$ ratio**



The $\chi 2$ of the fit is 30 units better when correlations are included
The difference in PDFs is small for the gluon
But can be larger in the d-quark sector
Remember the goal for PDF precision is ~1% for M_W and $\sin^2\theta_W$ measurements if BSM effects are to be seen by the deviations of these parameters from their SM values

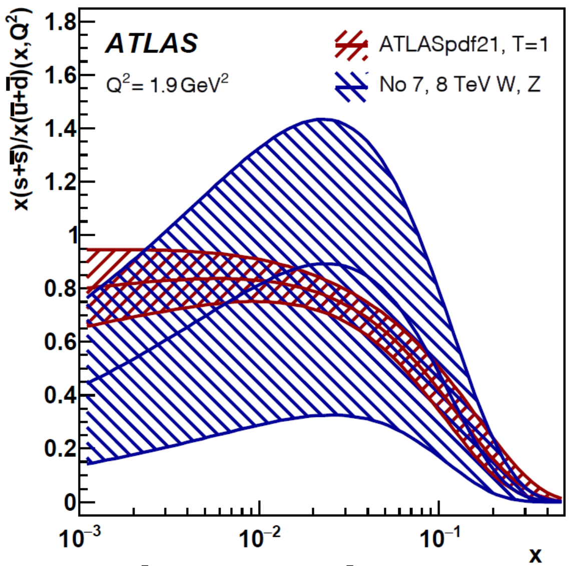
Correlations can be important

Investigate the impact of each class of data

Impact of inclusive W,Z production data

NOTE: these plots show experimental uncertainties with $\Delta\chi^2 = 1$

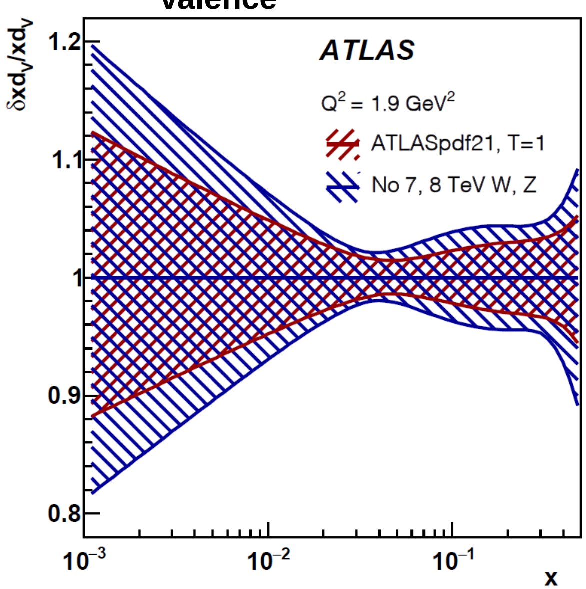
strange/light quarks



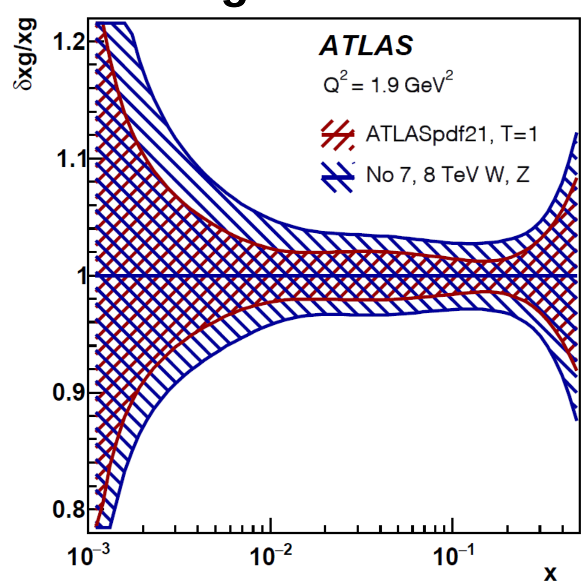
Inclusive W,Z production data at 7 and 8 TeV are removed from the fit
Without these data the ratio of strange to light quarks is very poorly determined

W,Z data also reduce the uncertainties of the valence quarks and the gluon considerably

$d_{valence}$ ratio

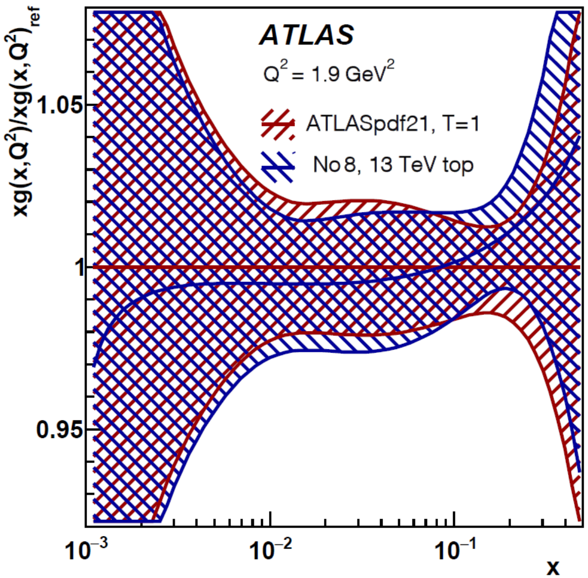


gluon ratio

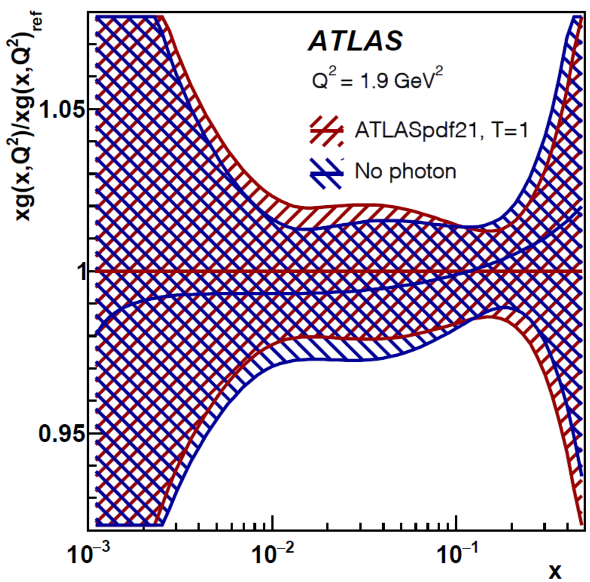


Impact of t-tbar production, direct photon production, and inclusive jet production …………….. Is mostly on the high-x gluon

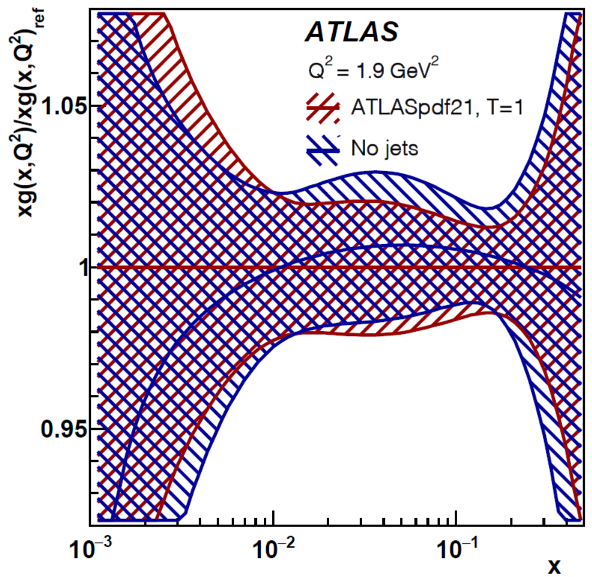
T-tbar production



Direct photon



Inclusive jets



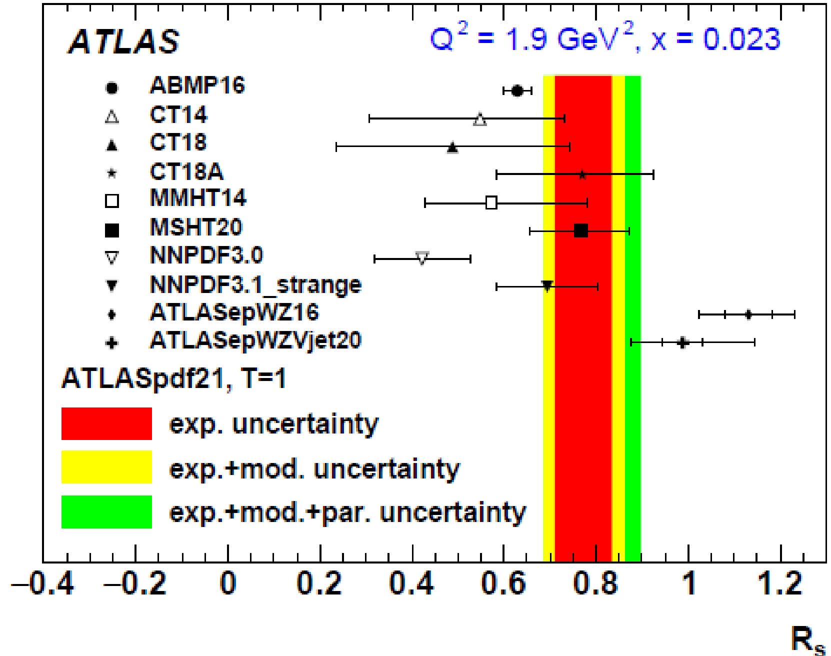
The gluon PDF uncertainty with/without these data is shown to illustrate the slight pulls on the shape and the more significant decrease in uncertainty—mostly from inclusive jets

Strangeness

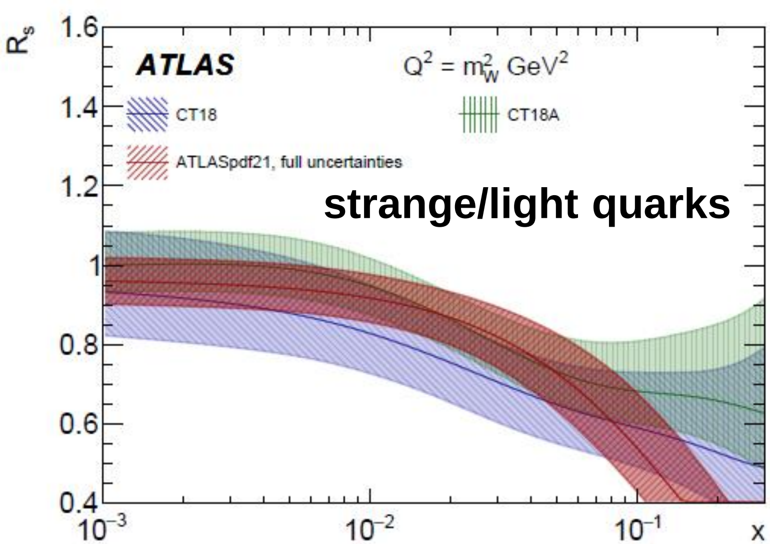
The information on strangeness has often been presented at a single x,Q^2 point and compared to the result of global PDFs

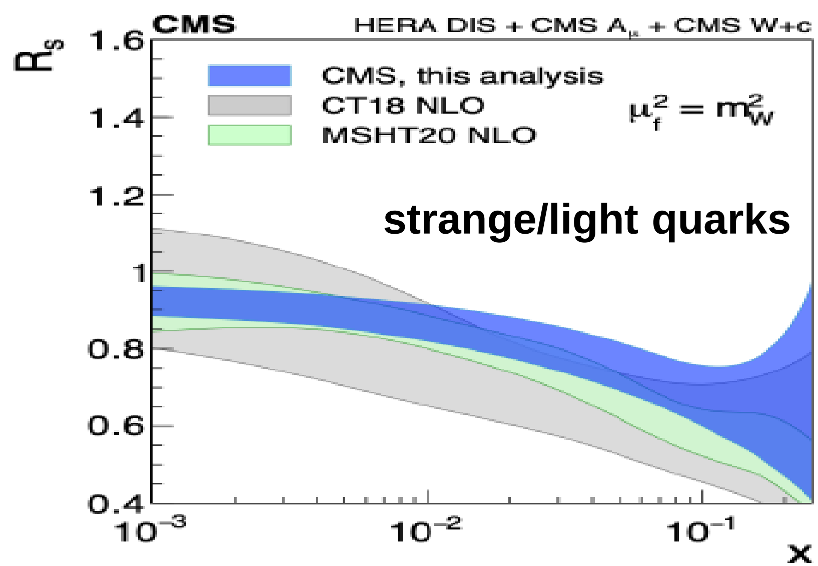
Note that older PDFs CT14, MMHT14,NNPDF3.0 all had Rs~0.5 at low scale (Q^2=1.9GeV^2) **BUT** this has moved up to ~0.8 for CT18A, MSHT20, and NNPDF3.1_strange after ATLAS W,Z 7 TeV data was included (not for CT18 which does not include these data)

ATLAS older fits had Rs~1.0 and have moved down to Rs~0.8 due to input of new data, V+jets and W,Z 8 TeV and greater flexibility of low-x parametrisation

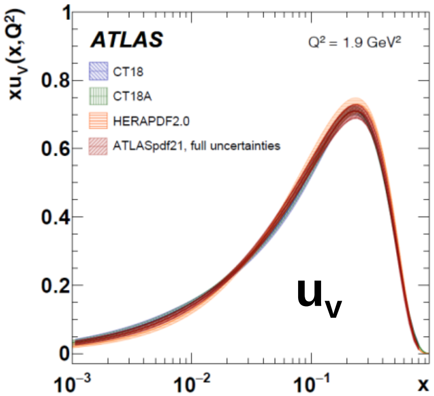


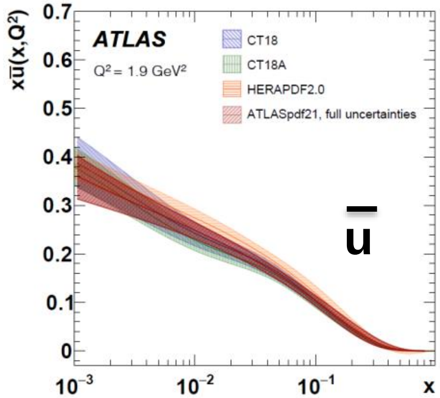
See the more interesting shape of the strangeness ratio at higher scale m_W^2 compared to CT18 and CMS

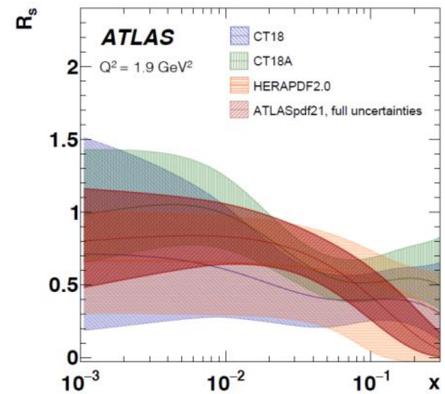


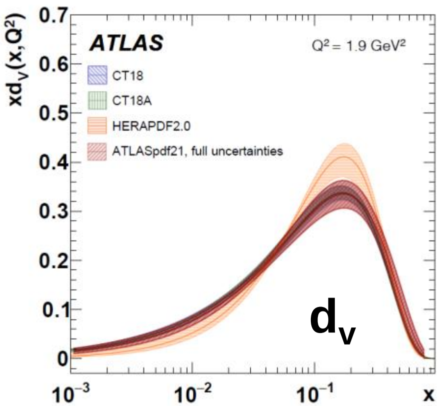


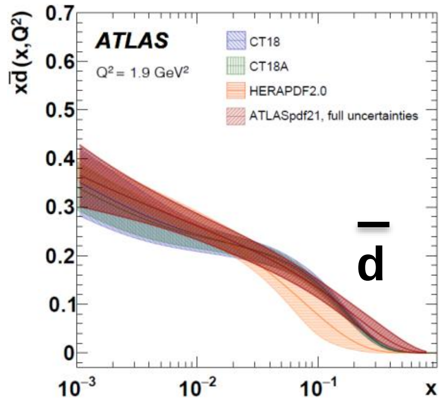
Now compare ATLASpdf21 PDFs to modern global PDFs

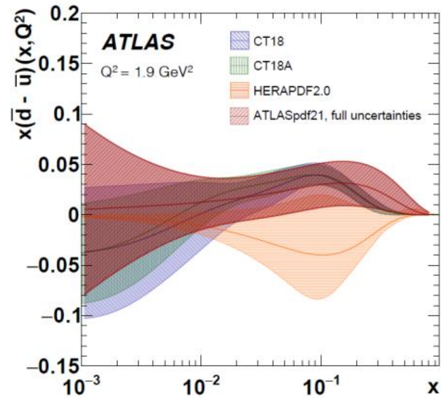


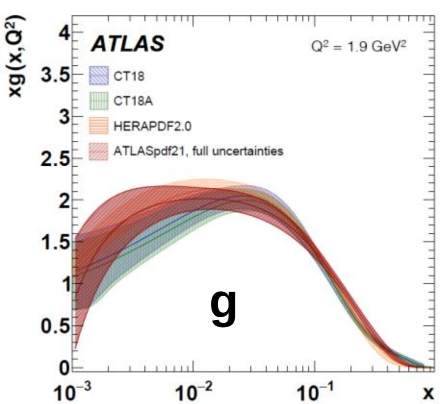


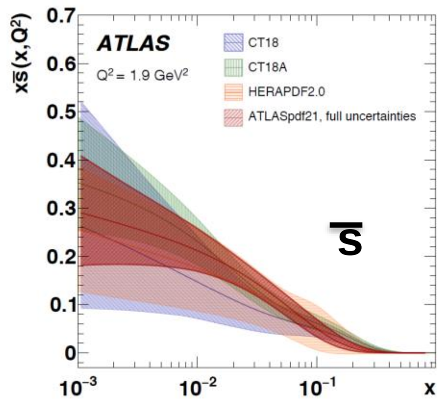








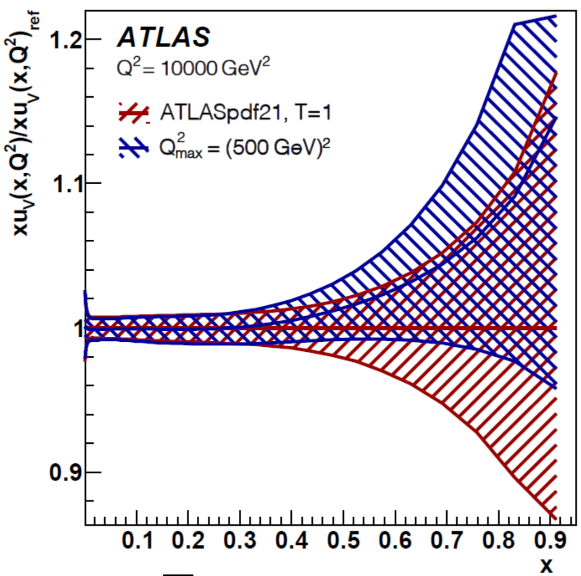




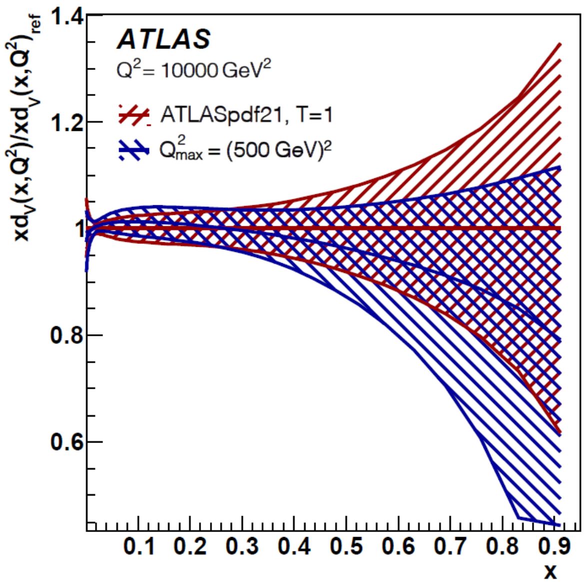
NNPDF, MSHT, ABMP shown in the paper,
ATLAS PDFs agree with the other PDFs as well as the other PDFs agree with each other!

