Z+charm at the LHC: implications for proton structure

Juan Rojo, VU Amsterdam & Nikhef

LHC Electroweak Working Group Meeting:

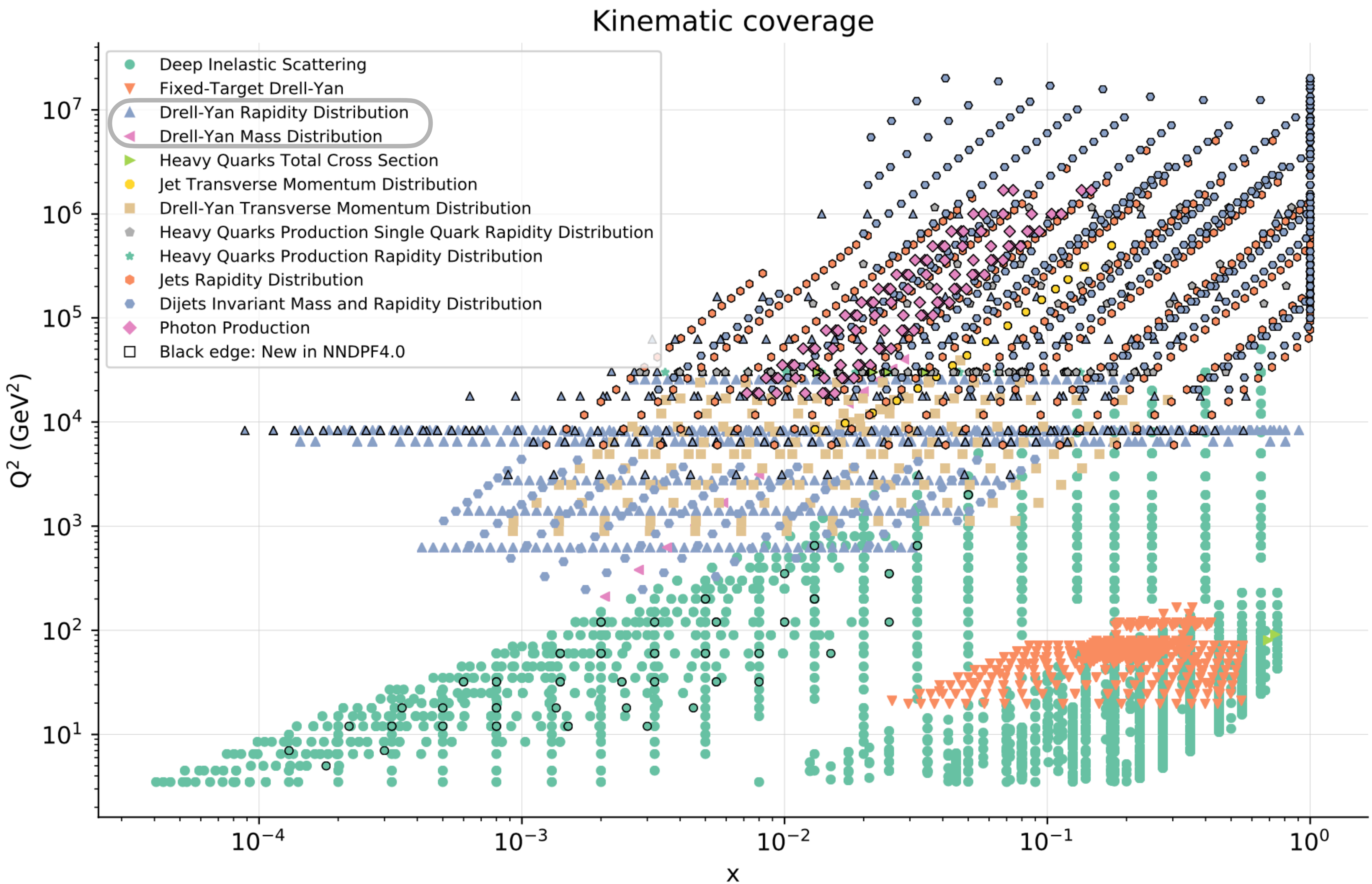
Jets and EW bosons

18th October 2021

Heavy quark production in perturbative QCD

V+jets in PDF fits

Inclusive measurements of ***W* and *Z* boson production**: key building block of global fits

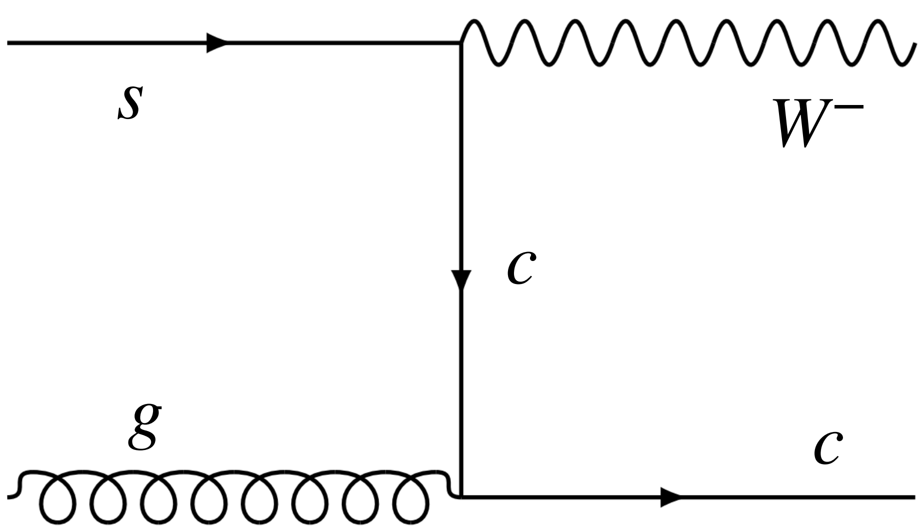


V+jets in PDF fits

Inclusive measurements of ***W* and *Z* boson production**: key building block of global fits

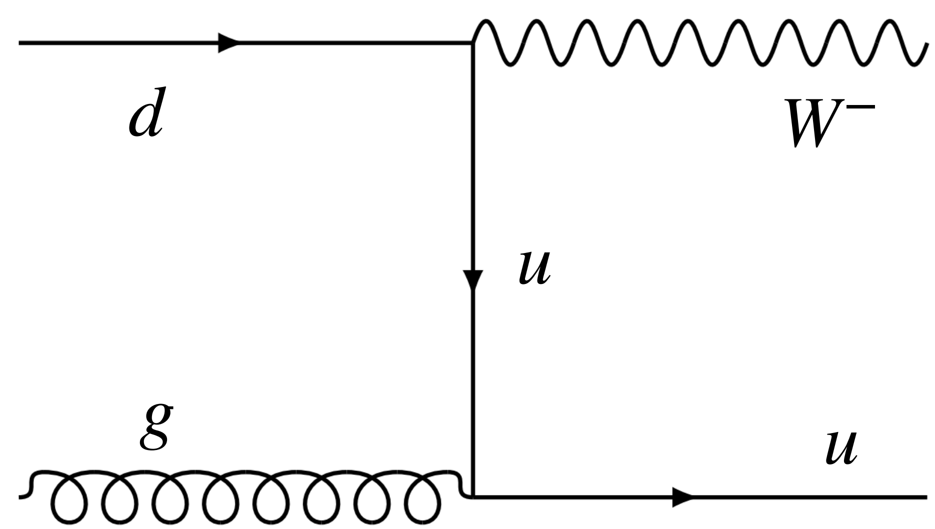
Measurements of ***W* and *Z* boson in association with jets** provide sensitivity to complementary PDF flavour combinations

W+charm



Direct handle on strangeness

W+jets



Light quark flavour separation

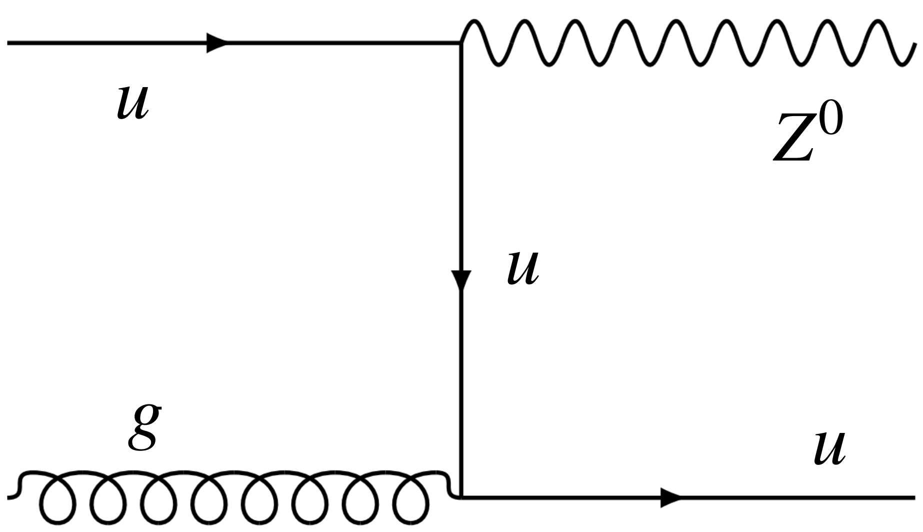
Both types of processes can be treated using **NNLO QCD theory** and are part of the NNPDF4.0 determination

V+jets in PDF fits

Inclusive measurements of ***W* and *Z* boson production**: key building block of global fits

Measurements of ***W* and *Z* boson in association with jets** provide sensitivity to complementary PDF flavour combinations

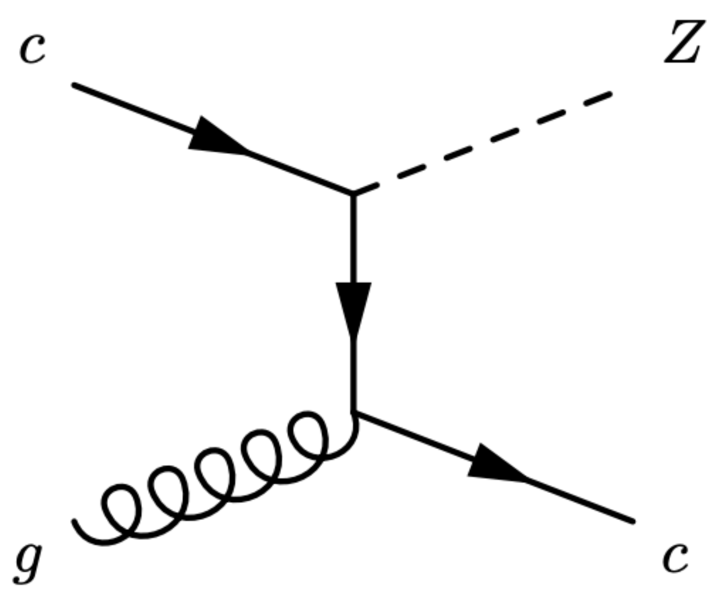
Z+jets & Z p_T



Constrains the gluon PDF for x ≃ 0.01, relevant for Higgs production in gluon fusion

nb Z p_T vanishes at Born level

Z+charm



Direct handle on the charm content of the proton

The **Z p_T distributions** from ATLAS and CMS are routinely added in global fits