A fit was made cutting data for which the scale > 500 GeV, to check if the PDFs differ if we cut out possible hidden new physics in the high scale data

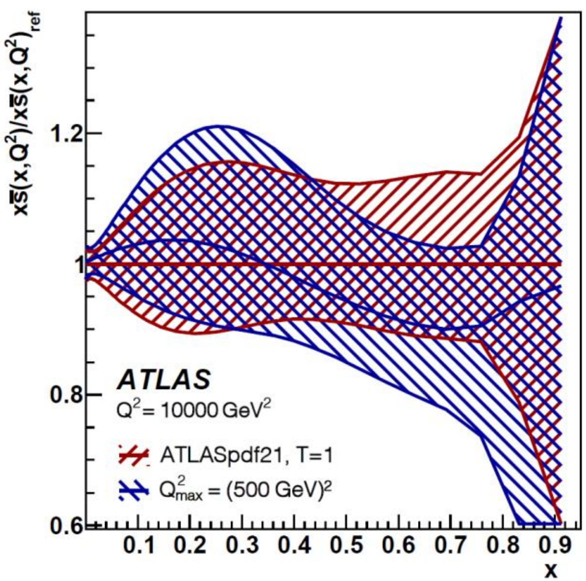
$u_{valence}$ ratio



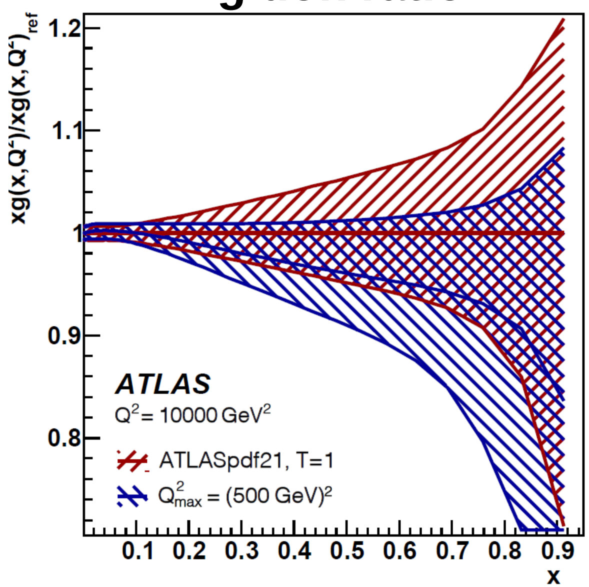
$d_{valence}$ ratio



$\bar{s}$ ratio

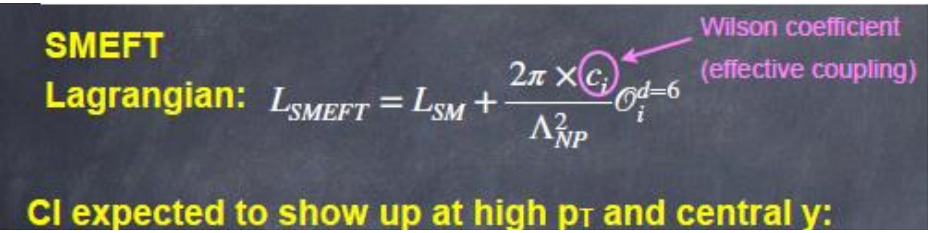


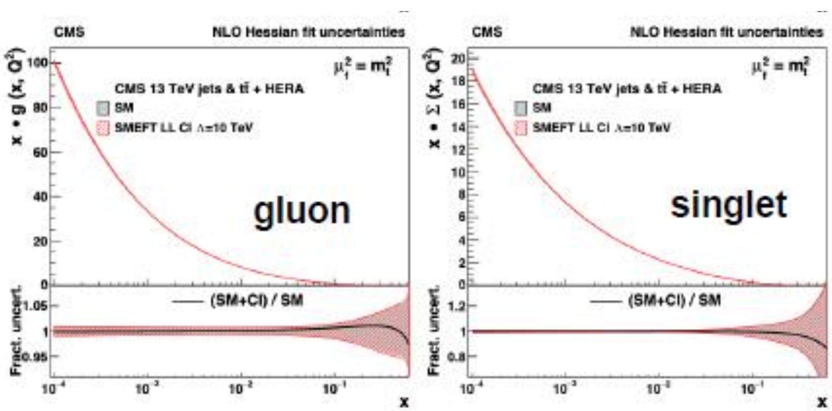
gluon ratio

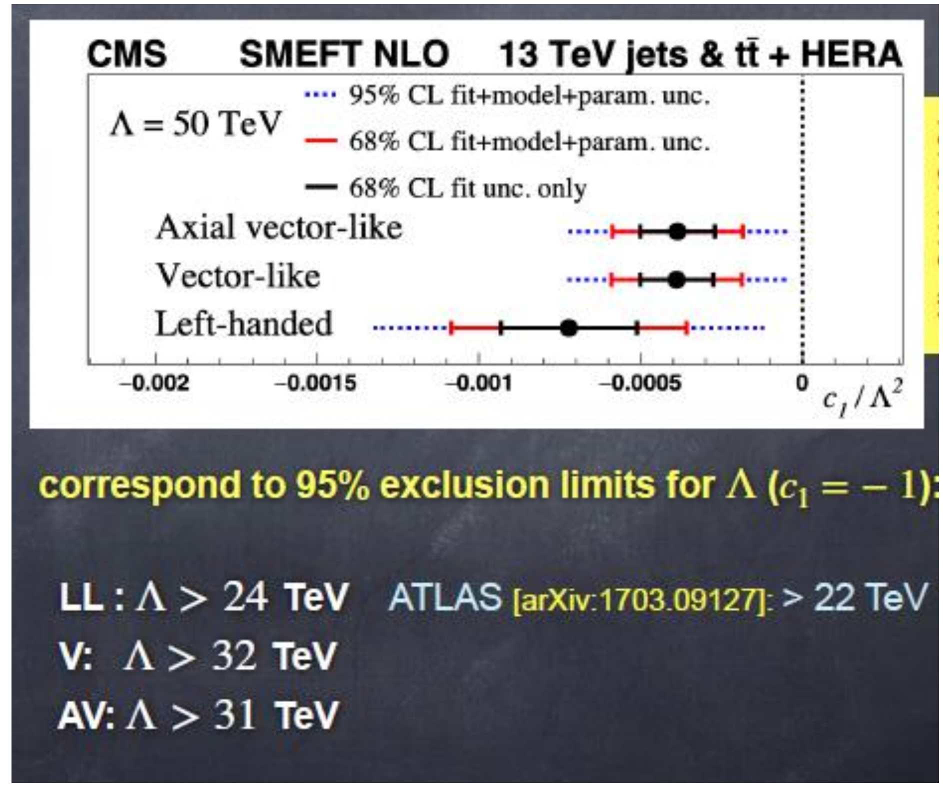


This cut mostly removes inclusive jet production data.
The effect is only seen at high x –note linear x scale- PDFs are not significantly changed
These changes would barely show up on our usual log scale in x

This is corroborated by a **CMS** study arXiv:2111.10431 in which parameters for BSM physics are fitted simultaneously with PDF parameters (and m_{top} and $\alpha_S(M_Z)$) using Jet data and top-antitop production data







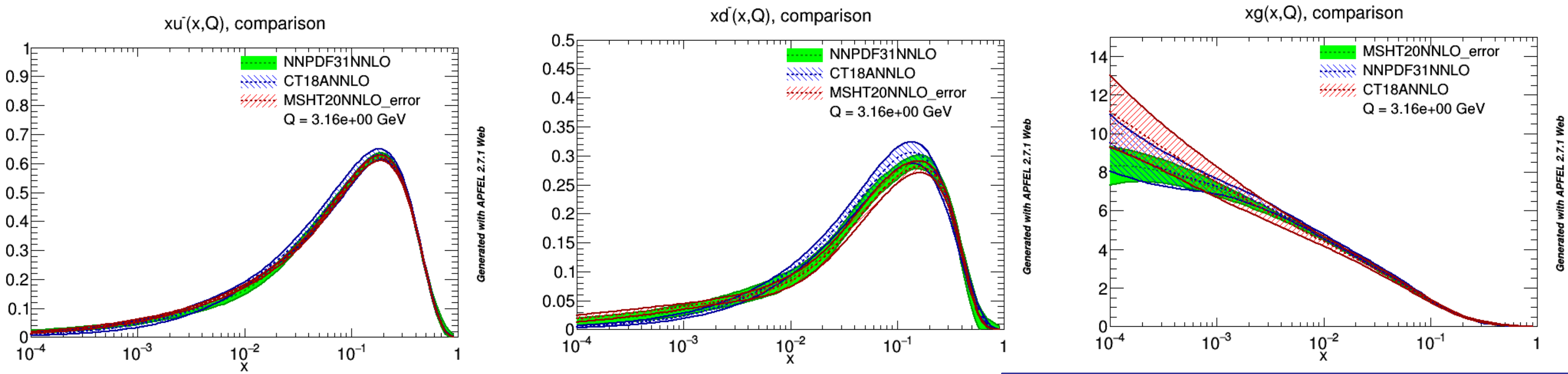
PDFs almost unchanged with or without the BSM term

CMS do not produce PDFs but rather PDF studies– they usually combine a new data set (or maybe two) with the HERA data to study its impact. They do not combine many data sets, or assess correlations between them—(to date)

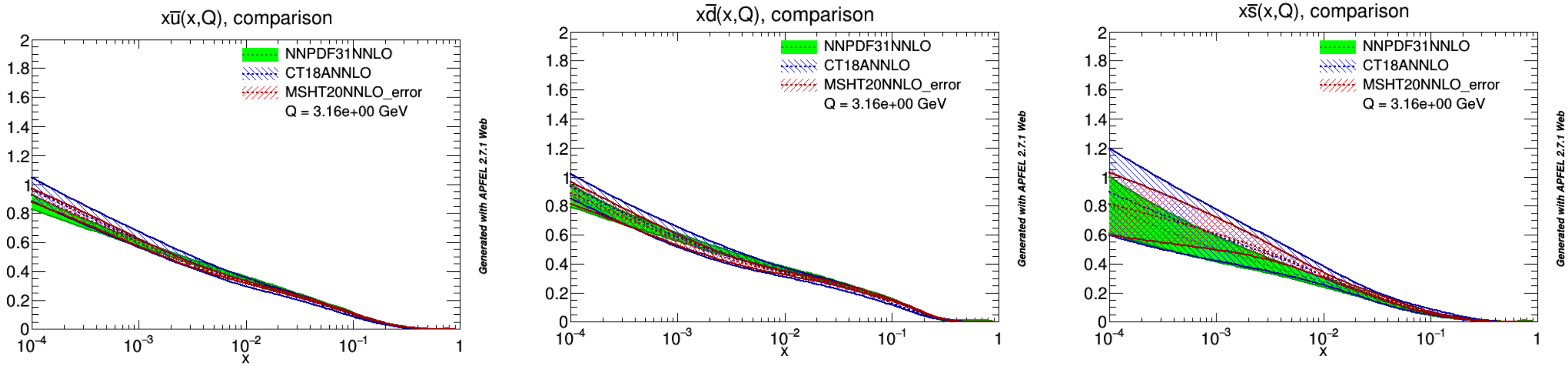
BSM parameters close to zero

NOTE study is NLO in pQCD

Let's see how we are doing--- PDF comparisons at NNLO in pQCD

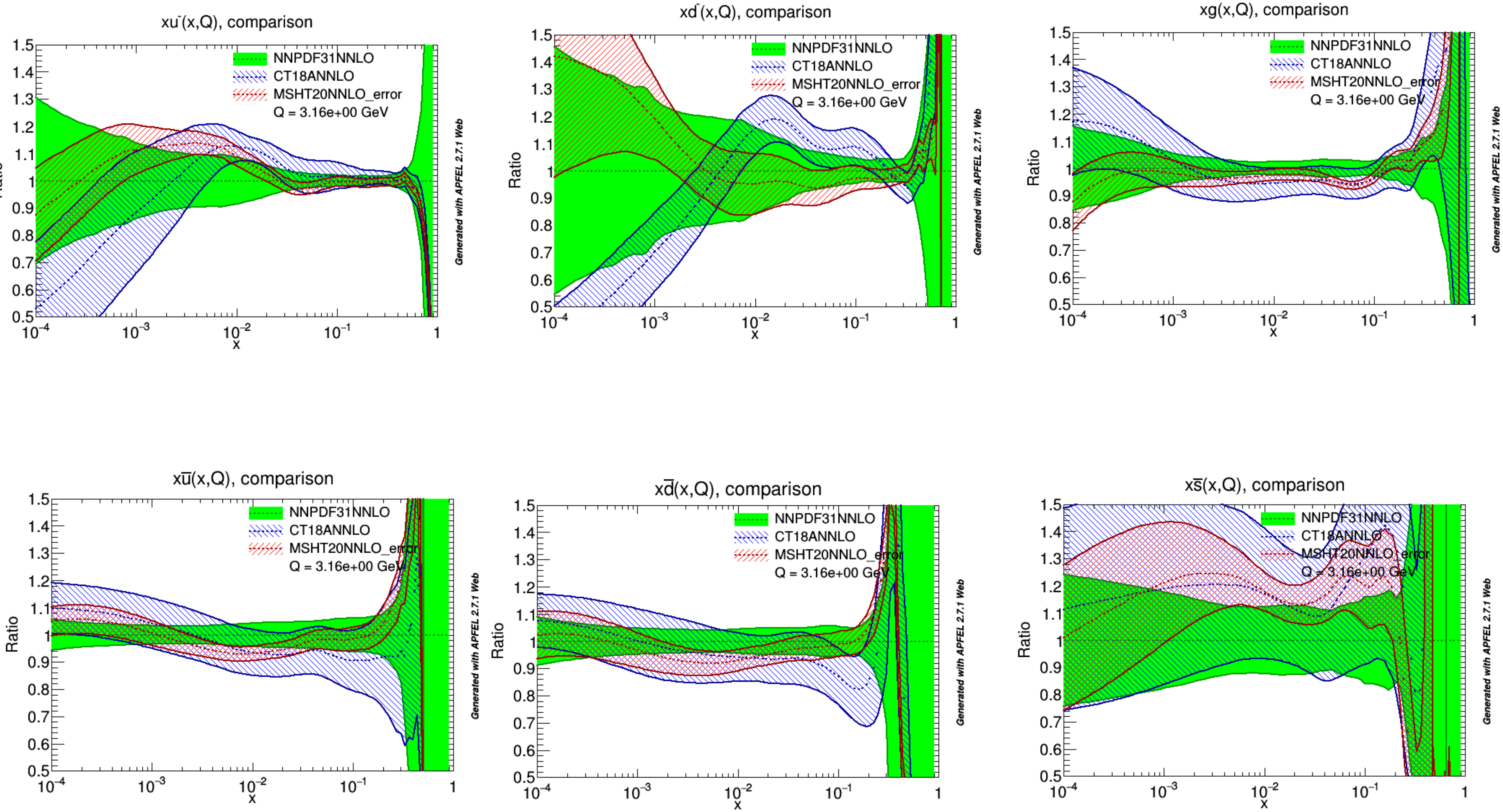


I will come to NNPDF4.0 later



Several groups extract PDFs and there are significant differences because of slightly different model choices:

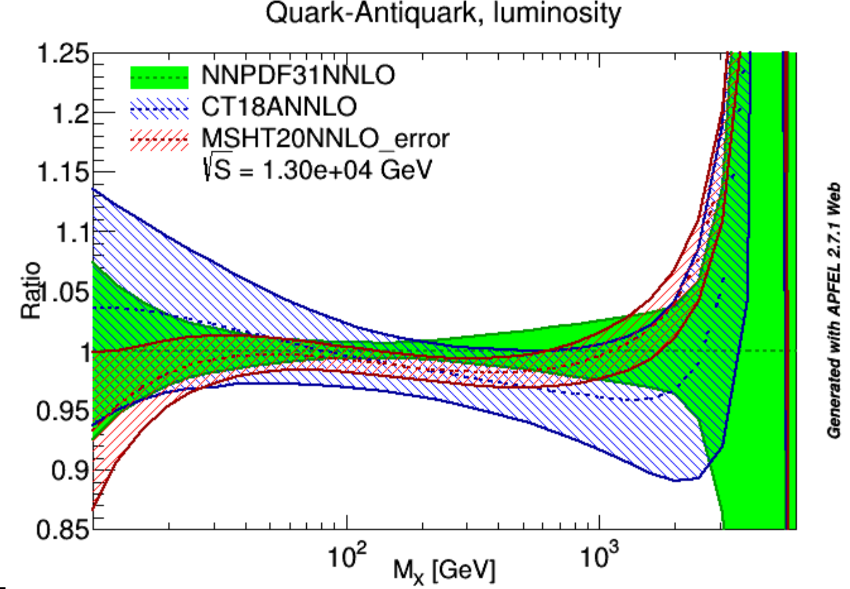
- Exact choice of data entering fit
- Choice of heavy quark masses, heavy quark schemes
- Choice of starting scale for QCD evolution, choice of parametrisation…etc, etc



Differences are more obvious in ratio

They are large at small-x and at high-x

One way to see the impact of the uncertainties on the parton distribution functions at the LHC is in terms of parton-parton luminosities, which are the convolution of the purely partonic part of the sub-process cross-section.