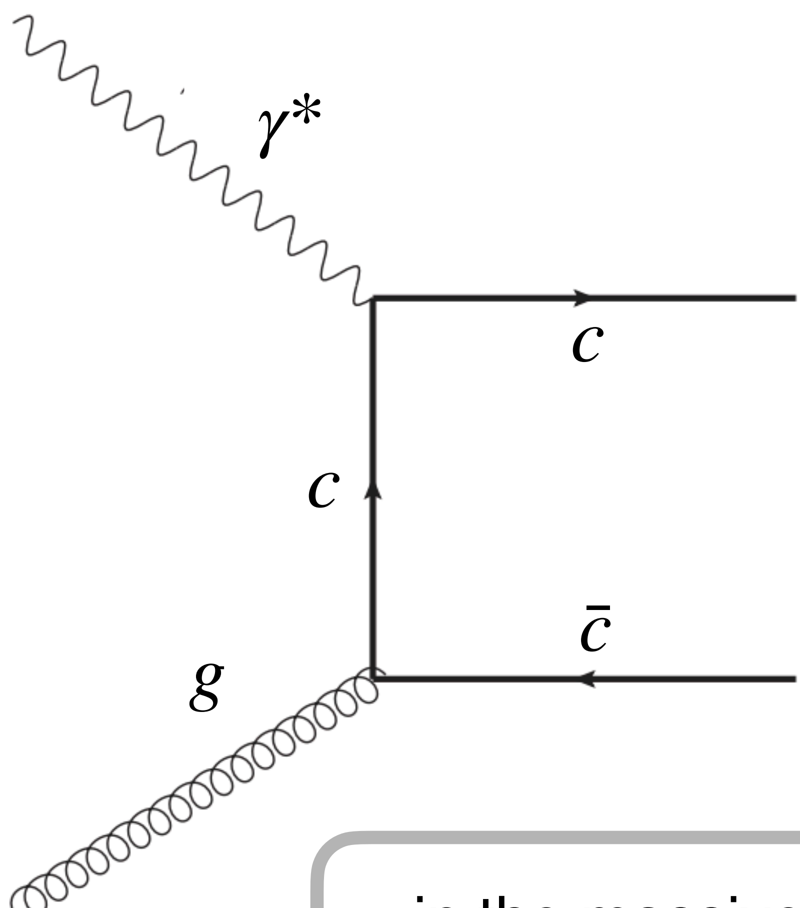
What about Z+charm? What do we really learn by measuring the charm PDF?

A tale of three schemes

say you want to evaluate the **charm DIS structure function**. You have three options

Let us assume for the time being that there is no ``Intrinsic Charm'': the non-perturbative proton wave function does not have a charm component

☑ **Fixed-flavor scheme:** no charm PDF, charm mass effects accounted for exactly



$$F_2^c(x,Q^2) \propto \sum_{i=g,u,d,s} C_i^{(n_f)}(\alpha_s, Q^2/m_c^2) \otimes f_i^{(n_f)}$$

no charm PDF

exact in **threshold region**

not appropriate to describe ***$Q^2 >> m_c^2$*** region due to l**arge unresummed logs**

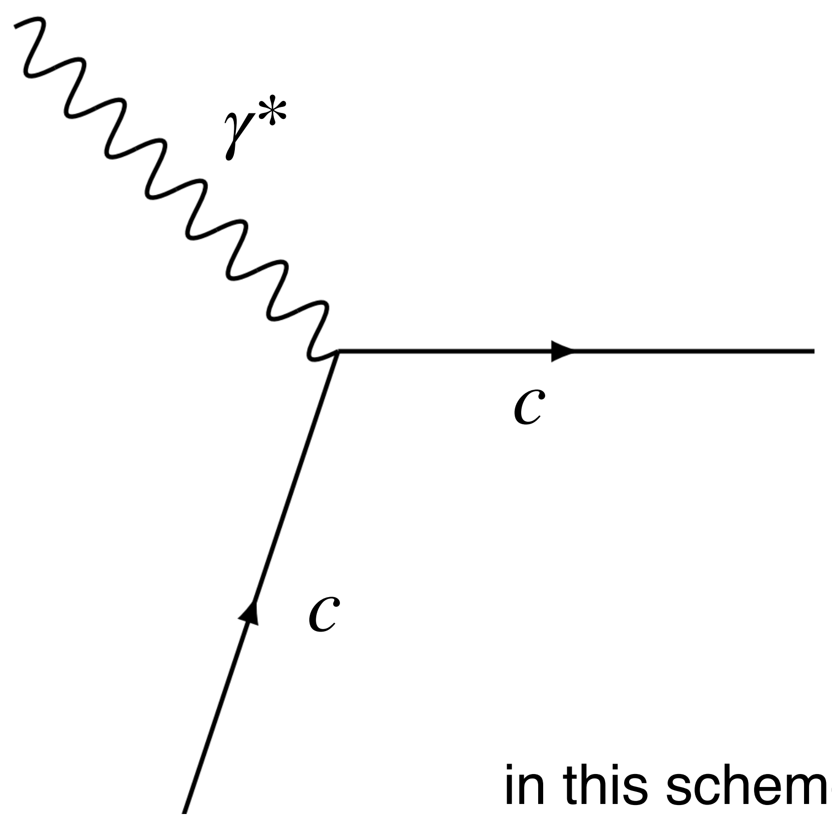
in the massive FFN scheme, **nothing to say** about the charm PDF

The charm PDF does not exist here!

A tale of three schemes

say you want to evaluate the **charm DIS structure function**. You have three options

☑ **Zero-mass scheme:** charm PDF treated on the same footing as all other quark flavours



$$F_2^c(x, Q^2) \propto \sum_{i=g,u,d,s,c} C_i^{(n_f+1)}(\alpha_s) \otimes f_i^{(n_f+1)}$$

exact far from threshold region

not appropriate to describe $\boldsymbol{Q^2 \lessapprox m_c^2}$

region due to lack of **massive corrections**

in this scheme, the charm PDF above threshold is constructed from the nf=3 PDFs via **matching**. If there is no charm PDF for $\boldsymbol{n_f=3}$ (assumption) then

$$f_c^{(n_f+1)} \propto \alpha_s \ln \frac{Q^2}{m_c^2} \left(P_{qg} \otimes f_g^{(n_f+1)} \right) + \mathcal{O}\left(\alpha_s^2\right)$$

the charm PDF is deterministically **generated from the gluon** (and light quark) PDFs

Here charm PDF is ``trivial": little to learn from its measurement

A tale of three schemes

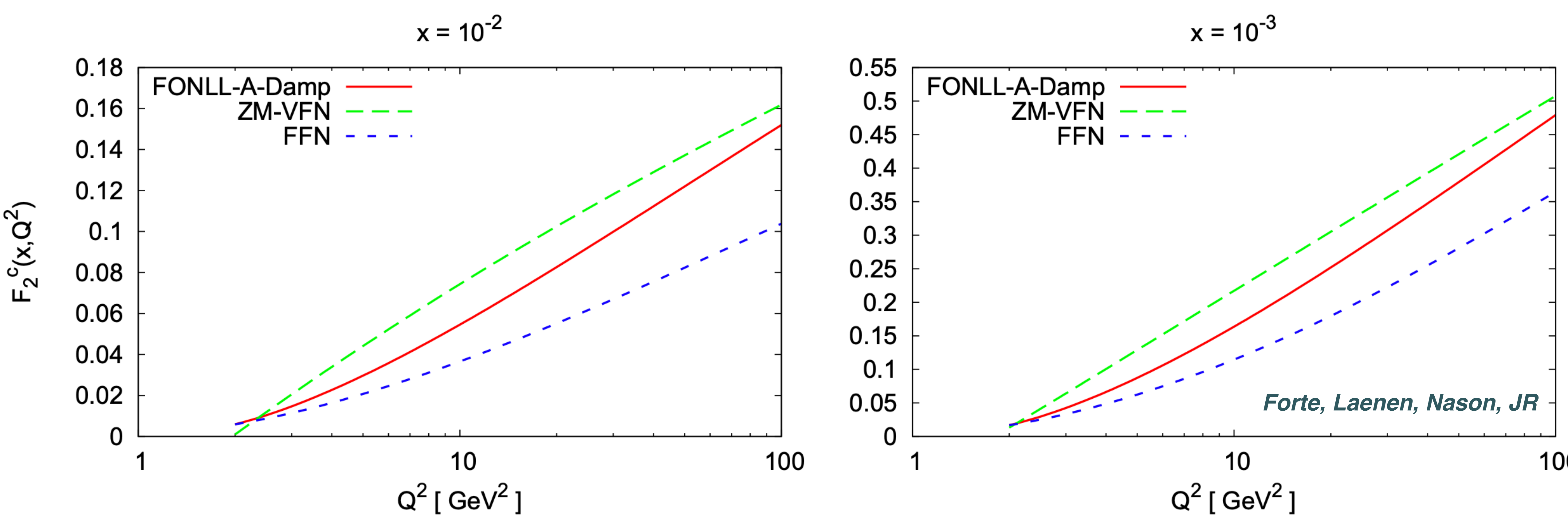
say you want to evaluate the **charm DIS structure function**. You have three options

- **General-mass VFN scheme:** charm PDF treated on the same footing as all other quark flavours, massive effects included in coefficient functions

$$F_2^c(x,Q^2) \propto \sum_{i=g,u,d,s,c} C_i^{(\mathrm{GM})}(\alpha_s, Q^2/m_c^2) \otimes f_i^{(n_f+1)}$$

charm PDF included in sum

Systematically improvable, **reliable for all values of *Q*^{2}** from threshold to collider scales



The charm PDF

A tale of three schemes

so far we have assumed that charm is **purely perturbative**. But it does not need to be so! An **intrinsic component** is allowed and even predicted in many models of nucleon structure

THE INTRINSIC CHARM OF THE PROTON

S.J. BRODSKY [1]
Stanford Linear Accelerator Center,
Stanford, California 94305, USA

and

P. HOYER, C. PETERSON and N. SAKAI [2]
NORDITA, Copenhagen, Denmark

Received 22 April 1980

Recent data give unexpectedly large cross-sections for charmed particle production at high x_F in hadron collisions. This may imply that the proton has a non-negligible uudc$\bar{c}$ Fock component. The interesting consequences of such a hypothesis are explored.

Valence-like charm PDF predicted, peaked at x=0.4

Ultimately, **data has to decide**: a phenomenological determination of the charm PDF will quantify the relative weight of the perturbative vs intrinsic components

$$c^{(n_f=5)}(x,Q) \simeq c^{(n_f=5)}_{\rm (pert)}(x,Q) + c^{(n_f=5)}_{\rm (intr)}(x,Q)$$

from pQCD evolution and matching

from intrinsic component $c^{(n_f=3)}_{\rm (intr)}(x) \neq 0$

perturbative, intrinsic, and fitted charm

The charm PDF in the **4-flavour scheme** (above charm threshold) can be determined as:

- ***Perturbative charm***: the charm PDF vanishes below threshold (no charm in **3-flavour scheme**), above threshold ($\mu_c \gtrsim m_c$) deterministically **generated from the gluon** (and light quark) PDFs

$$f_c^{(n_f)} = 0 \quad \rightarrow \quad f_c^{(n_f+1)} \propto \alpha_s \ln \frac{Q^2}{m_c^2} \left(P_{qg} \otimes f_g^{(n_f+1)} \right) + \mathcal{O}\left(\alpha_s^2\right)$$

3FNS charm *4FNS charm* *``trivial" charm PDF*

- ***Intrinsic charm***: a **model** for the charm PDF in the **3-flavour scheme** is assumed, then evolved with DGLAP: combination of (model-dependent) intrinsic and perturbative components

$$f_c^{(n_f)}(x, Q_0) = A x^2 \left[6x(1+x)\ln x + (1-x)(1+10x+x^2) \right]$$

BHPS model (scale independent)

the model parameters (*e.g.* normalisation) are extracted from comparison with data

- ***Fitted charm***: no assumptions on possible intrinsic component are made. The charm is parametrised **above threshold (4FNS)** in exactly the same was as all other quark PDFs