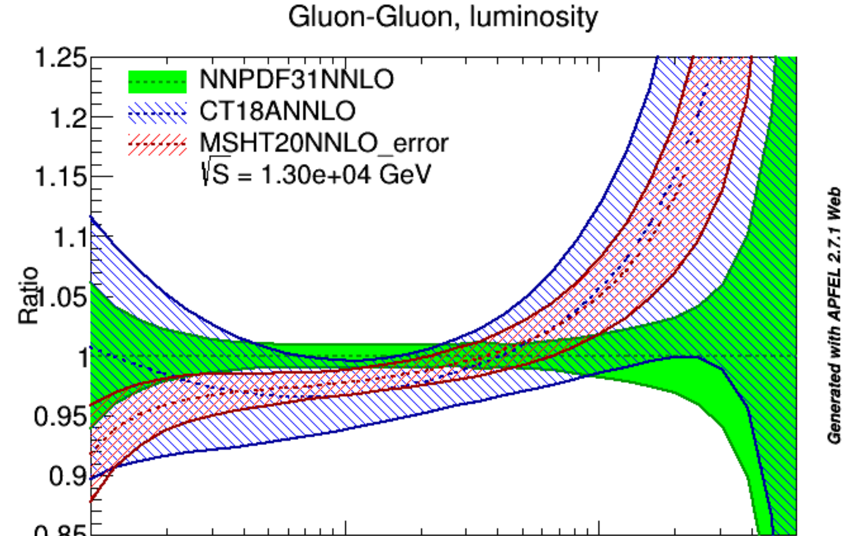


The quark-antiquark and gluon-gluon luminosities for various PDFs are compared here for 13 TeV LHC running in terms of the centre of mass energy of the parton sub- process M_X

$$\frac{\partial \mathcal{L}_{\Sigma_q(q\bar{q})}}{\partial \hat{s}} = \frac{1}{\hat{s}}\int_\tau^1 \frac{dx}{x} \sum_{q=d,u,s,c,b} [f_q(x,\hat{s}) f_{\bar{q}}(\tau/x,\hat{s}) + f_{\bar{q}}(x,\hat{s}) f_q(\tau/x,\hat{s})],$$

Small M_X corresponds to small x and
Large M_X to large x

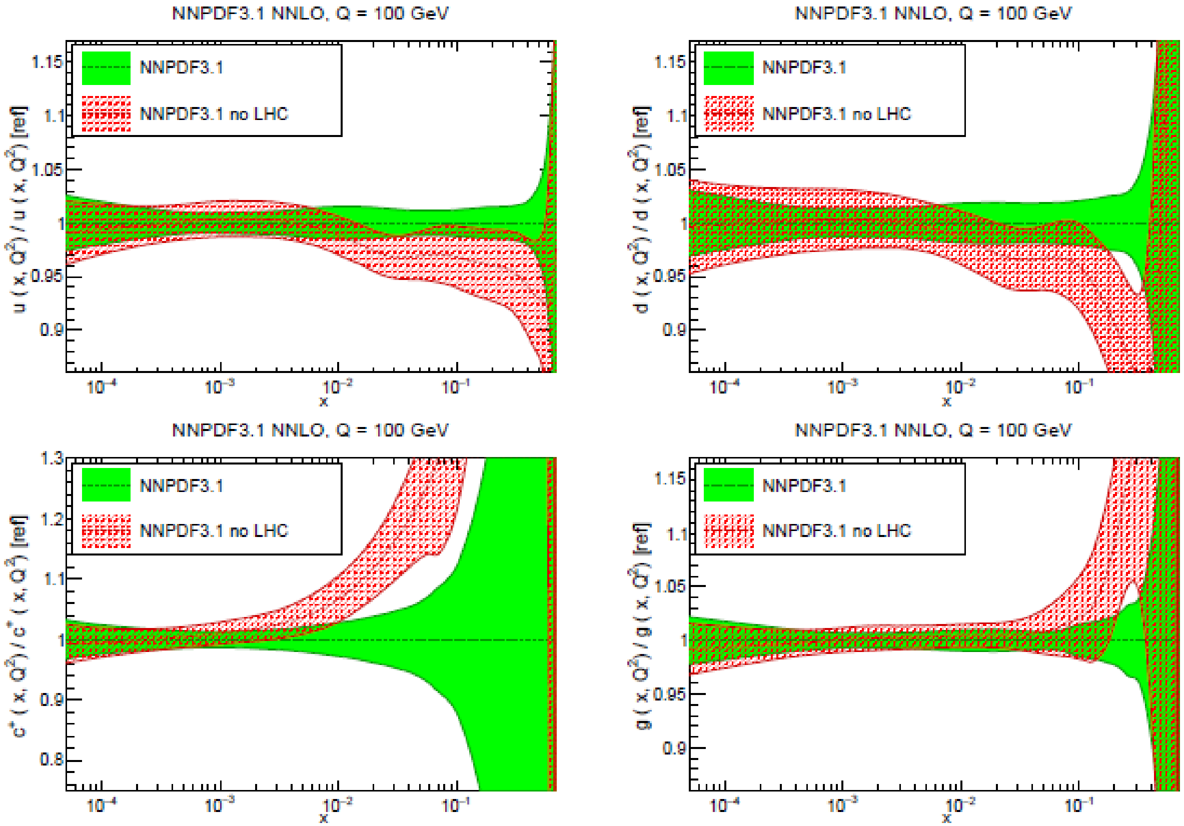
So for quark-antiquark production of W or Z bosons ----at Mx ~80,90 GeV
Or for gluon-gluon production of Higgs at ---Mx~125 GeV
the parton-parton luminosities are fairly well known….but not as well known as we'd like
This is much worse for higher mass particles that could be produced by 'Beyond' Standard Model (BSM) physics



$$\frac{\partial \mathcal{L}_{gg}}{\partial \hat{s}} = \frac{1}{\hat{s}}\int_\tau^1 \frac{dx}{x} f_g(x,\hat{s}) f_g(\tau/x,\hat{s}),$$

Let's see how much LHC data has improved PDFs

NNPDF3.1 includes modern LHC data on W,Z + jets + top + Zpt from 7 and 8 TeV running. Compare PDFs with and without LHC



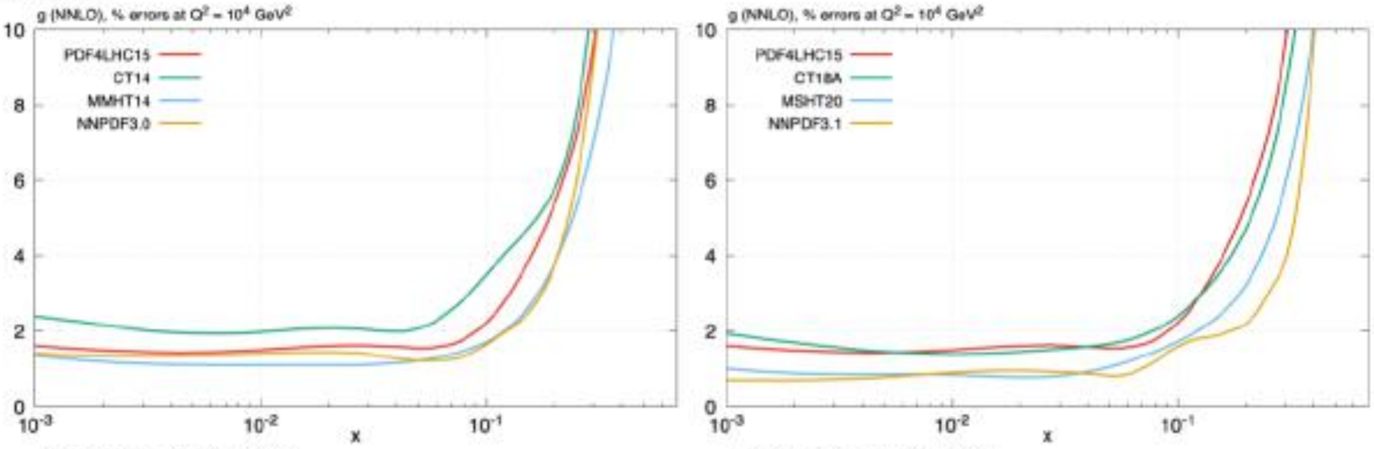
Some of the data input to NNPDF3.1 –like the ATLAS W,Z data have already reached a limit of how accurate they could be. The experimental uncertainties of O(1%) are limited by experimental systematics not by statistics. This will not get better in the foreseeable future e.g. with the High Luminosity LHC

This looks good BUT specific choices were made by NNPDF e.g which top-quark differential distributions are used and of which jet data distributions are used etc., and what are the correlations between systematic uncertainties

Other PDF groups are making slightly different choices—such differences could even increase the total uncertainty due to differences between PDF sets

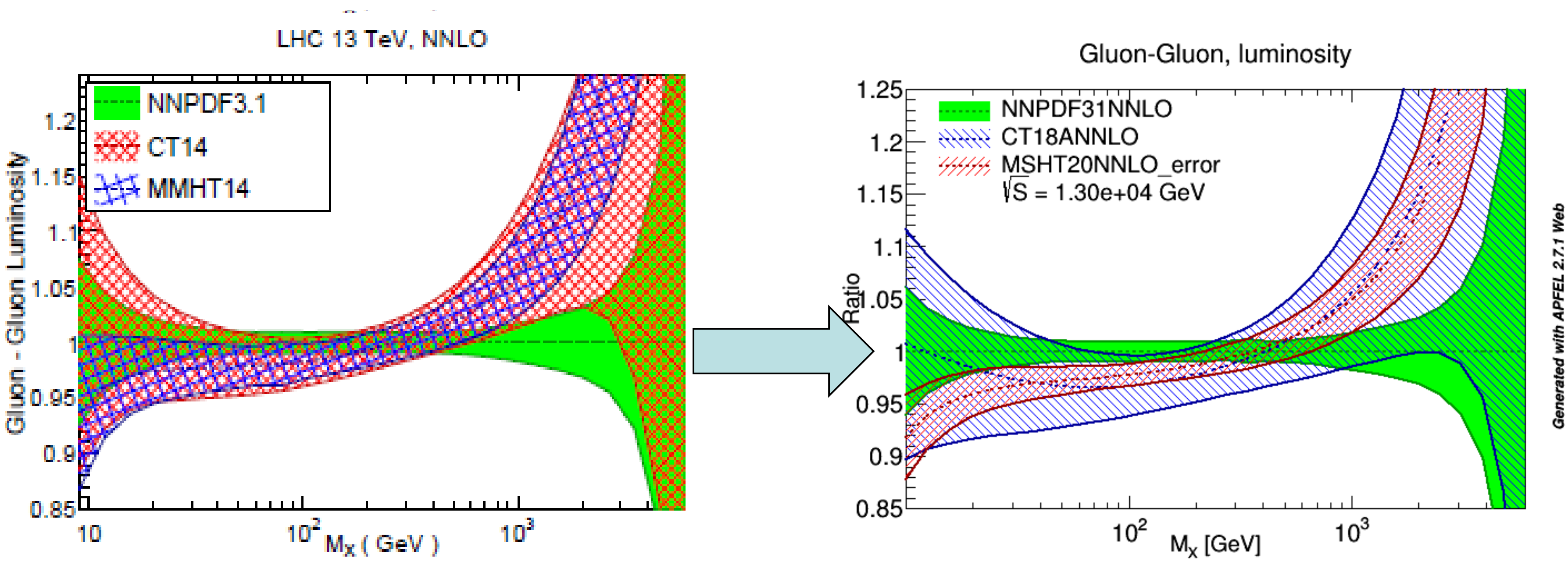
(I will come back to NNPDF4.0)

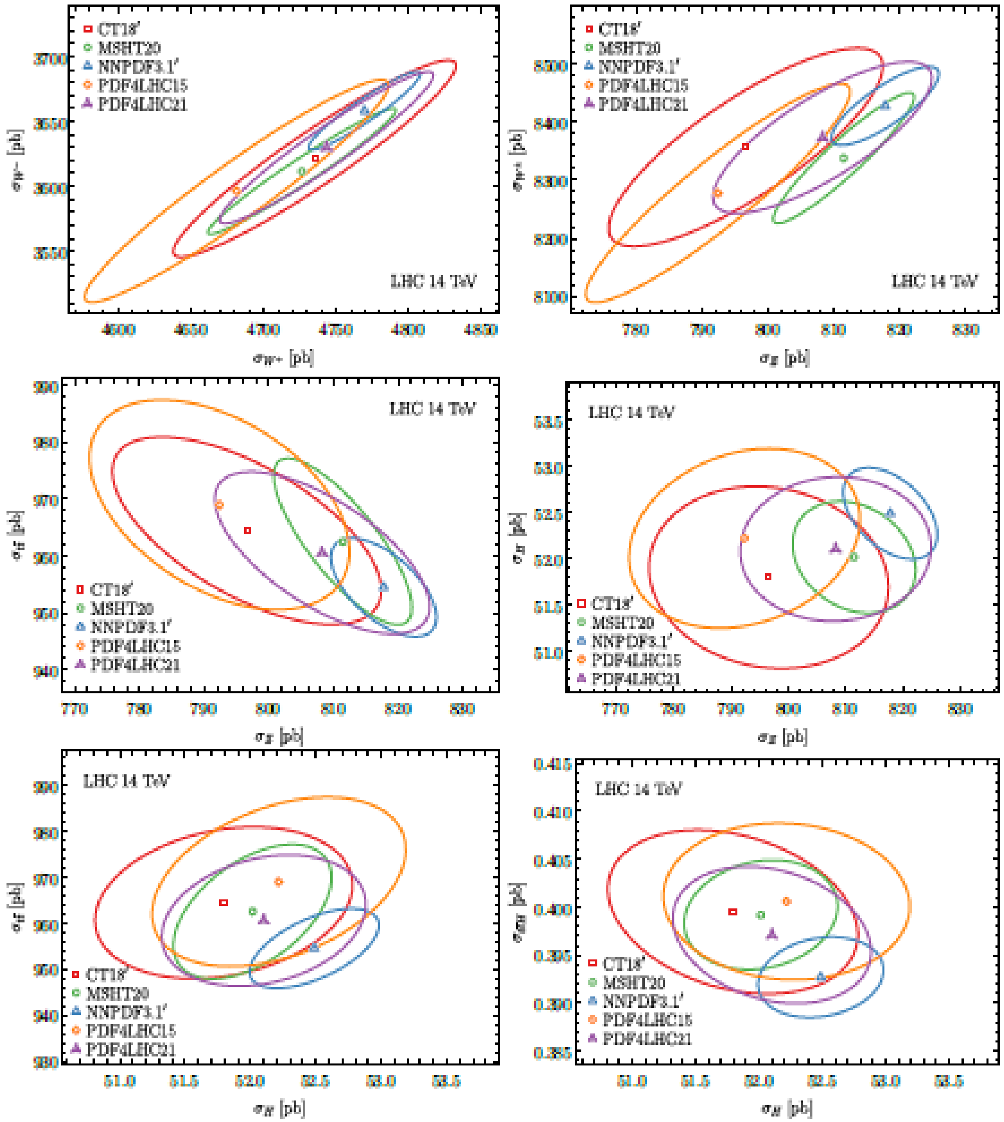
Gluon



IS THERE PROGRESS?

As the uncertainties of each individual PDF decrease with the input of more information, the divergence of the PDFs from each other has increased



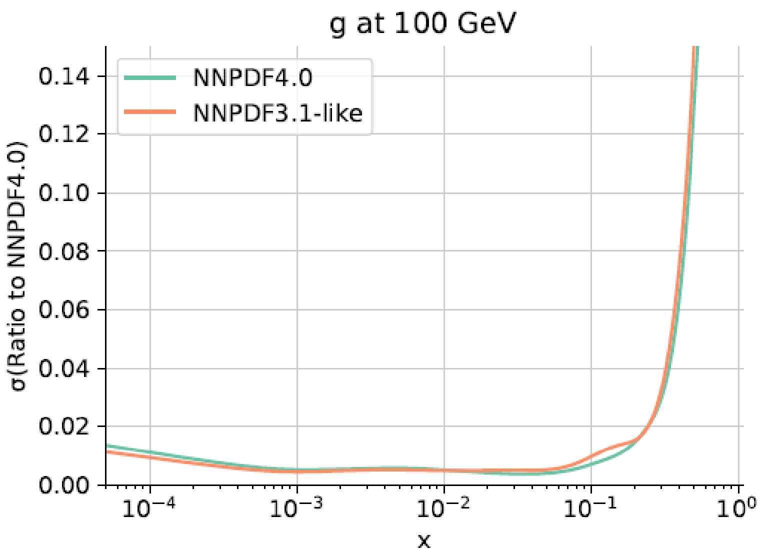


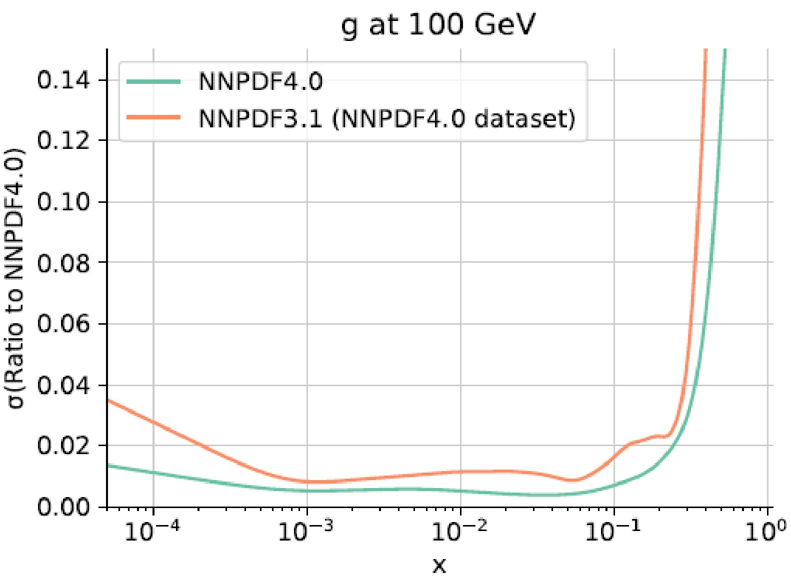
The PDF4LHC group makes combinations of the PDFs from the three main fitting groups NNPDF, CT and MSHT

The PDF4LHC15 combination has now been superseded by the PDF4LHC21 combination (issued in 2022!) arxiv: 2203.05506

There IS an improvement in uncertainty BUT this is not enough e.g.