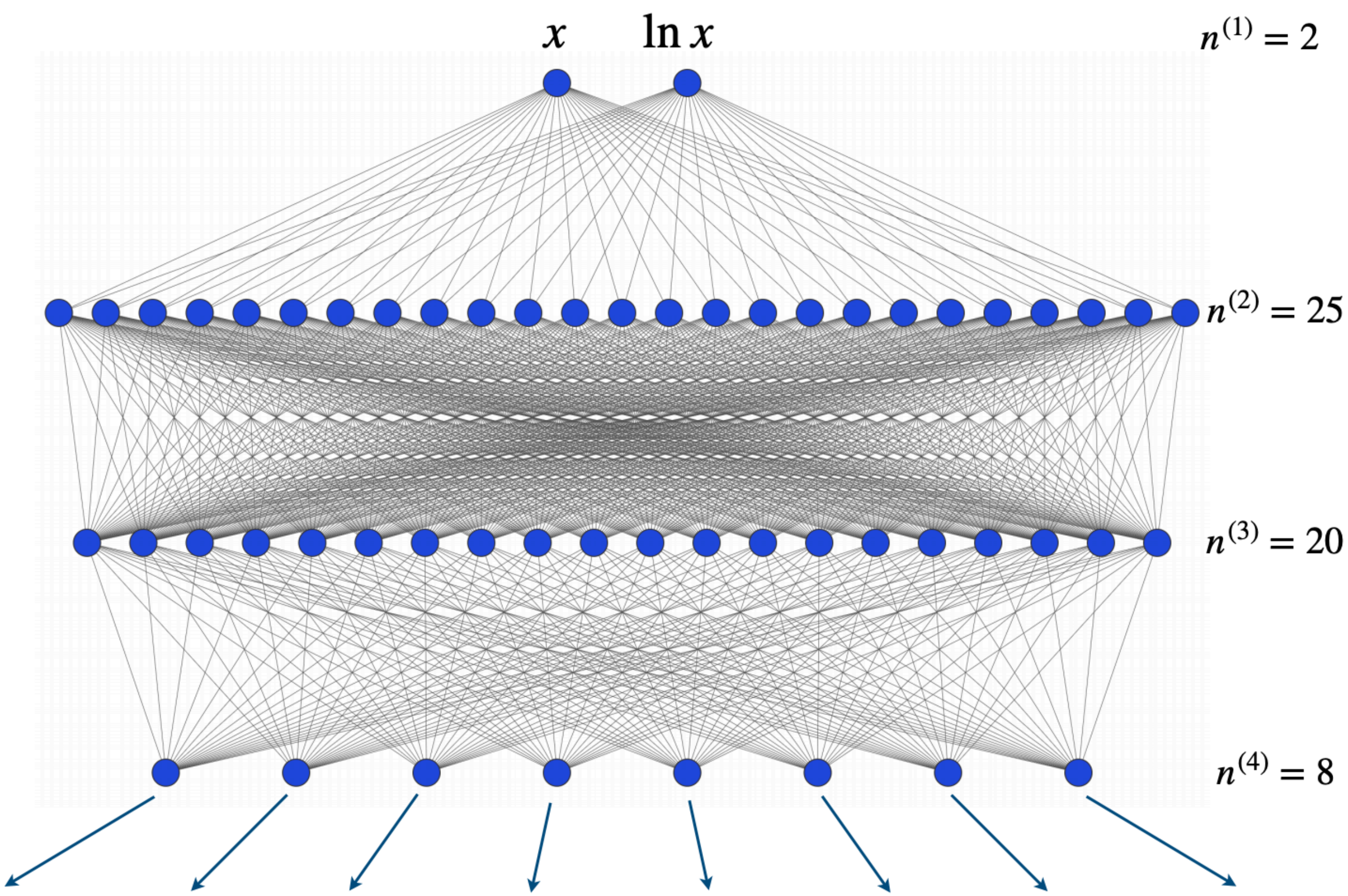
NNPDF approach

$$f_c^{(n_f+1)}(x, Q_0) = x^{-\alpha_c}(1-x)^{\beta_c}\mathrm{NN}(x)$$

Let the data tell us whether or not there is an intrinsic charm component!

n.b. the GM-VFN structure functions need to be modified for a non-zero charm PDF in the n_f=3 scheme

perturbative, intrinsic, and fitted charm



NNPDF4.0 PDF parametrisation

$$Q_0 \geq m_c$$

$xg(x, Q_0)$	$x\Sigma(x, Q_0)$	$xV(x, Q_0)$	$xV_3(x, Q_0)$	$xV_8(x, Q_0)$	$xT_3(x, Q_0)$	$xT_8(x, Q_0)$	$xT_{15}(x, Q_0)$	***evolution basis***
$xg(x, Q_0)$	$xu(x, Q_0)$	$x\bar{u}(x, Q_0)$	$xd(x, Q_0)$	$x\bar{d}(x, Q_0)$	$xs(x, Q_0)$	$x\bar{s}(x, Q_0)$	$xc^+(x, Q_0)$	***flavour basis***

- ***Fitted charm***: no assumptions on possible intrinsic component are made. The charm is parametrised **above threshold (4FNS)** in exactly the same was as all other quark PDFs

NNPDF approach

$$f_c^{(n_f+1)}(x, Q_0) = x^{-\alpha_c}(1-x)^{\beta_c}\mathrm{NN}(x)$$

Let the data tell us whether or not there is an intrinsic charm component!

Why fitting charm?

1) Fit quality of global analysis with **fitted charm** superior to that with **perturbative charm**

Process	Dataset	$n_{\rm dat}$	$\chi^2_{\rm base}$	$\chi^2_{\rm pr}$	$\chi^2_{\rm str}$	$\chi^2_{\rm str\_s\_hat}$	$\chi^2_{\rm str\_pch}$
νDIS ($\mu\mu$)		76/76/95/91/95	0.70	0.71	0.53	0.52	0.63
	NuTeV [9]	76/76/76/76/76	0.70	0.71	0.53	0.55	0.61
	NOMAD [10]	—/—/19/15/19	[9.0]	[8.8]	0.55	0.35	0.69
W, Z (incl.)		327/418/418/418/418	1.38	1.40	1.40	1.39	1.40
	ATLAS [12]	—/61/61/61/61	3.22	1.65	1.67	1.64	1.80
$W{+}c$		—/37/37/37/37	[0.76]	0.68	0.60	0.66	0.68
	CMS [17,18]	—/15/15/15/15	[1.10]	0.98	0.96	1.00	1.00
	ATLAS [16]	—/22/22/22/22	[0.53]	0.48	0.42	0.43	0.46
$W{+}$jets	ATLAS [15]	—/32/32/32/32	[1.58]	1.18	1.18	1.18	1.18
Total		3917/4077/4096/4092/4096	1.17	1.17	1.17	1.17	1.20

NNPDF3.1 strangeness study

Fitted charm ***Perturbative charm***

consistent improvement in fit quality, not driven by one specific class of processes

Similar picture in **NNPDF4.0 global analysis**

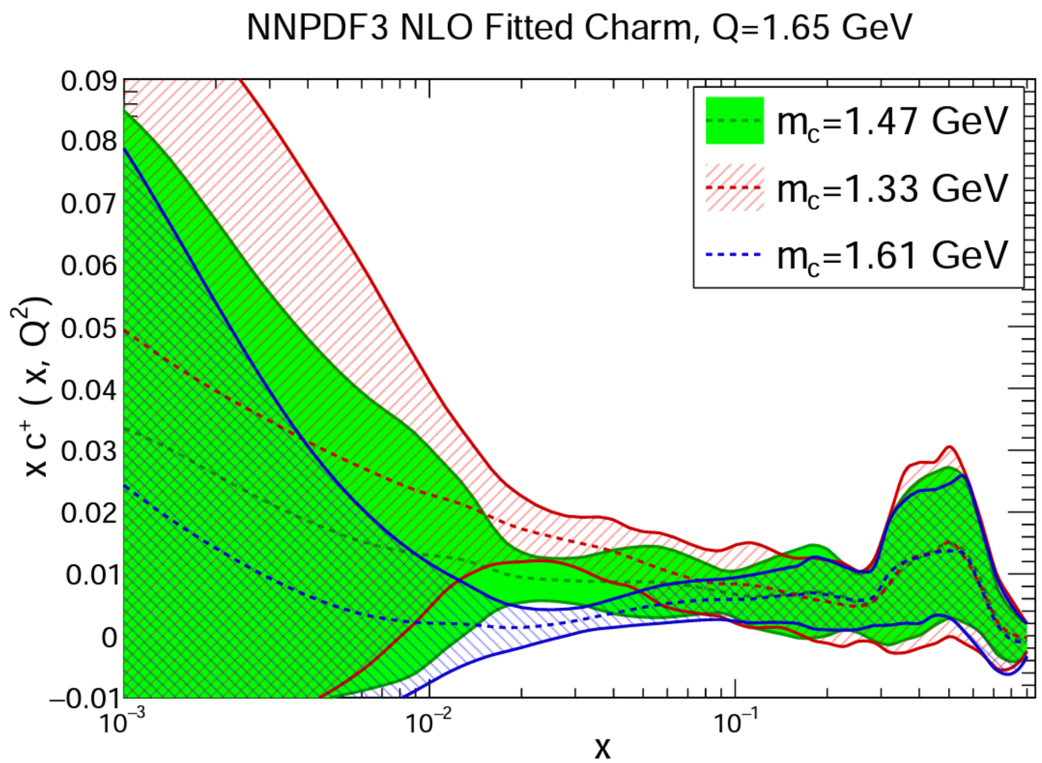
$$\chi^2_{\rm tot} = 1.162\,, \quad \chi^2_{\rm DY} = 1.26\,, \quad \chi^2_{\rm DISnc} = 1.22$$ ***Fitted charm***

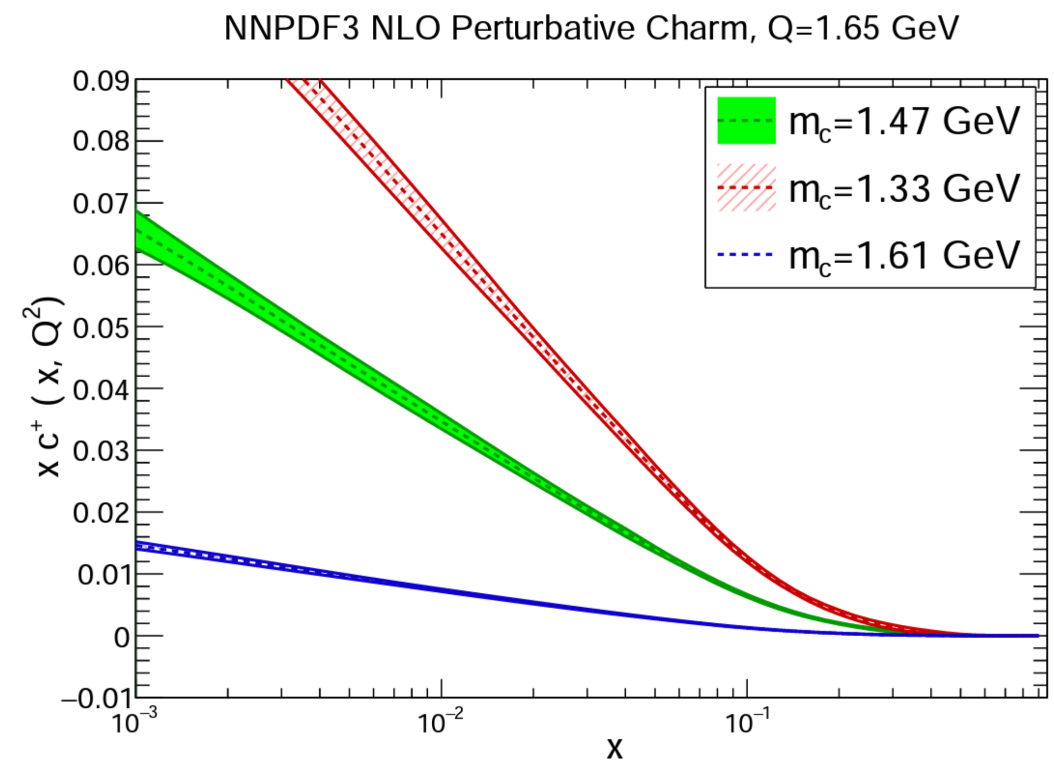
$$\chi^2_{\rm tot} = 1.981\,, \quad \chi^2_{\rm DY} = 1.31\,, \quad \chi^2_{\rm DISnc} = 1.28$$ ***Perturbative charm***

Why fitting charm?