1. to reduce the PDF uncertainty on on LHC measurements of SM parameters such as m_W, sufficiently to compete with the CDF uncertainty- we need more than this…
2. PDF uncertainty at high x still limits our ability to see new physics





Since the issue of PDF4LHC21 there has been a new PDF set from NNPDF4.0
This has a lot of new data from the LHC

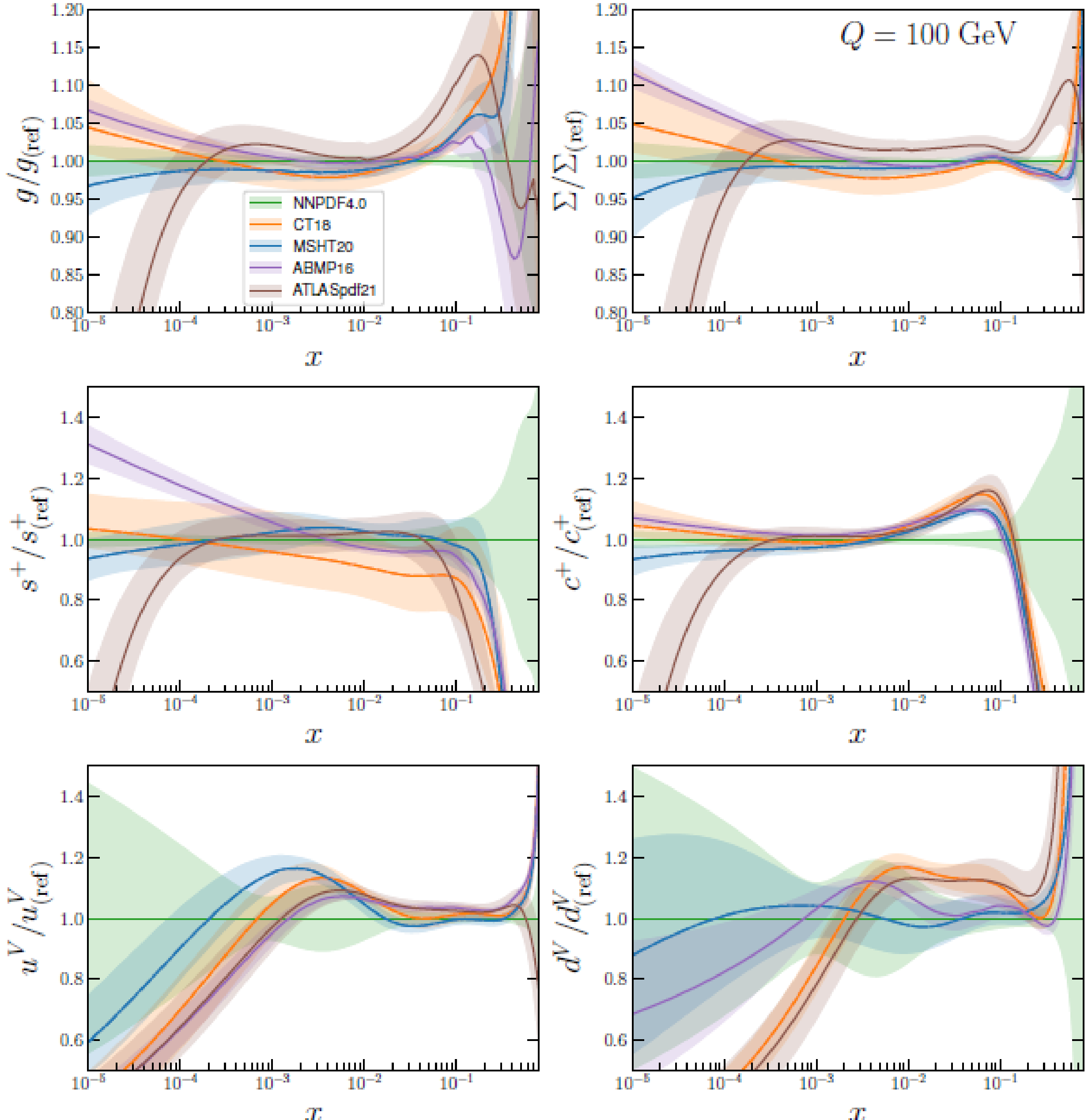
Nevertheless the improvements in uncertainty are not much due to these data, they are more **due to improvements in their procedure**

The top plot compares the uncertainties of NNPDF4.0 and 3.1 data sets using the **SAME new methodology**

The bottom plot shows the impact of the methodology on the **SAME new data** set
4.0 shows new methodology and 3.1 here shows old methodology on new data-set

There is currently a lot of debate in the PDF community over the new methodology.
But even if it is accepted this does not help much if one is trying to combine with other PDFs MSHT20 and CT18 with different central values and larger uncertainties.

A closer look at modern PDFs going down to VERY low-x for Q=100, central LHC probes only down to $x \sim 10^{-3}$



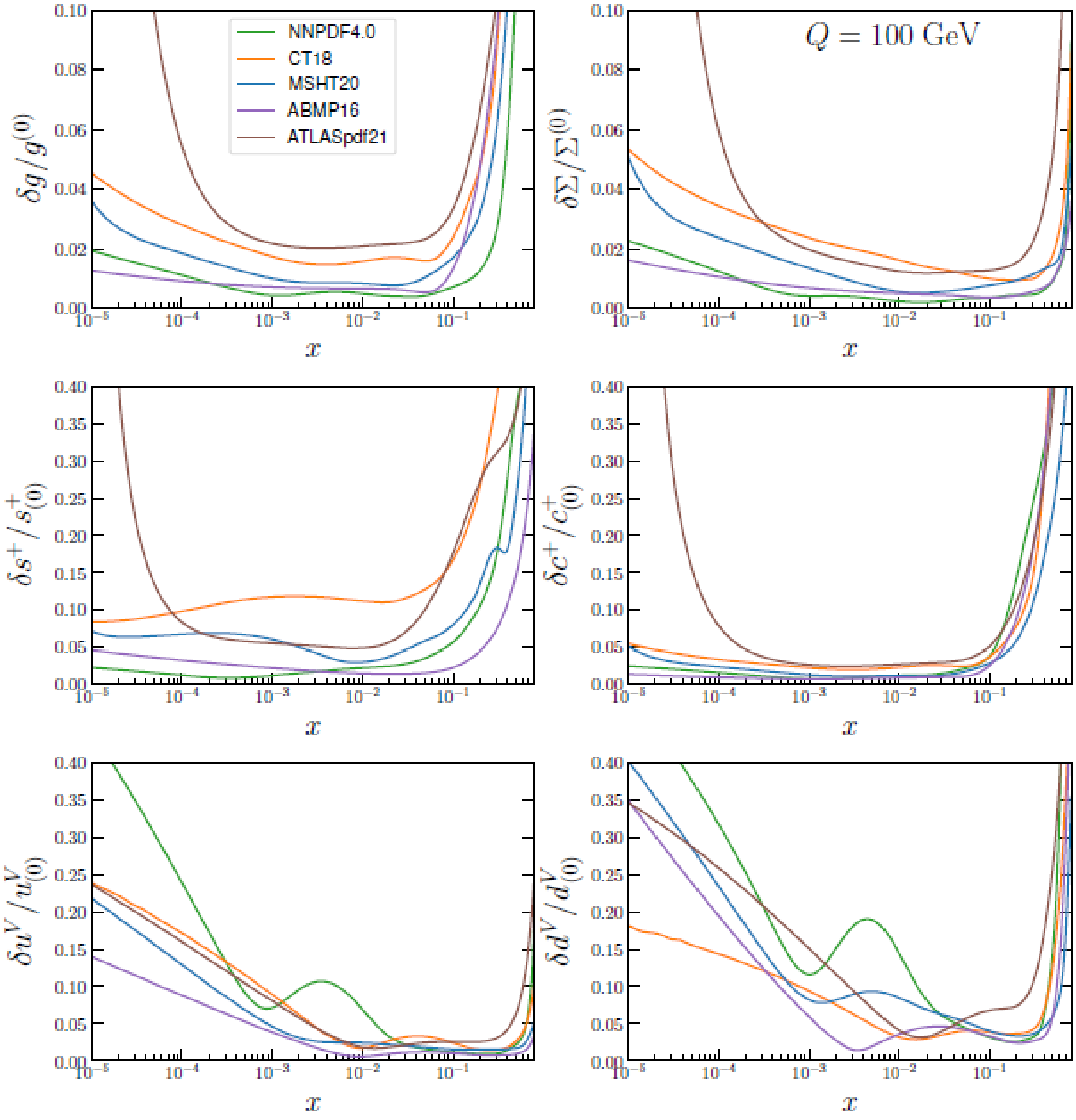
In ratio to NNPDF4.0

We are not so surprised by differences at high-x, though they can be outside uncertainties

e.g.NNPDF has intrinsic charm. But also less strange suppression

Differences in low-x valence are also unsurprising, when little is known on valence at very low-x

Let us look at low-x gluon



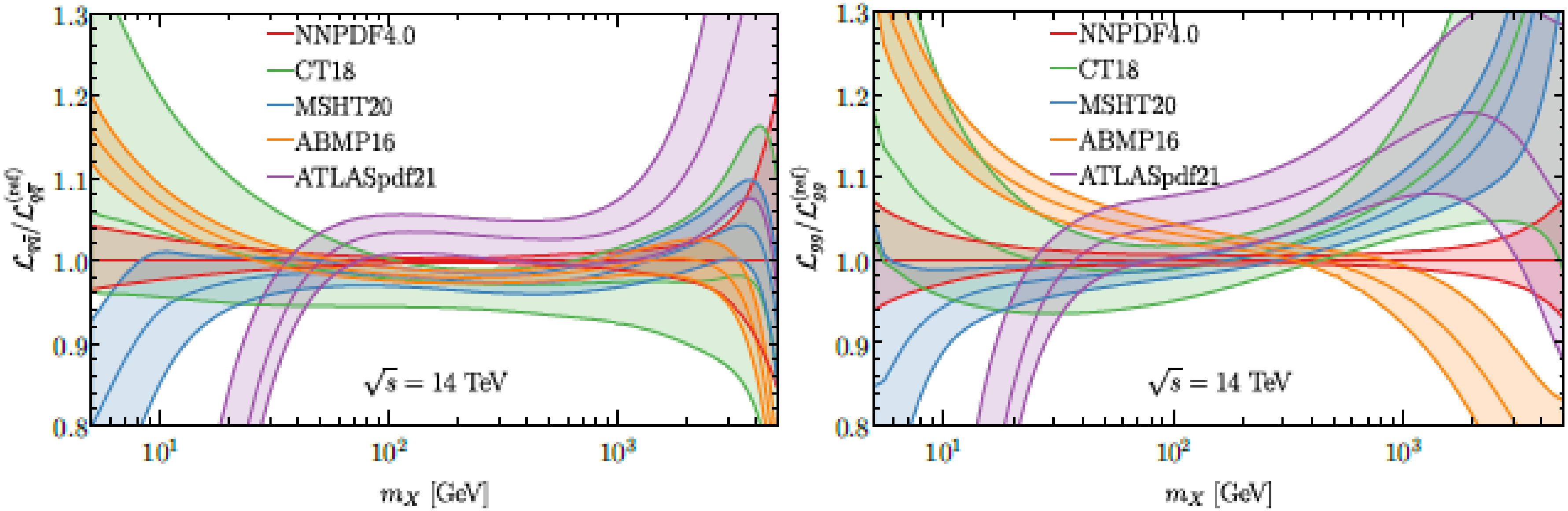
But first look at uncertainties

NOTE ABMP16 is relatively small in regions where similar amounts of data are used, because Δχ2=1 is used rather than a higher tolerance

ATLASpdf21 is larger at low and small x because less data are used

CT18 is often the larger of CT, MSHT because of a larger tolerance than MSHT

NNPDF4.0 has generally very small uncertainties in the data region--- new procedure, positivity, integrability etc..



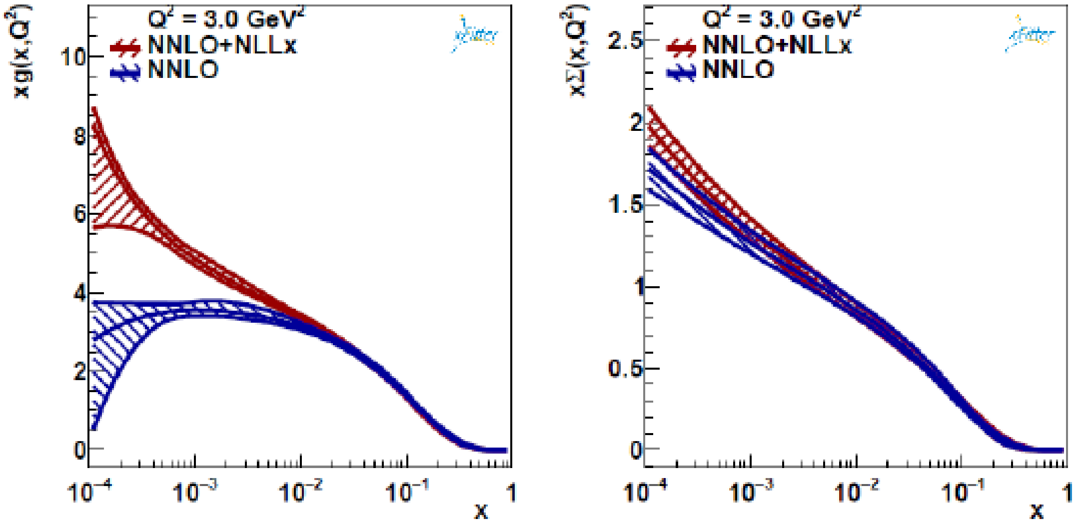
Now let us see the consequences of differences at low-x for LHC luminosities:
q-qbar left, g-g right in ratio to NNPDF
We don't often worry about mX < 30 GeV at the LHC

But if we did.. We need to worry about the low-x theory

i) The HERA data at $Q^2 < 10$ GeV^2 ($x < \sim 10^{-4}$) were cut precisely to avoid this problematic region– (the HERA data are still the main data which probe this region) but it turns out that this is almost exactly the wrong thing to do at NNLO– a better approximation to 'the truth' would be got by fitting down to lower Q^2 and putting up with the larger $\chi2$– as is done by MSHT (who have a similar gluon parametrisation), CT, NNPDF

What do I mean by 'the truth'– well I mean what one might get at higher order or with BFKL ln(1/x) resummation

There has long been an issue that at low-x one should probably be resuming ln(1/x) terms as well as ln(Q^2) terms –this is BFKL resummation and is beyond DGLAP
This has been done by NNPDF- NNPDF3.1sx 1710.05935
And on the HERAPDF using xFitter 1802.00064 (using HELL, Bonvini 1805.08785)



What does it do?
It turns blue into red– dramatic change on the low-x gluon

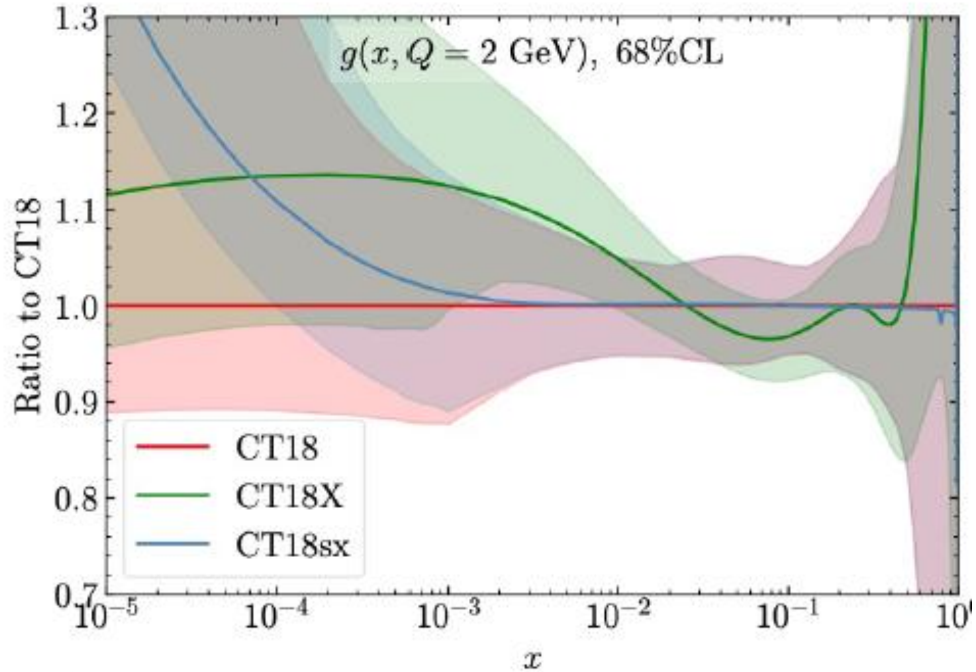
Gets a better χ2 for the low-x,Q2 HERA data by ~70 units

CT have now also done this in CT18sx

But there is another thing one needs to consider– **high density effects** when the gluon gets large such that gluons may recombine, as well as split, and this may lead to gluon saturation. CT have modelled this with an x dependent scale for DIS in CT18X
Not Q^2 BUT

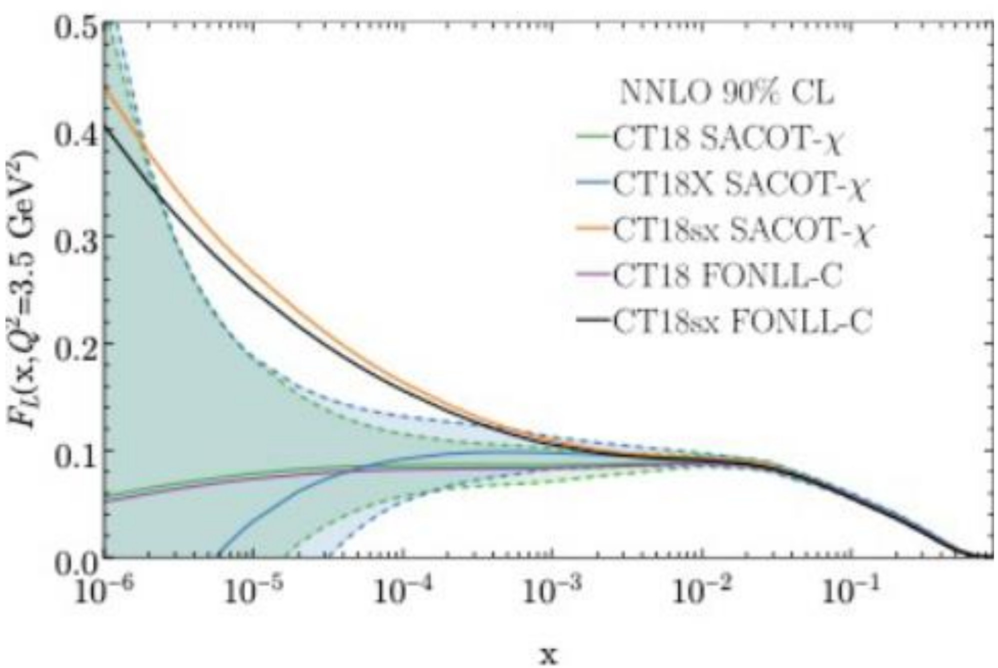
$$\mu^2_{DIS,X} = 0.8^2\left(Q^2 + \frac{0.3\ GeV^2}{x^{0.3}}\right)$$