2) Stability of fit results with respect to the value of the **charm mass**

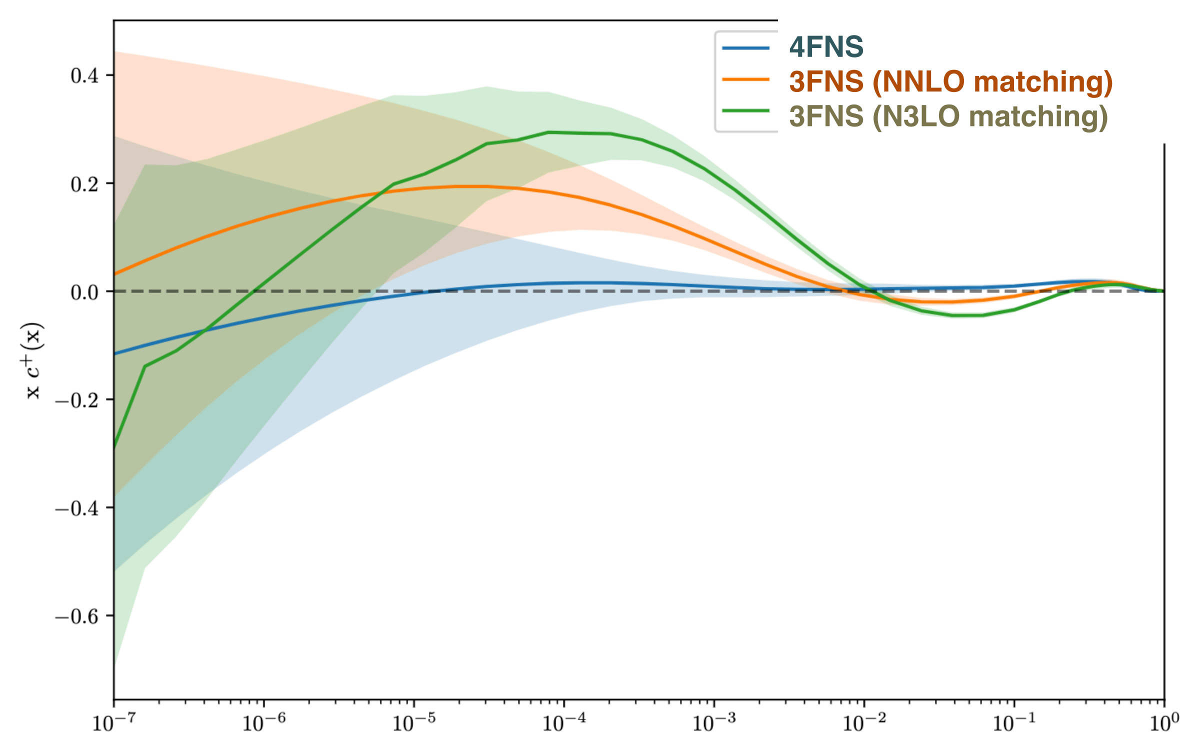
3) Unrealistically **small uncertainties** of charm PDF when generated perturbatively