This also enhances the low-x gluon—--
And it gives a similar decrease in χ2 for the low-x,Q2 HERA data by ~70 units



Compare gluon shapes at low scale

CT18X is a variant which uses a scale intended to mimic saturation
CT18sx is a variant with ln(1/x) resummation

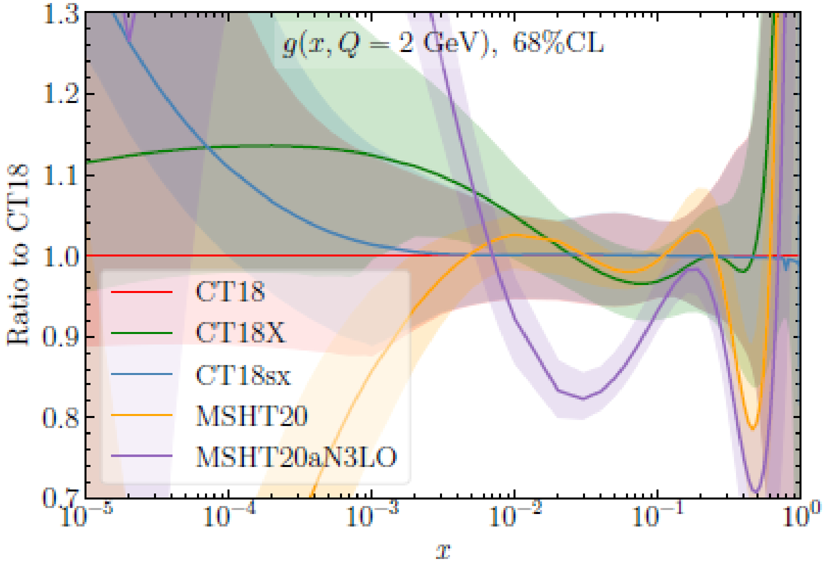


How to tell these apart?
FL measurements at low-x

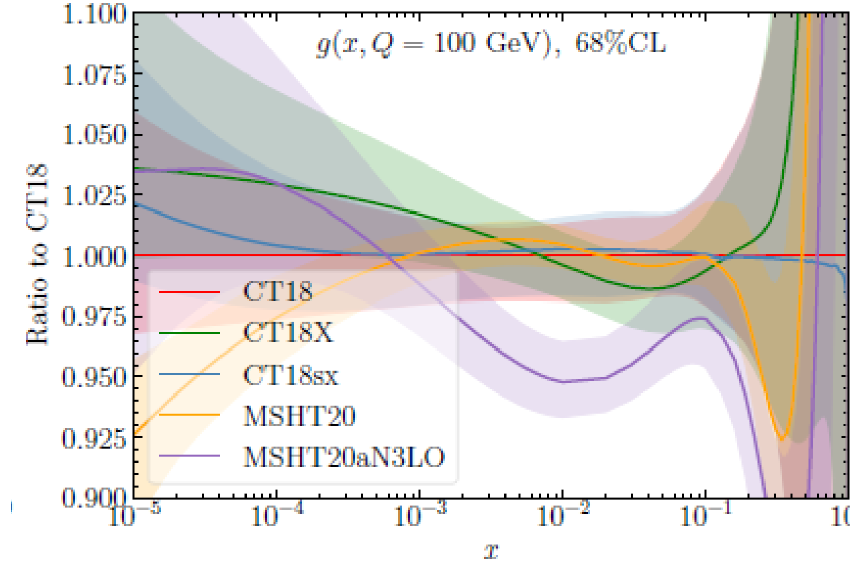
Returning to the LHC there has been a parallel development --- N3LO

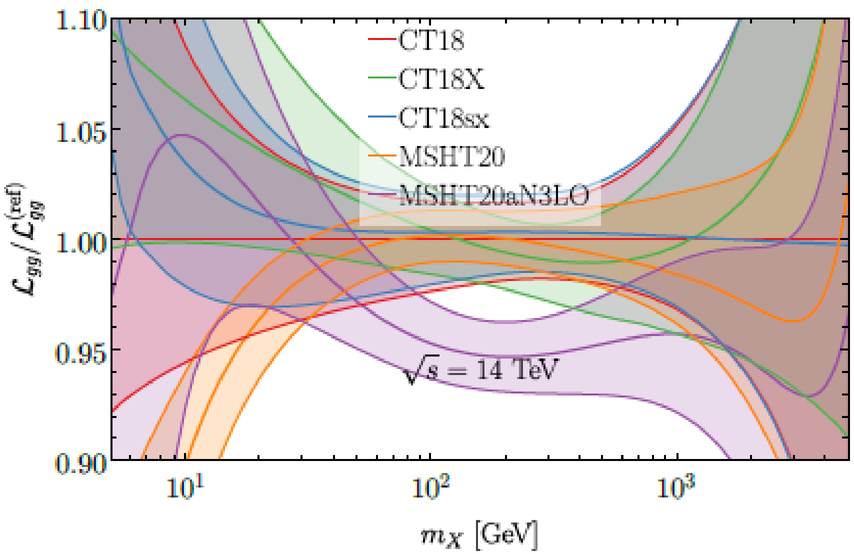
Well at least approximately

This has an astounding effect on the low-x gluon at low scales



Which persists to LHC scales

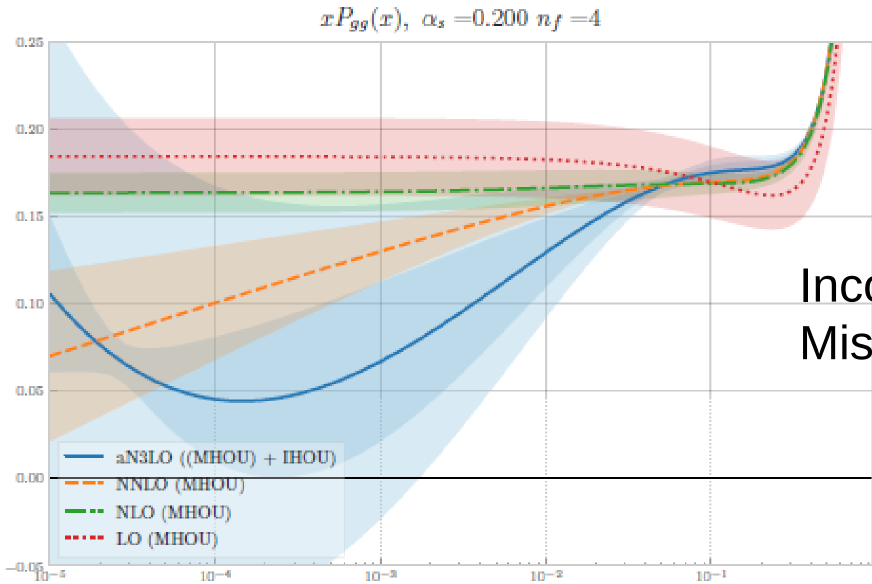




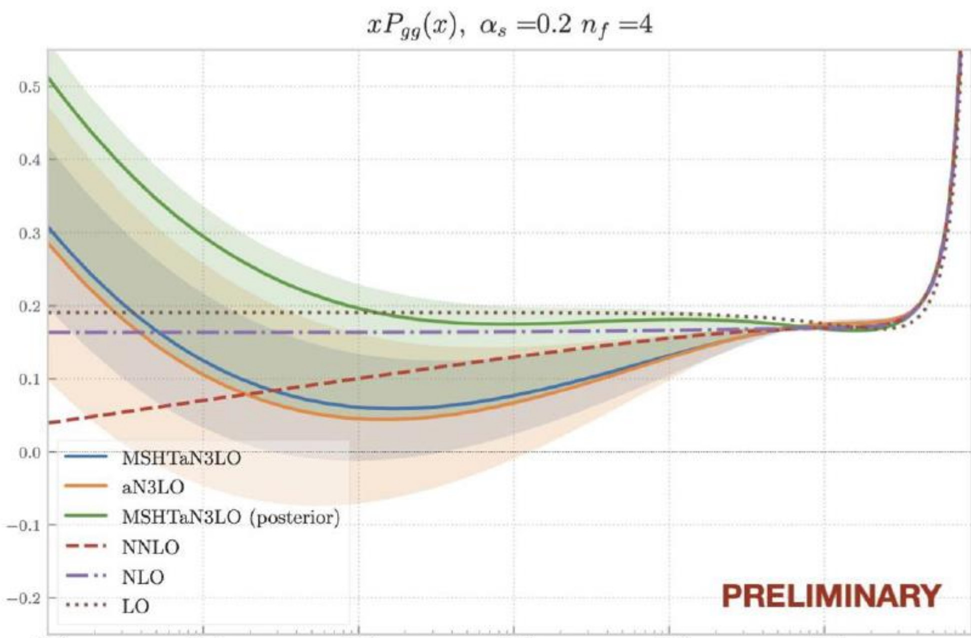
Contrast the MSHT20 NNLO
With the MSHT20aN3LO
But also note it is much stronger than the changes of CT18 to either CT18sx or CT18X– although there are similarities in a rise of the low-x gluon

More alarming is the 'knock-on' effect on the gluon-gluon luminosity -- a decrease of ~5% at the Higgs scale

Now NNPDF have issued some work at N3LO arXIV:2306.15294 and discussion at Les Houches 2023



Incomplete Higher Order Uncertainties (light)
Missing Higher Order Uncertainties (dark)



There is not complete agreement on N3LO between MSHT and NNPDF – may rescue the dramatic effect on gg luminosity somewhat

MSHT (prior) ≃ **NNPDF**

MSHT (posterior) shifts within uncertainty band
(absorbs some low-x logs?)

We need N3LO benchmarking

Summary/ Things to worry about

- PDF improvement is not just a matter of more data
- Consistency of data matters
- Knowledge of common systematic uncertainties matters
- Real data are always more problematic than pseudo-data projections

- Differences in the PDFs are not just about choice of data set—PDF4LHC reduced data sets still have some difference--
- There are irreducible methodological differences between the PDFs
- Sometime this is just a matter of model choices that can be made consistent------- heavy quark masses, $\alpha_s(M_Z)$.
- But sometimes the choices are made for ‘ideological reasons’—paramterisations, NNs, heavy quark treatment/intrinsic charm
- Greatest differences in definitions of how to set uncertainties – choice of $\chi2$ tolerance /NN method

- N3LO /N3LO benchmarking
- Ln(1/x) resummation, recombination/saturation

Clearly we are going to have to consider these both

- But ALSO – we need to worry about the consequences at the Higgs scale

Back-up

The PDF parametrisation

valence quark distributions (xu_v, xd_v) and the light anti-quark distributions ($x\bar{u}$, $x\bar{d}$, $x\bar{s}$).

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

And for the gluon an extra negative term is added, which gives more flexibility at low-x

$$xg(x) = A_g x^{B_g} (1-x)^{C_g} P_g(x) - A'_g x^{B'_g} (1-x)^{C'_g},$$

(consequences of not adding this extra term have been explored and mostly affect quality of fits to HERA data)

The *A* parameters for the valence quarks and the gluon are set by the sum-rules.
The *A,B,C* parameters of the light quark sea are all free parameters
as are the *B,C* parameters of the valence and the *B,C,A',B'* parameters of the gluon.

D,E,F parameters are added until there is no further significant improvement in χ2

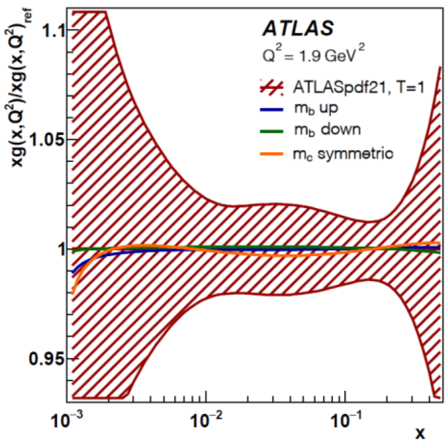
$$P_{u_v}(x) = 1 + D_{u_v} x + E_{u_v} x^2, \quad P_{d_v}(x) = 1 + D_{d_v} x \text{ and } P_g(x) = 1 + D_g x,$$

This results in a 21 parameter fit.
In addition further parameters, F_{uv} *and* D_{dbar}, are considered as part of a parametrisation uncertainty. These parameters are the only further *D,E,F* parameters which give a visible (if small) shape change even though χ2 is not significantly improved.

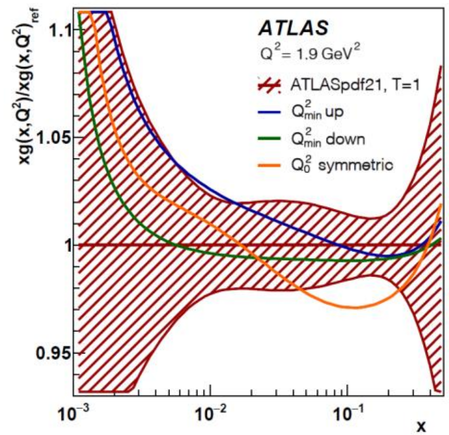
In addition to this a fit using 21 parameters with Chebyshev polynomials was considered, the χ2 of this fit was somewhat worse than our central fit and the PDF shapes are within the uncertainty bands.

Further uncertainties from model assumptions and parametrisation variation

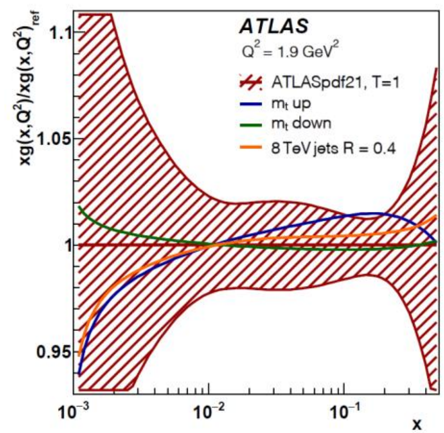
gluon ratio



gluon ratio



gluon ratio



Model uncertainties illustrated in ratio to the central gluon and its experimental uncertainties

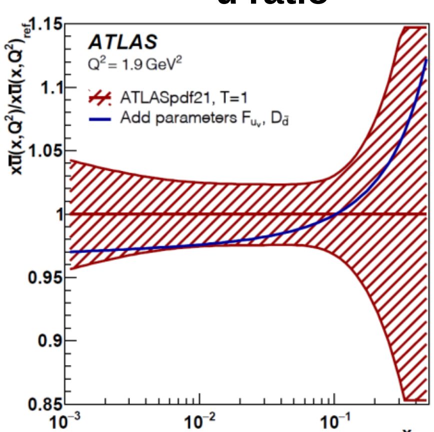
Model uncertainties from the choice of:

- Minimum Q^2 entering the fit
- Heavy quark masses (including top)
- Starting scale for evolution
- Jet radius choice R=0.4 instead of R=0.6 for inclusive jets

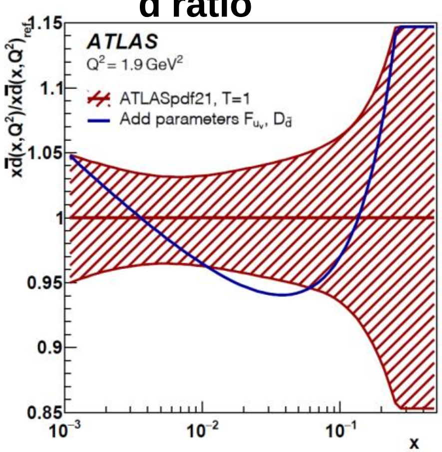
Parametrisation variation including extra F_{uv} and D_{dbar} parameters

Central χ^2/NDF	2010/1620
Model variations	
$Q^2_{\text{min}} = 12.5\ \text{GeV}^2$	1947/1571
$Q^2_{\text{min}} = 7.5\ \text{GeV}^2$	2076/1660
$m_c = 1.45$ GeV (sym)	2025/1620
$Q^2_0 = 1.6\ \text{GeV}^2$ (sym)	2018/1620
$m_b = 4.3$ GeV	2016/1620
$m_b = 4.1$ GeV	2014/1620
$m_t = 175.0$ GeV	2063/1620
$m_t = 172.5$ GeV	2018/1620
$R = 0.4$	2080/1620
Parameter variations	
$F_{u_v}, D_{\bar{d}}$	2007/1620

$\bar{u}$ ratio



$\bar{d}$ ratio

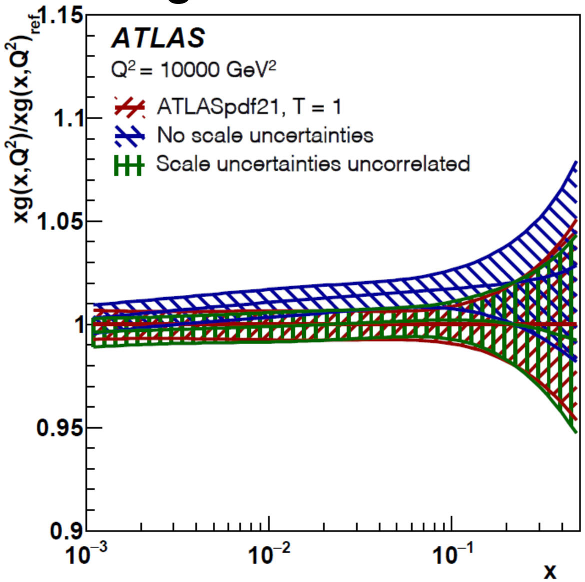


Parametrisation uncertainties illustrated in ratio to the central d-valence and dbar and their experimental uncertainties

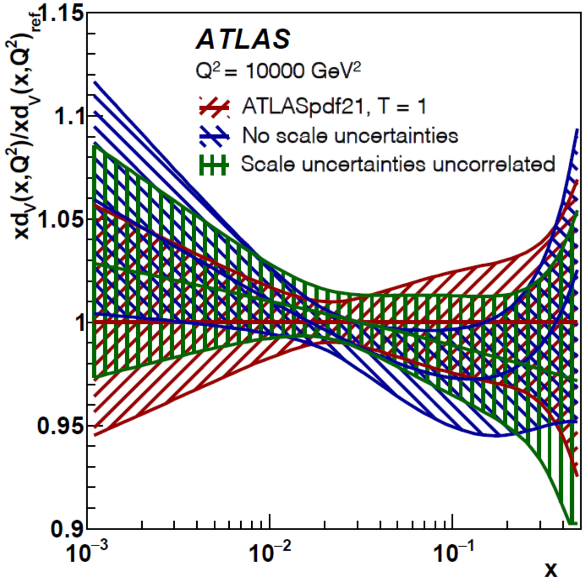
Impact of scale uncertainties

For the inclusive W,Z production at 7 and 8 TeV the experimental uncertainties are comparable to the scale uncertainties and thus the scale uncertainties are included as theoretical uncertainties in the fit. By default they are correlated between the W and Z data and between the 7 and 8 TeV data in the ATLASpdf21 fit.

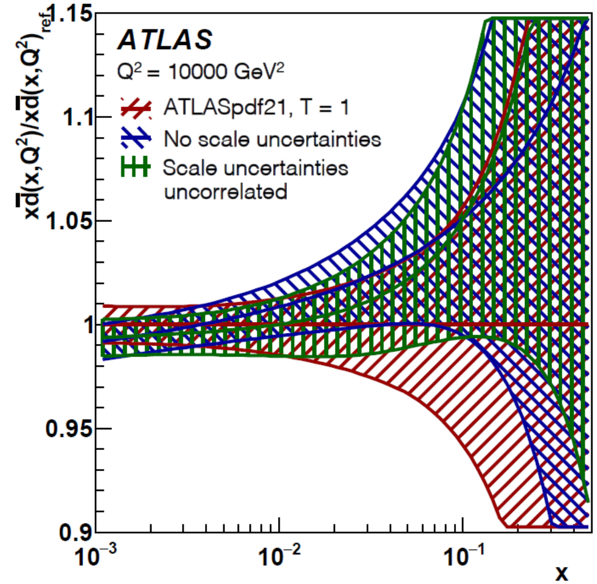
gluon ratio



$d_{valence}$ ratio



$\bar{d}$ ratio



Here we show the ratios of the gluon, d-valence and dbar PDFs with (red) and without (blue) these scale uncertainties included

In green we show the effect of including scale uncertainties but not correlating them between 7 and 8 TeV data

Clearly scale uncertainties can be important if 1% precision is sought

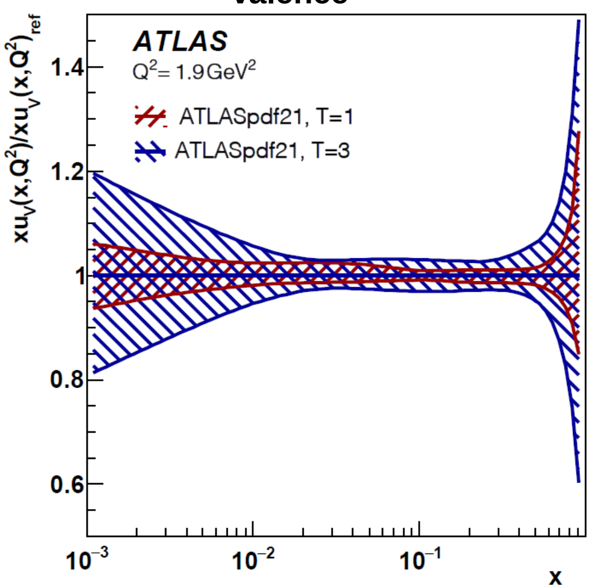
χ2 tolerance,

So far we have applied the conventional $\Delta\chi 2 = T^2 = 1$ for 68%C.L.

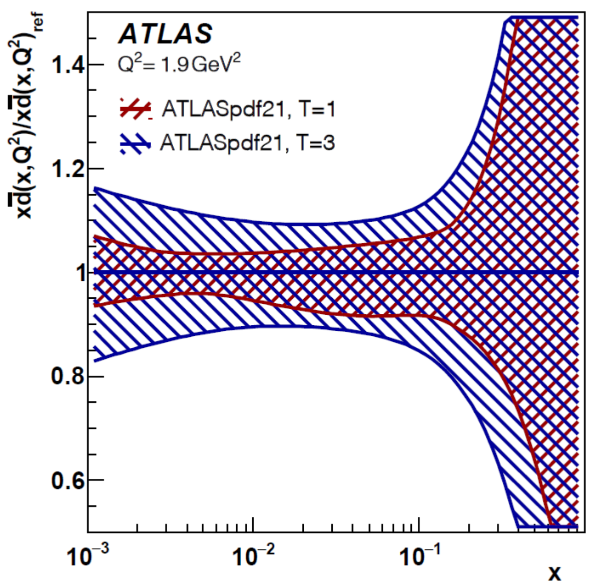
When diverse data sets are included global fitters consider $T > 1$.Historically values of T~7- 10 have been considered by MSHT and CT.

We have followed the MSHT dynamic tolerance procedure (first used for MSTW2008) and obtained an appropriate tolerance for T=3 for the ATLAS data sets included in the fit such that all data sets are fitted within their 68%CL for variations of all eigenvectors.

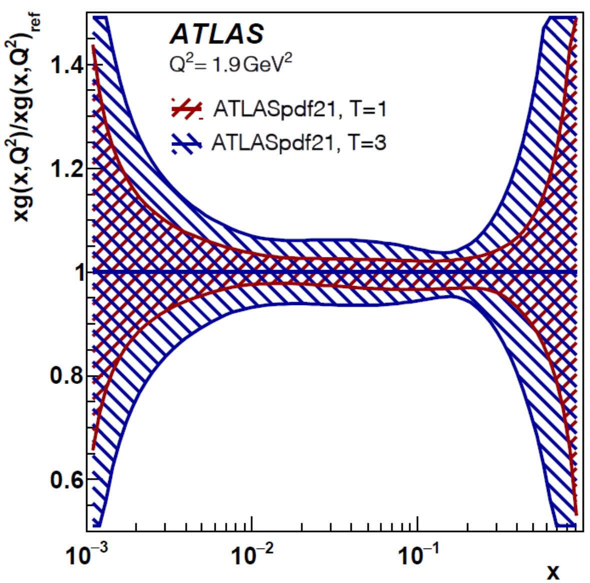
$u_{valence}$ ratio



$\bar{d}$ ratio



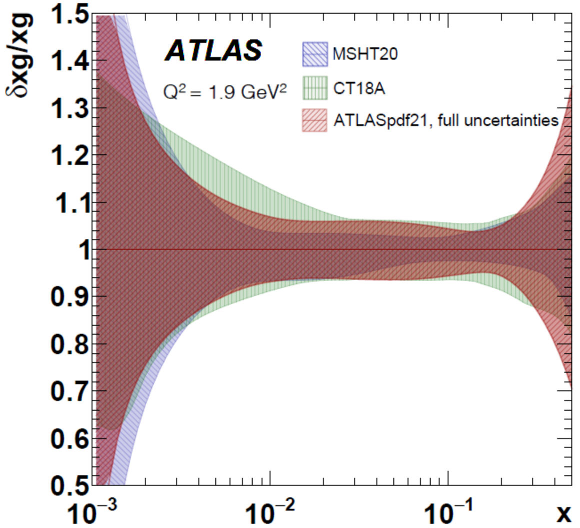
gluon ratio



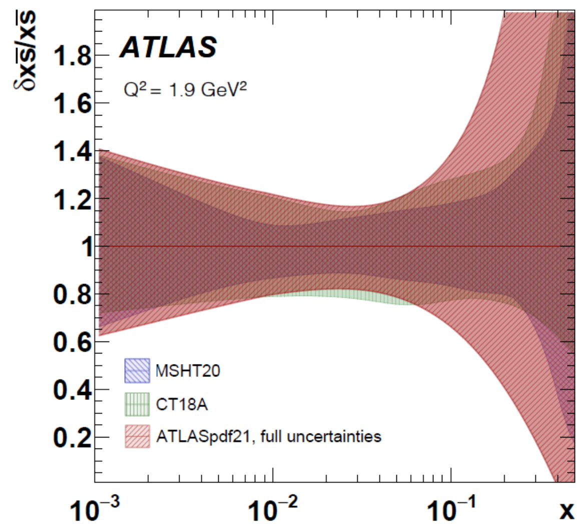
Change from T=1 to T=3 for various PDFs for Full uncertainties including model and parametrisation

Now compare ATLASpdf21 uncertainties to those of modern global PDFs

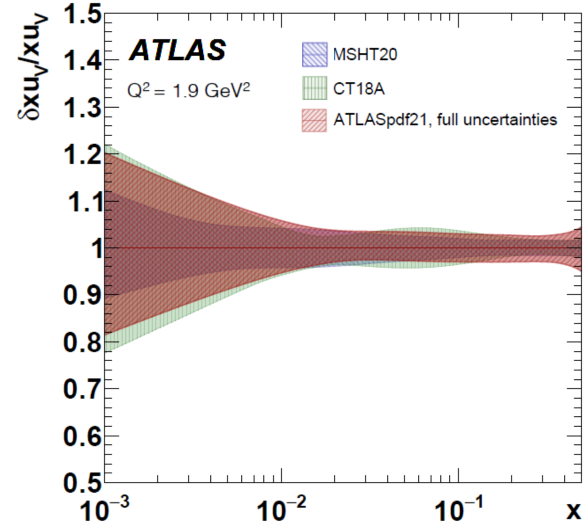
gluon ratio



$\bar{s}$ ratio



$u_{valence}$ ratio

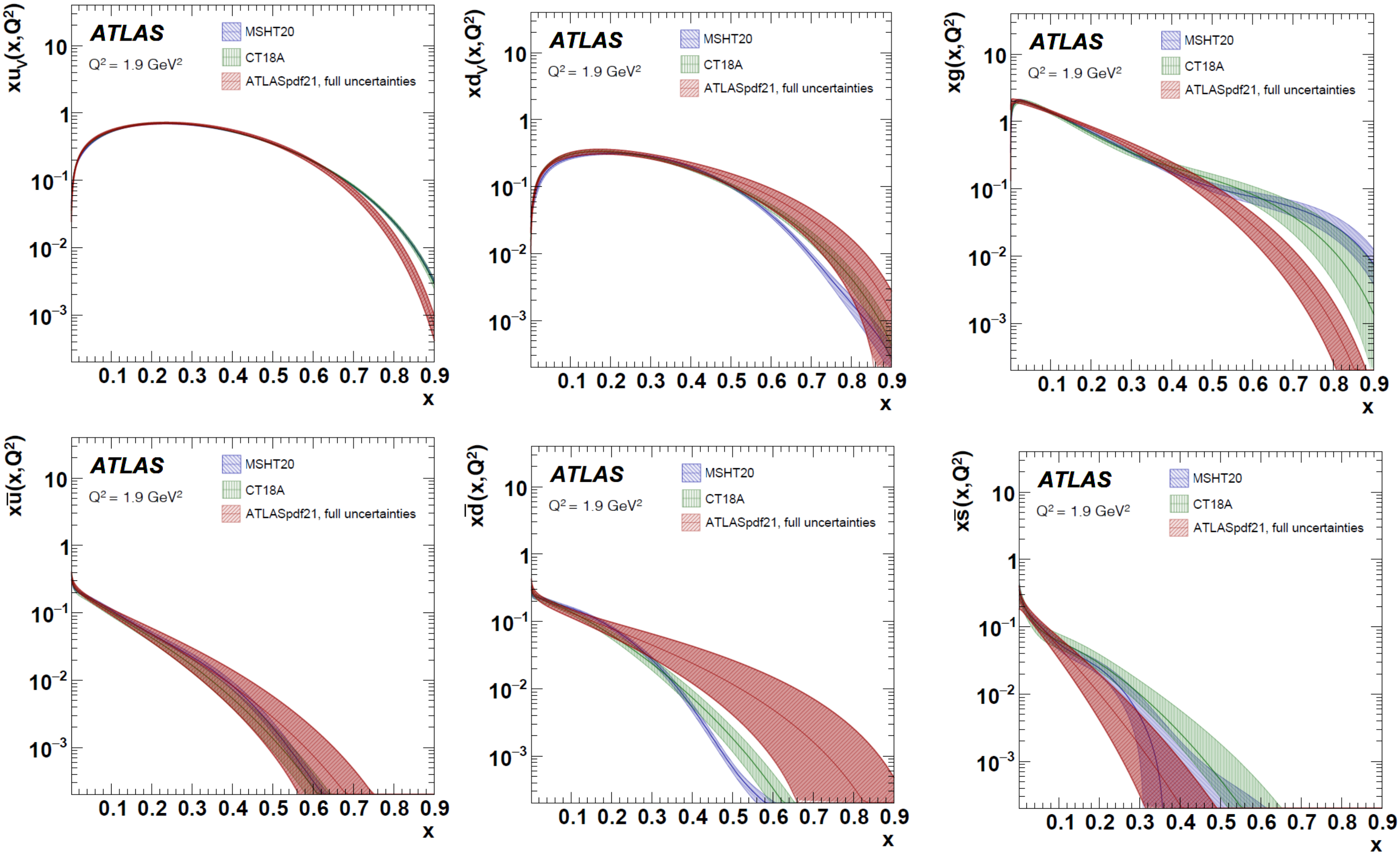


Uncertainties are competitive for x < 0.1 and somewhat larger at high-x depending on the PDF

Note that without enhanced tolerance we would be claiming to do better than global fits for x> 0.1 where we have much less constraining data.

Global fits use older DIS fixed target data and Drell-Yan data as well as Tevatron data, which constrain high-x. However, they are subject to unknown systematic uncertainties and/or unknown correlations of systematic uncertainties on these older data.

ATLASpdf21 is able to describe the most significant of these older data sets well--we no longer need them to set the central values of the PDFs at high-x---
the price we pay for not using them is the larger high-x uncertainties, this should improve with future ATLAS data.

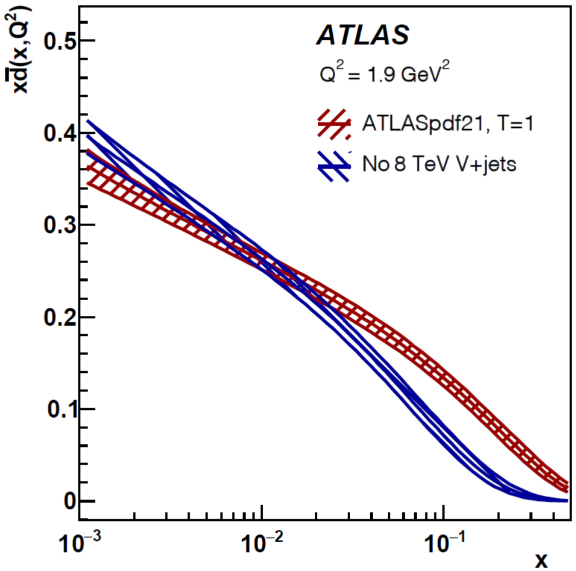


High- x comparison with other PDFs

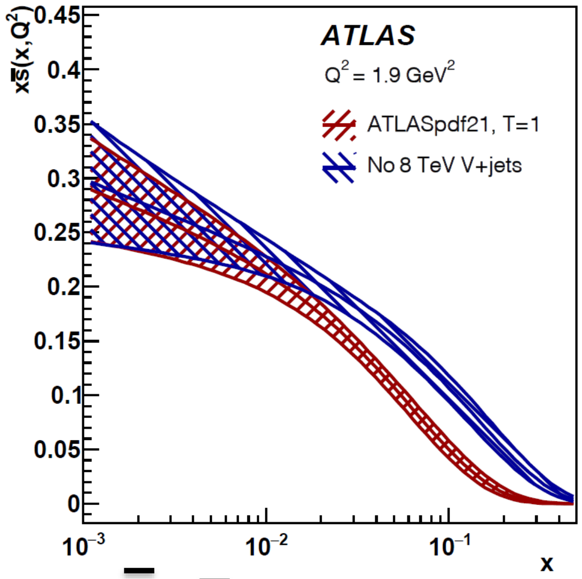
Impact of V + jets data

Is to increase dbar and decrease sbar at high-x

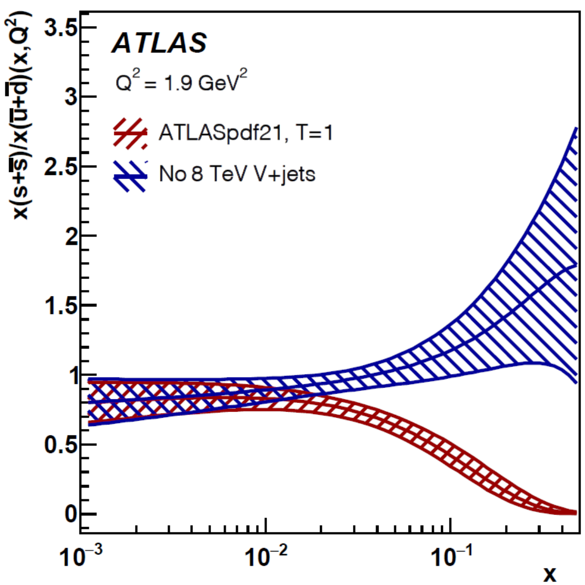
$\bar{d}$



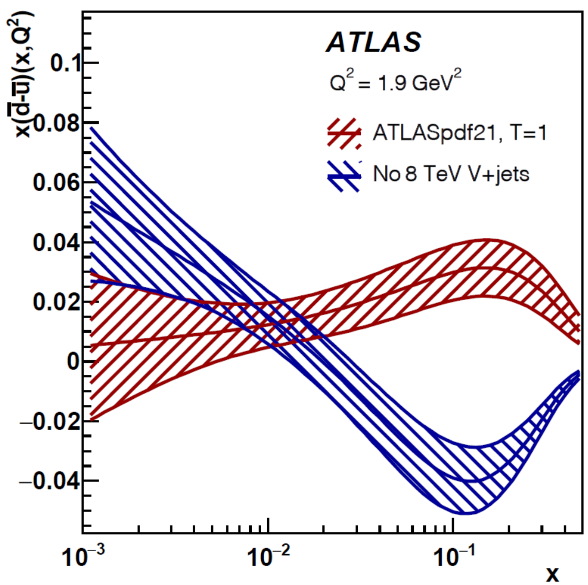
$\bar{s}$



strange/light quarks



$\bar{d} - \bar{u}$



This change looks so dramatic because the V+jets data resolves a double minimum in the rest of the data which are almost equally happy with the blue or red PDFs.

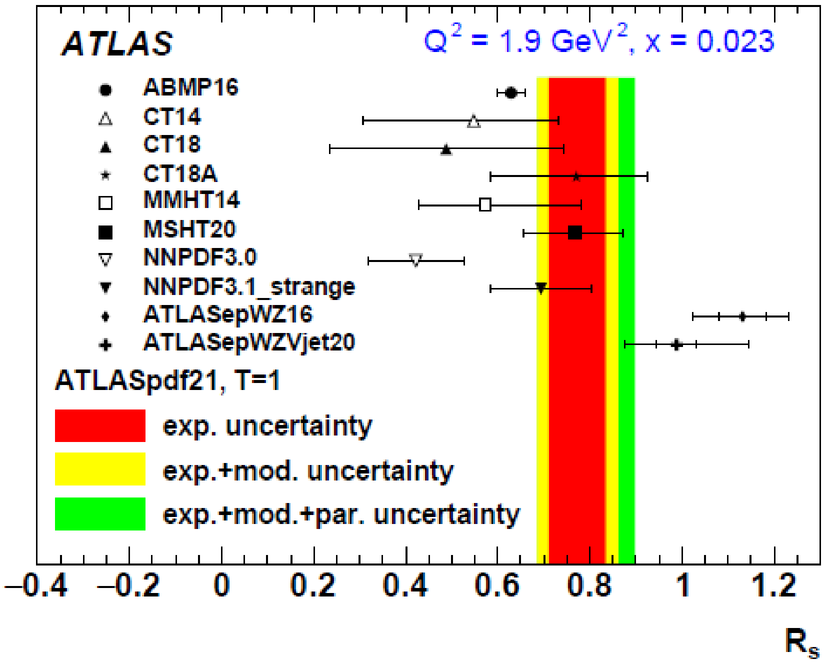
Remember that only experimental uncertainties are shown here. The total uncertainties including model and parametrisation choices are much larger for the blue PDFs, because some of them go to the alternative minimum **See back-up.**

Strangeness

The information on strangeness has often been presented at a single x,Q^2 point and compared to the result of global PDFs

Note that older PDFs CT14, MMHT14,NNPDF3.0 all had Rs~0.5 at low scale (Q^2=1.9GeV^2) **BUT** this has moved up to ~0.8 for CT18A, MSHT20, and NNPDF3.1_strange after ATLAS W,Z 7 TeV data was included (not for CT18 which does not include these data)

ATLAS older fits had Rs~1.0 and have moved down to Rs~0.8 due to input of new data, V+jets and W,Z 8 TeV and greater flexibility of low-x parametrisation



The history of these changes is:

ATLASepWZ16 to ATLASepWZVjets20
Input of V+jets data suppresses Rs at high-x and this has a knock-on effect at x=0.023 so that Rs~ 1.15 →1.0

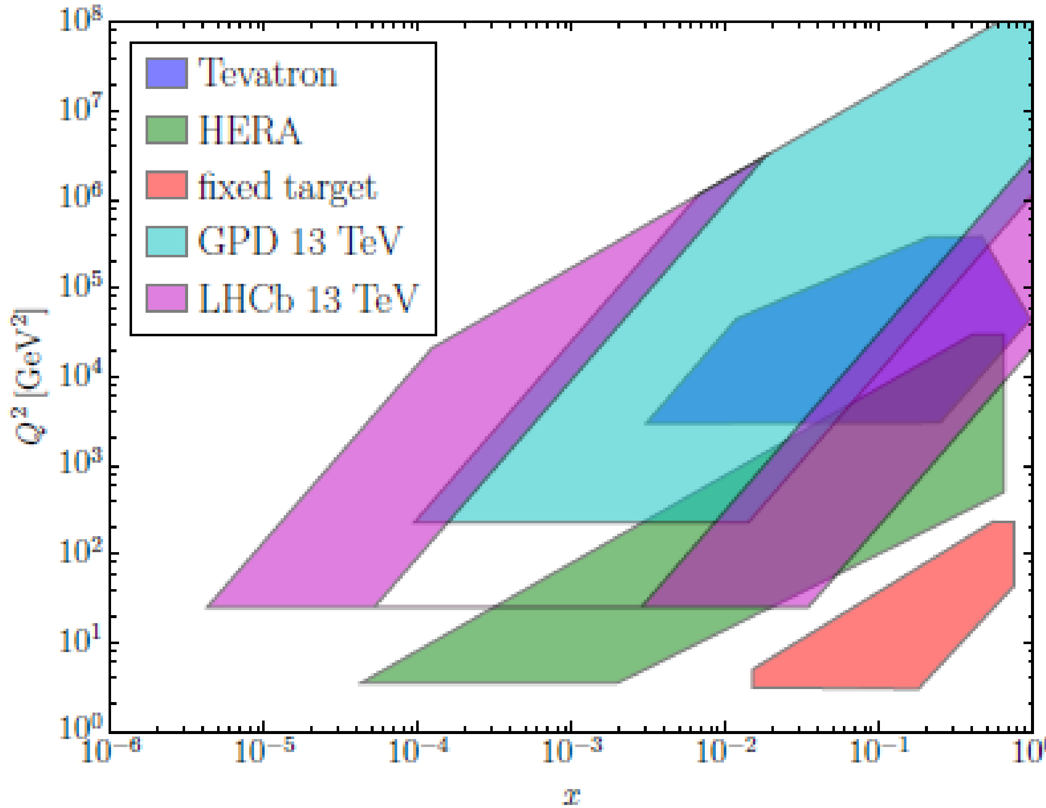
ATLASepWZVJets20 to ATLASpdf21

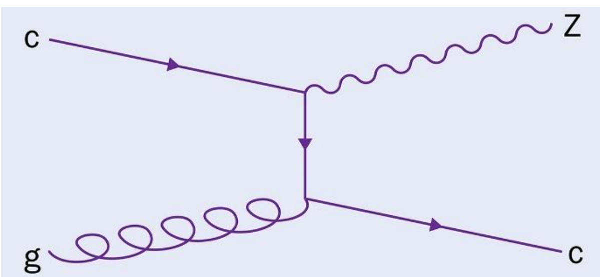
- More flexible low-x parametrisation corresponds to lower edge of Vjets20 error-band Rs~ 1.0 →0.85
- Addition of W,Z 8 TeV data Rs~ 0.85 →0.8

Flavour 2. Charm

The kinematic reach of LHCb goes to both higher and lower rapidity and hence to higher and lower x than CMS or ATLAS (labelled as GPD general purpose detectors)

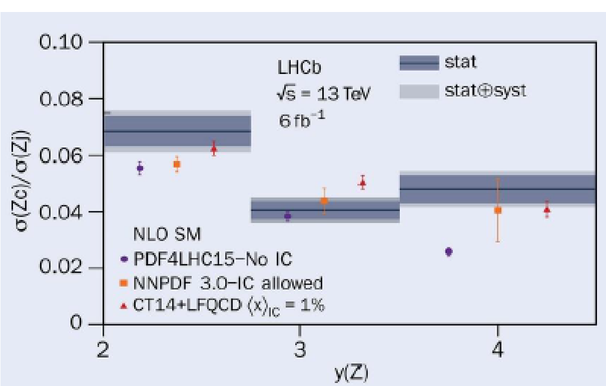
Hence they may be able to look into intrinsic charm in the nucleon

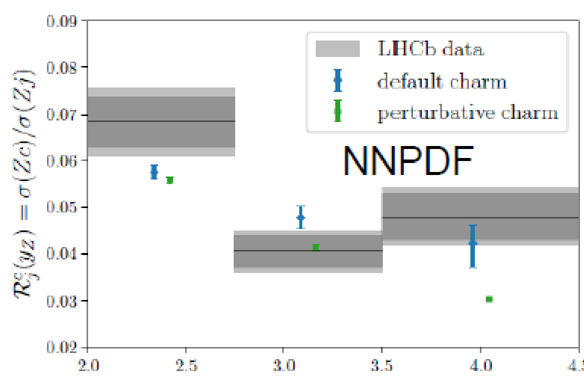




Clearly you need Z+c or γ+c

It's probably smart to take ratios like Z+c/Z+jets and do it at high rapidity to reach high x





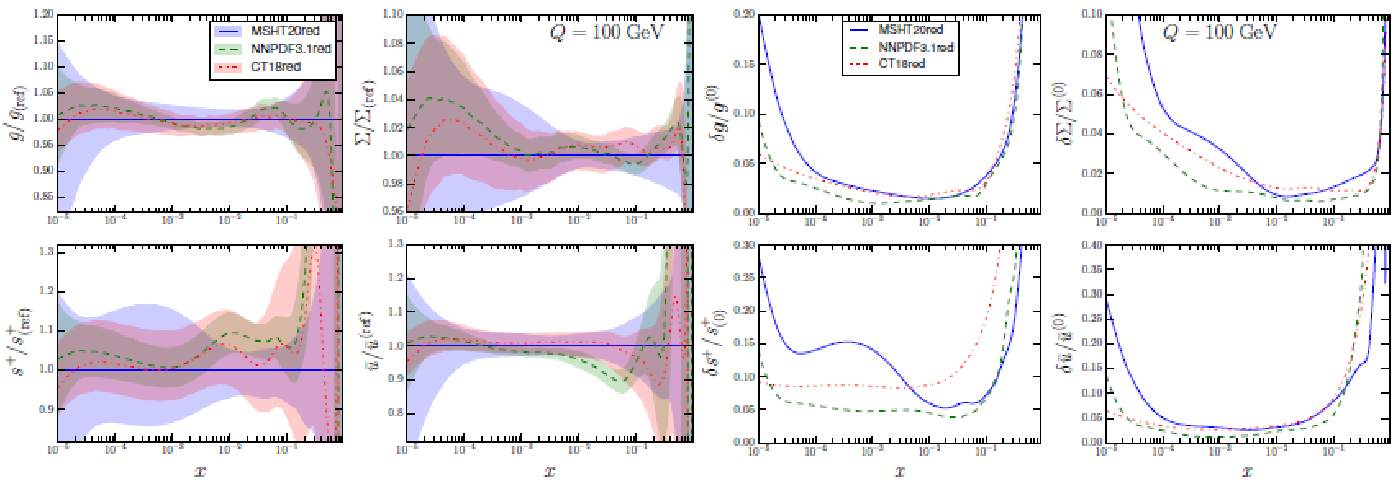
14

The PDF4LHC group makes combinations of the PDFs from the three main fitting groups NNPDF, CT and MSHT

First try to understand differences by using a common data set and common settings for heavy quark masses and alphas

PDF Benchmarking: Reduced Fits

- Use fits to reduced common datasets and common theory settings.



- Very good agreement within uncertainties, including gluon.
- Similar size uncertainties in data regions, differences outside this, reflecting remaining methodological and other choices.
- Agreement much improved relative to global PDFs.
- Same data and theory settings → consistent PDFs. Smaller remaining differences, e.g. in errors, reflect methodological choices.

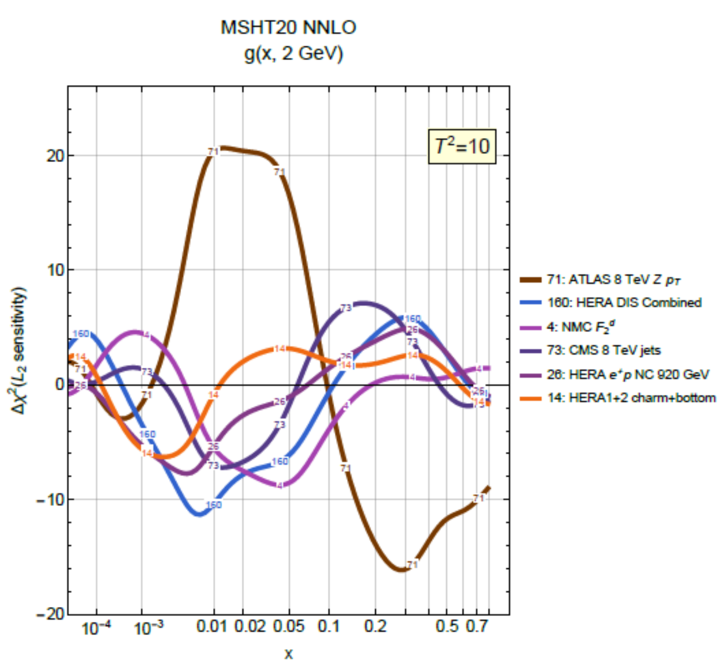
It is not recommended to use these reduced fits, greater consistency does not mean greater accuracy—the differences in the main fits are there for a reason!

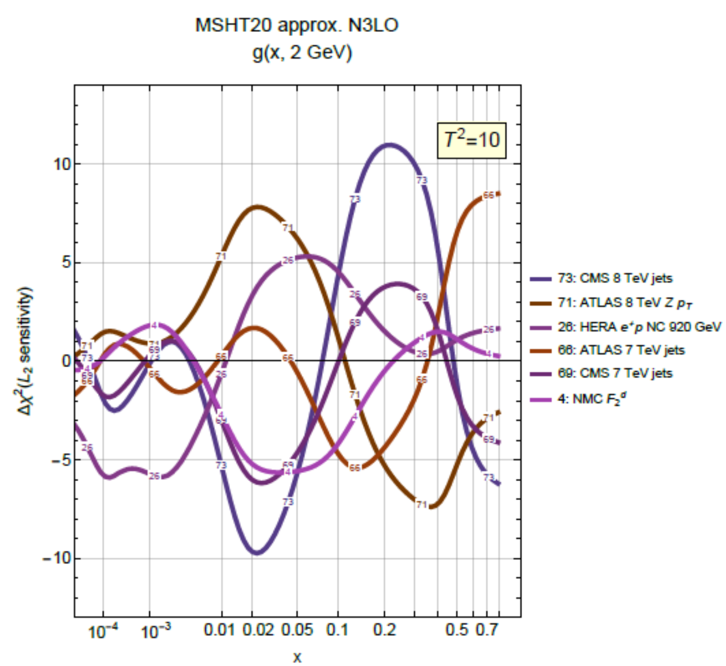
New tools to asses data sensitivity/inconsistency arXIV:2306.03918
L2 sensitivity combines information on the sensitivity of a measurement to a PDF in principle, and the precision of a particular measurement

$$S^{\mathrm{H}}_{f,L2}(E) \equiv \frac{\vec{\nabla}\chi^2_E \cdot \vec{\nabla} f}{\delta_{\mathrm{H}} f} = \left(\delta_{\mathrm{H}}\chi^2_E\right)\ C_{\mathrm{H}}(f,\chi^2_E)\ , \qquad (5)$$

where $C_{\mathrm{H}}(f,\chi^2_E)$ represents the cosine of the correlation angle between f and the χ^2 for experiment E, evaluated over the $2D$ Hessian eigenvector sets.

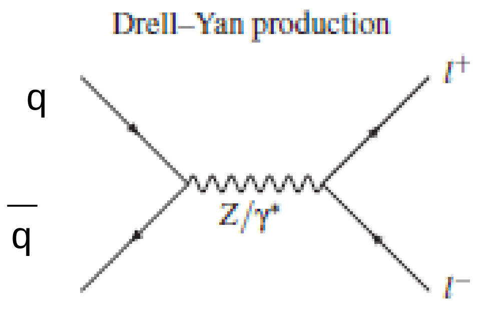
Sensitivity of the gluon to different data sets in the MSHT fit
Zpt VERY sensitive at NNLO……………..less so at N3LO—we will come back to N3LO

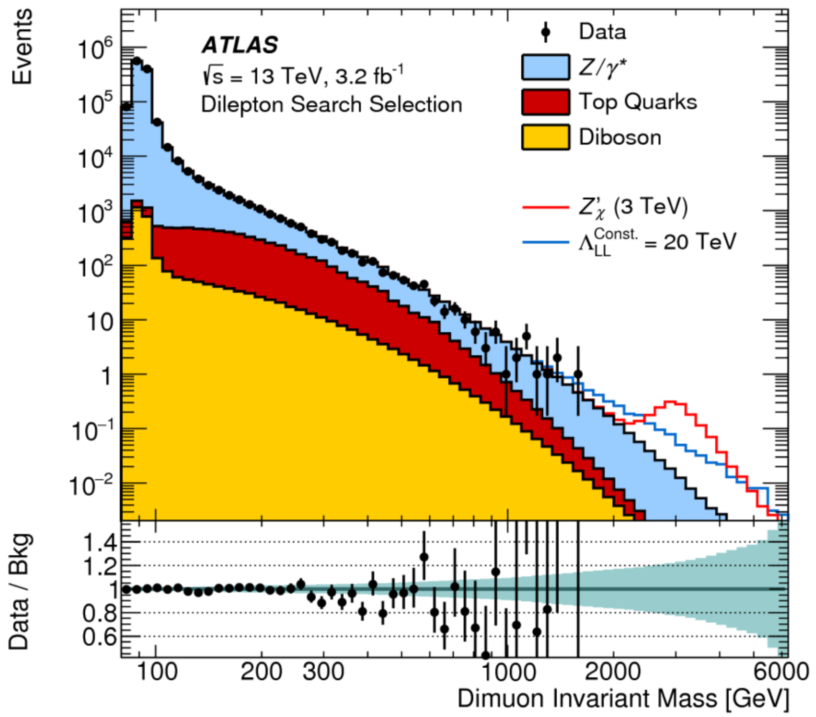




Uncertainty in the high-x sea?-one example

Current BSM searches in High Mass Drell-Yan are limited by high-x antiquark uncertainties as well as by high-x valence quark uncertainties

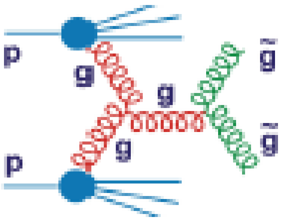


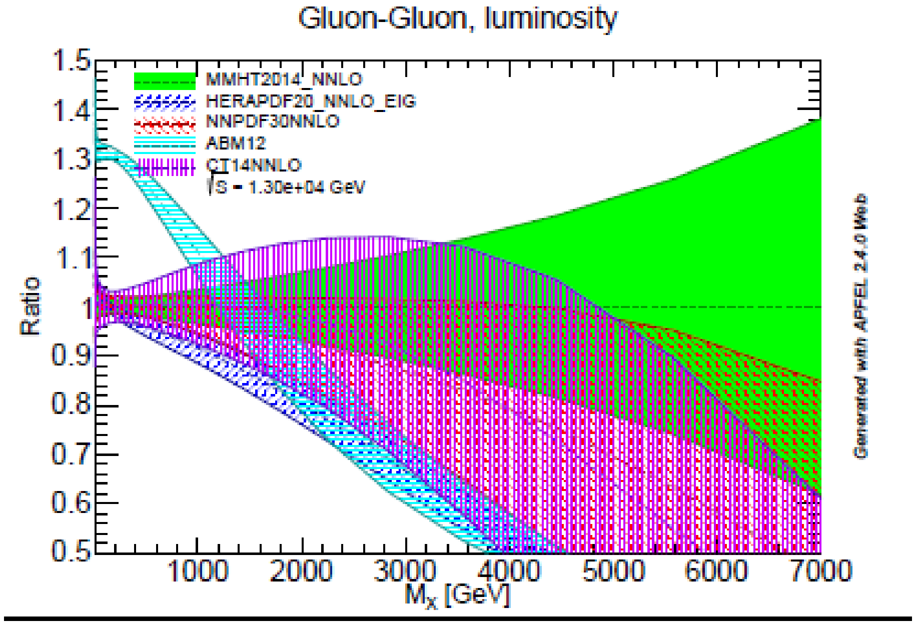


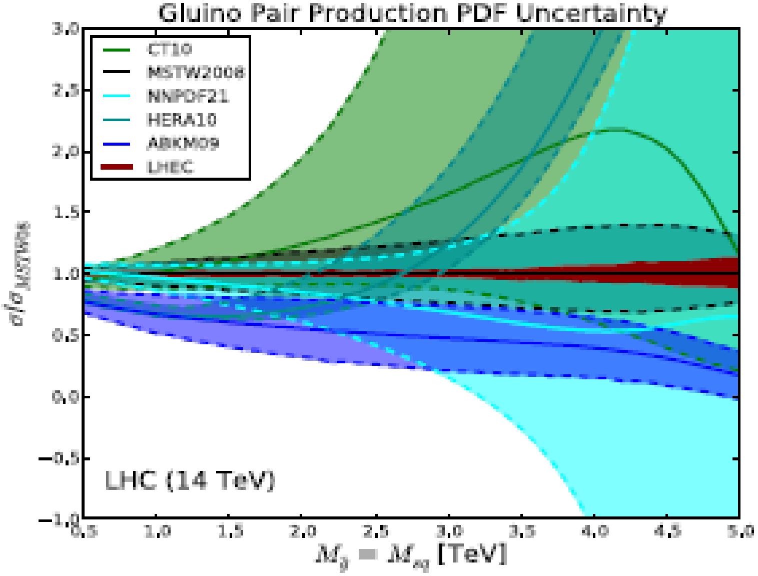
Drell-Yan is a term for q-qbar → $\mu^+ \mu^-$ collisions mediated by Z or virtual γ,Z bosons.
Some new theories predict higher mass Z' states, these have been excluded up to 2 TeV
The main reason we cannot do better is that the PDF uncertainty on the 'normal' Standard Model background is too big.

Consequence of uncertainty in the high-x gluon?-one example

Many interesting processes at the LHC are gluon-gluon initiated
…BSM processes like gluon-gluon → gluino-gluino
And the high-scale needed for this involves the high-x gluon
The gluon-gluon luminosity at high-scale is not well-known
This leads to uncertainties on the gluino pair production cross section

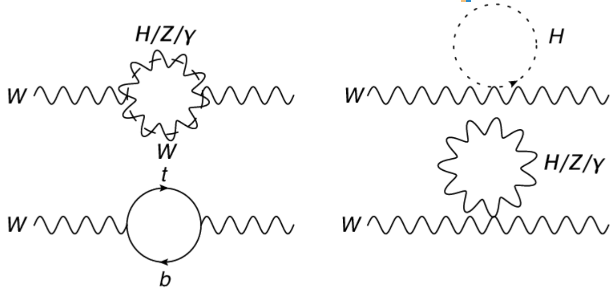




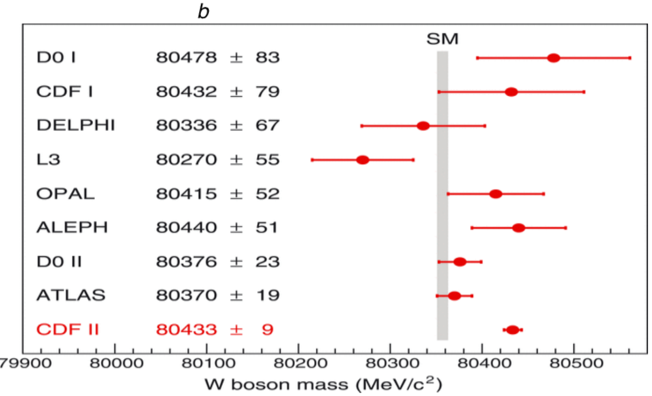


But as well as limiting our ability to identify BSM effects at high M_X , uncertainties on PDFs also limit indirect observations of new physics which we may hope to make by measuring discrepancies from the Standard Model (SM) values for fundamental parameters such as m_W – the W mass

$$m_W^2\left(1-\frac{m_W^2}{m_Z^2}\right)=\frac{\pi\alpha}{\sqrt{2}G_\mu}(1+\Delta r),$$



The W mass is predicted in the SM in terms of other SM parameters like the fine structure constant and the weak coupling G, but Δr represents higher order loops in the diagrams which are presently calculated with known particles like the top quark or Higgs, but could also contain BSM particles.



In that case the value of m_W would differ from its SM value
And indeed that is what we see in the latest Fermilab measurement!
BUT how can this be checked?

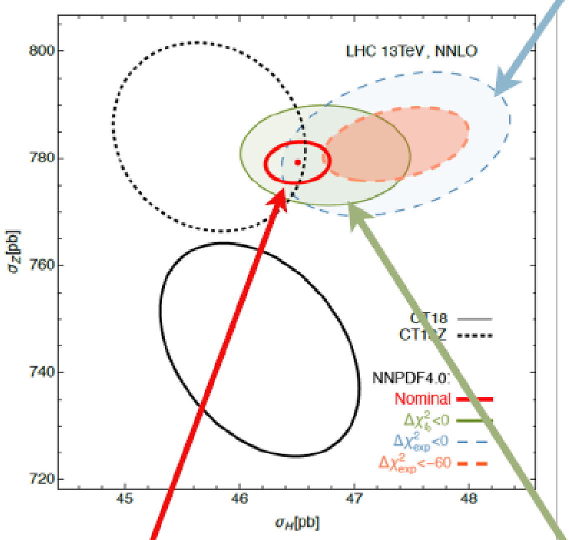
Well it can be checked at the LHC.
The most accurate LHC measurement to date is from ATLAS and is shown on the plot.
A major contribution to its uncertainty of 19MeV is the PDF uncertainty of 10 MeV.
LHC uses p-p not p-pbar and its kinematic reach is such that most collisions producing W are sea-quark collisions. It is not clear that the PDF uncertainties can be improved quickly .Recently an overall uncertainty of 16MeV was achieved—still not good enough

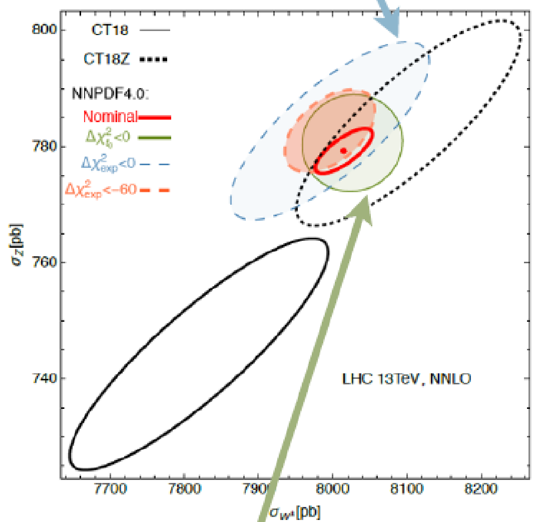
Monte-Carlo sampling sensitivity for PDFs

Debate on NNPDF4.0 uncertainties

Regions containing (very) good solutions according to the experimental form of χ^2

(is used in χ^2 summary tables of the NN4.0 article, was a default in the NN4.0 public code)





CT criticism ArxiV:2205.10444

Nominal NN4.0 Hessian or MC 68%cl

Region containing good solutions according to the most recent t_0 form of χ^2 (used to train NN4.0 replicas)

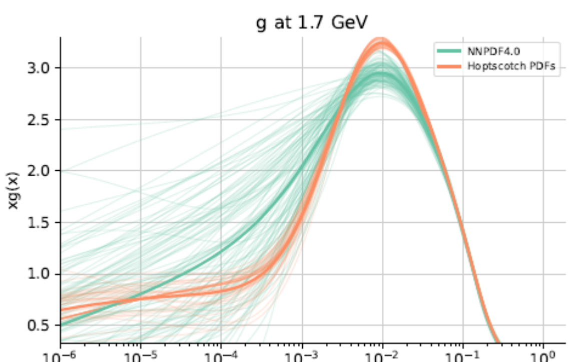
NNPDF prior disfavors some of the most displaced hopscotch replicas.

[M. Ubiali, HP2 2022 workshop, Durham, 2022/09]

The hopscotch replicas pass CT18 criteria for good individual solutions.

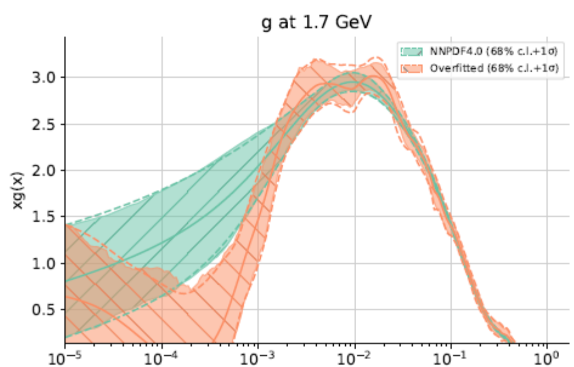
Observation 1

A comparison between NNPDF replicas and HS PDFs reveals a peculiar kink in HS PDFs



Observation 2

An overfitted gluon can be obtained with $\chi^{2(0,c)}$ smaller than NNPDF4.0 by $0.08 \times N_{\rm dat}$

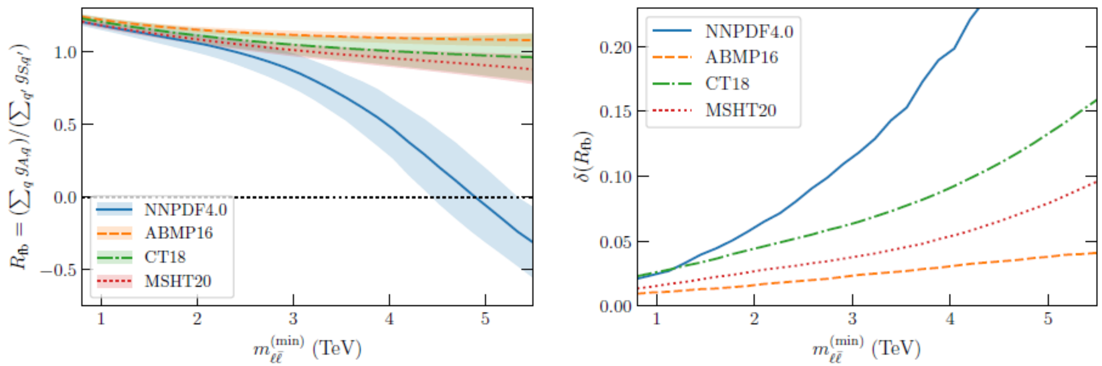


NNPDF reply at PDF4LHC2022… but it goes on

Comparisons at very high-x / High scale
AFB is very different for NNPDF4.0
NNPDF4.0 uncertainties remain large/largest beyond the current data region– but not large enough to cover this

Positive or negative asymmetry?

$$A_{\mathrm{fb}}(\cos\theta^*) \overset{\mathrm{LO}}{=} \frac{\cos\theta^*}{1+\cos^2(\theta^*)} \frac{g_A}{g_{S,q'}} \qquad g_{A,S} \propto \int \frac{dm_{\ell\bar{\ell}}}{m_{\ell\bar{\ell}}} \frac{dx_1}{x_1} \mathcal{L}_{\mathcal{A},\mathcal{S}}(m_{\ell\bar{\ell}}, x_1)$$



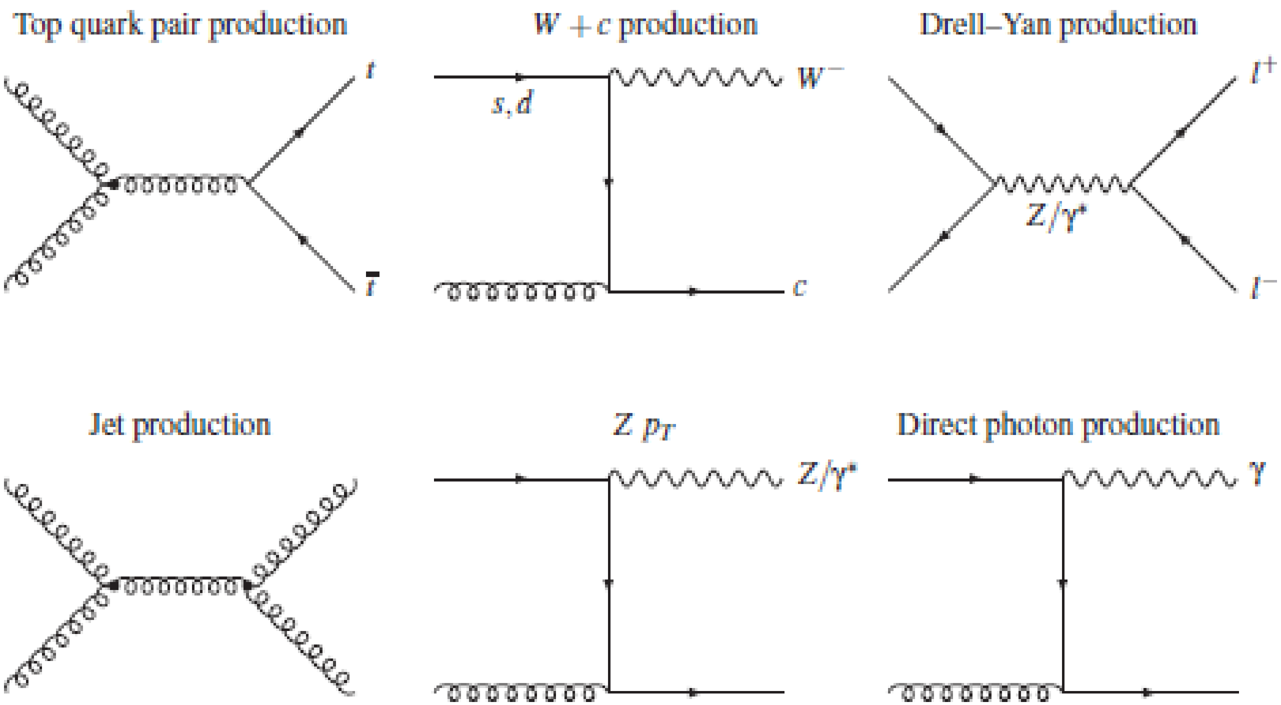
As the scale increases, larger x is probed

ABMP16, CT18 and MSHT20:
approximately scale independent

NNPDF: the coupling depends on the scale, larger uncertainty

Expectation: for NNPDF4.0, A_{fb} vanishes at large $m_{\ell\bar{\ell}}$

A study of potential improvements has been made using processes for which are now statistics limited, where the High-Luminosity LHC (HL-LHC) should help



Pseudo-data is generated for these processes assuming luminosity of 3 ab^{-1} for CMS and ATLAS and 0.3 ab^{-1} for LHCb

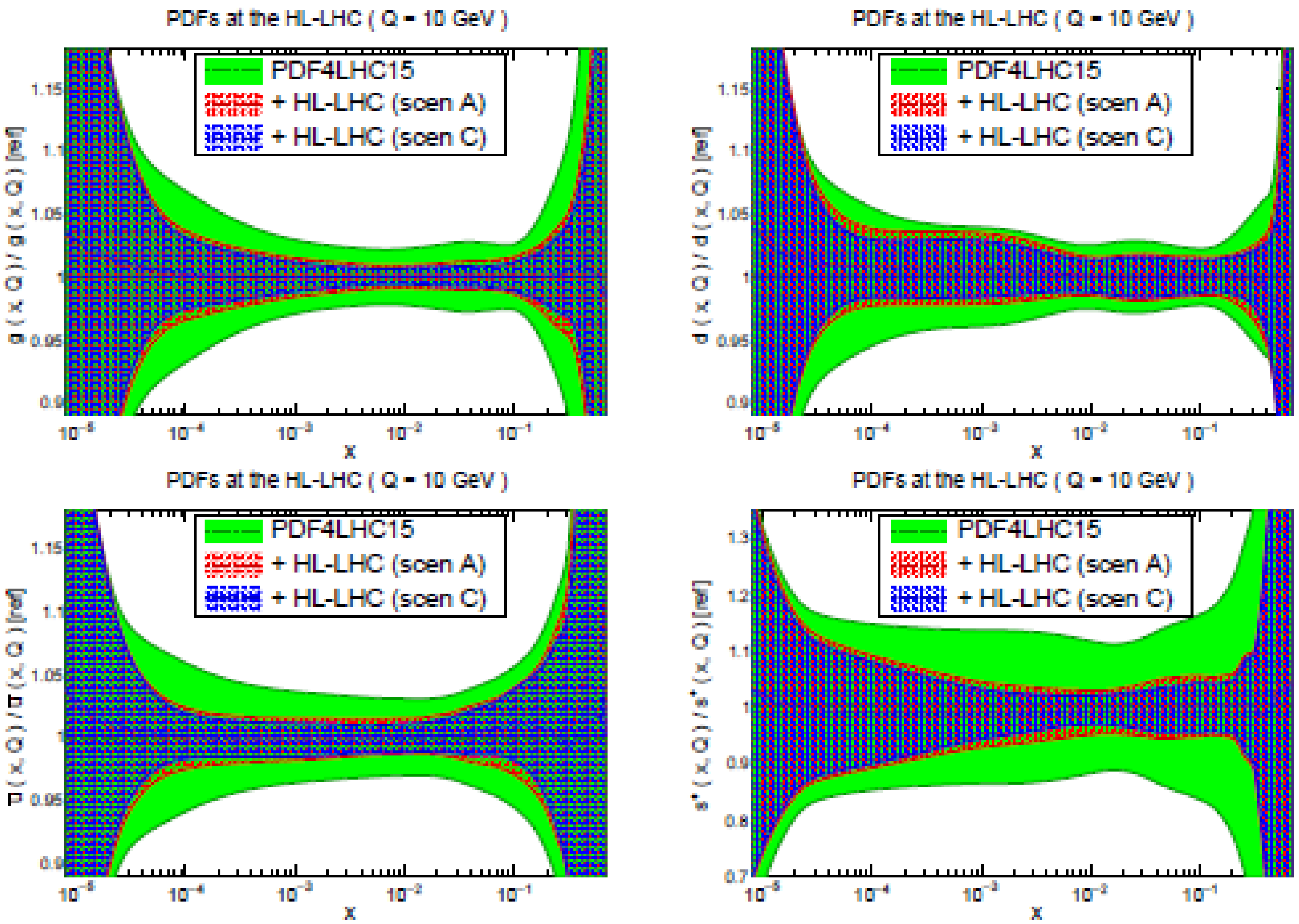
Pessimistic and Optimistic assumptions are made about systematic uncertainties based on experience with real data

Both about the effect of correlations-- typically, f_{corr} = 1, 0.25

And about possible reduction in uncertainty typically, f_{red} = 1, 0.4

This is about as good as you can do with pseudo-data but let's not forget that this is a somewhat ideal situation

So we see potential improvements in the PDFs



Where scenario A is pessimistic and scenario C is optimistic
--Such improvements could give up to a factor 2 improvement in the PDF uncertainty on something like m_W ----**but such estimates are unlikely to be fully realistic…**