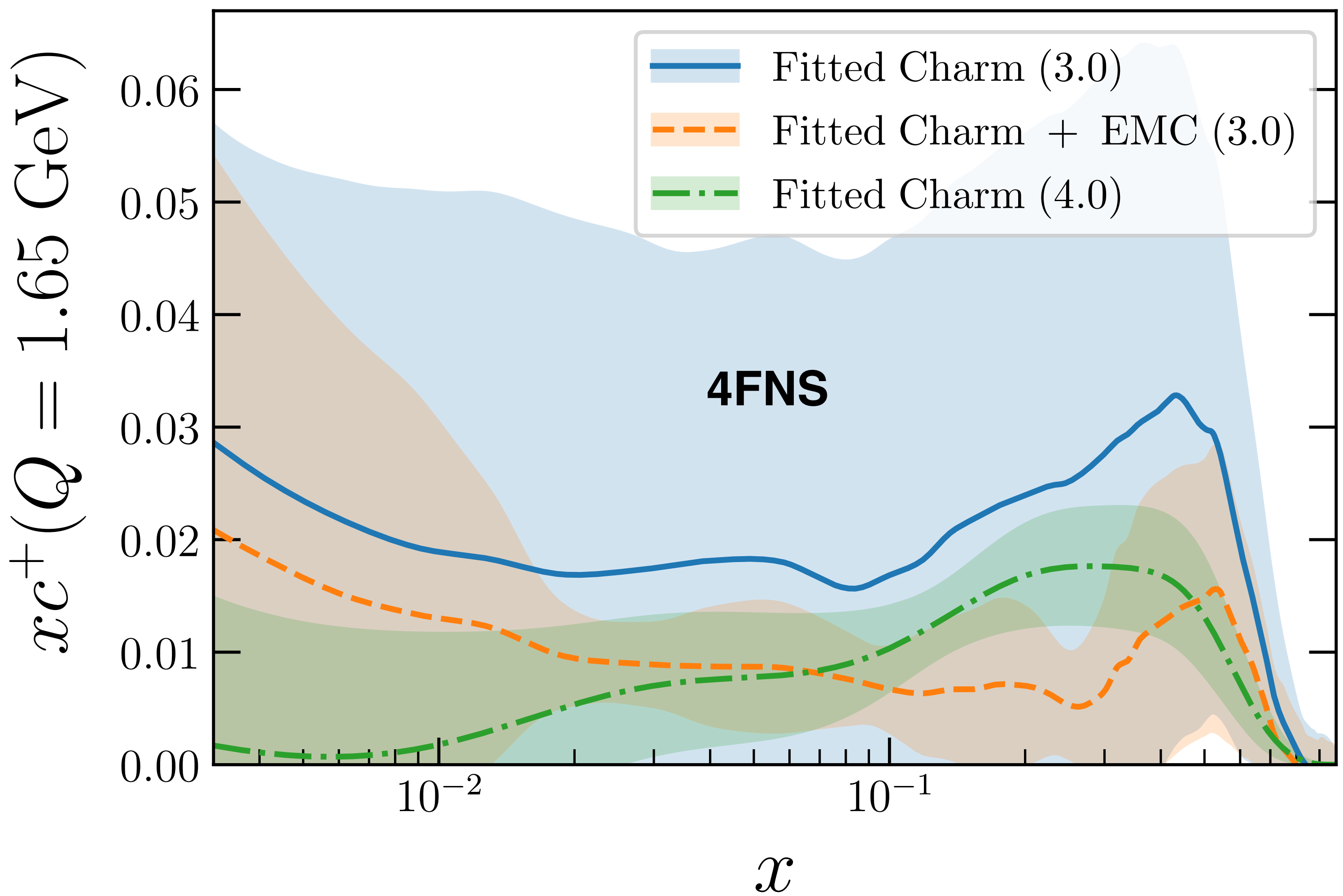


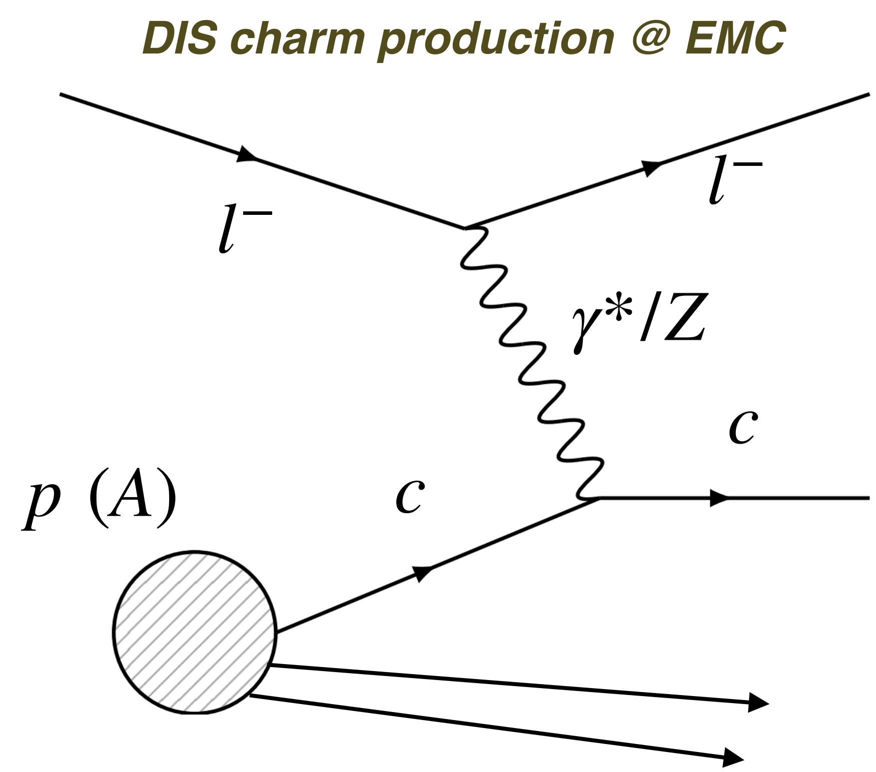
4) Very large MHOUs to the heavy quark matching conditions: **perturbatively unstable**

$$f_i^{(n_f+1)} = A_{ij}^{(n_f)} \otimes f_j^{(n_f)}$$
$$i, j = u, d, s, c, g, \ldots$$



Charm in NNPDF4.0



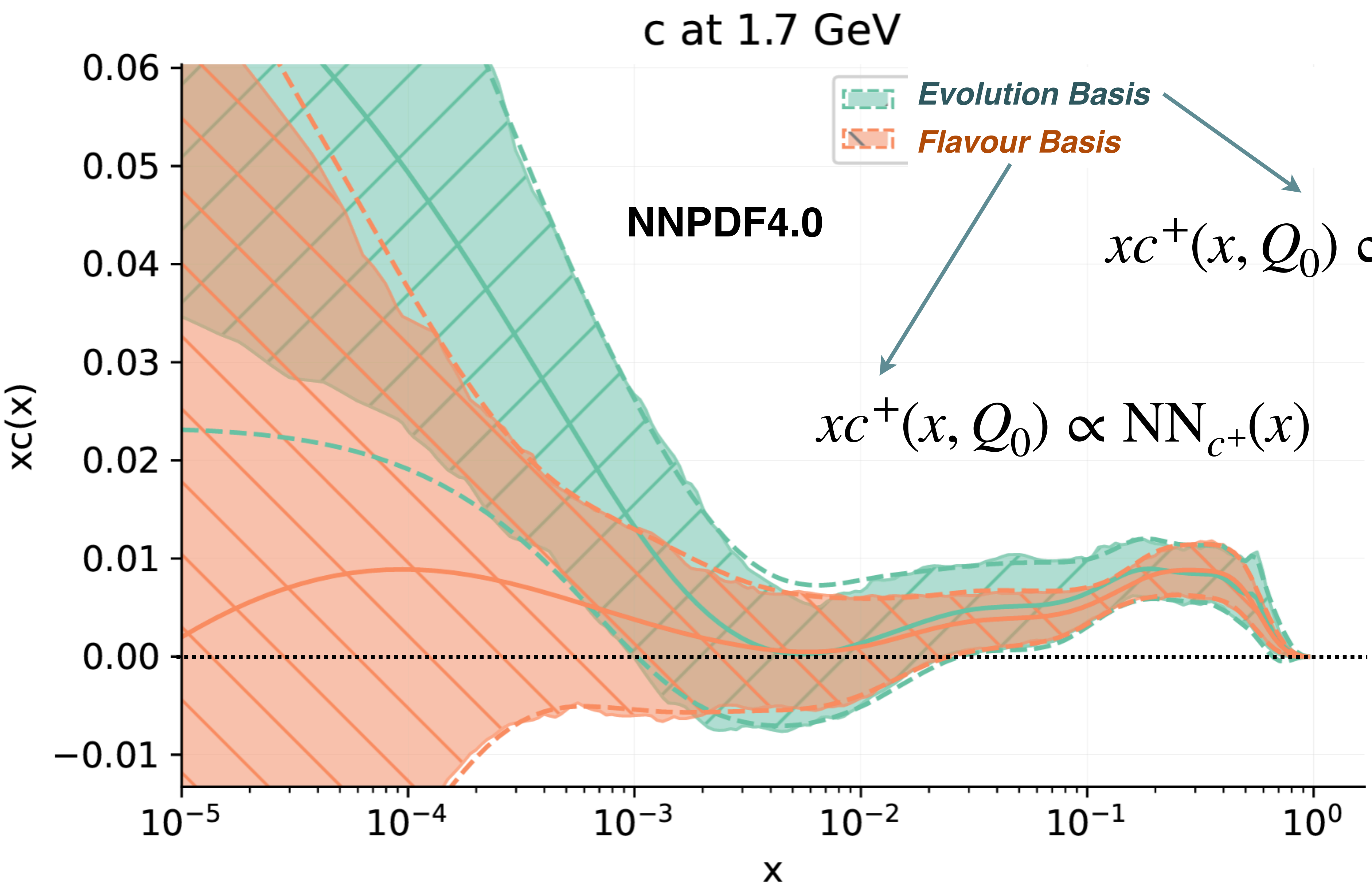


- As compared with the first NNPDF-based **fitted charm** determination, **greatly improved precision** in the recent NNPDF4.0 global analysis

- Bulk of constraints provided by new **precision LHC data**, complemented by fixed-target DIS

- The **EMC charm measurements** (early 80s) provided the dominant constraints on the charm PDF in NNPDF3.0.

Charm in NNPDF4.0

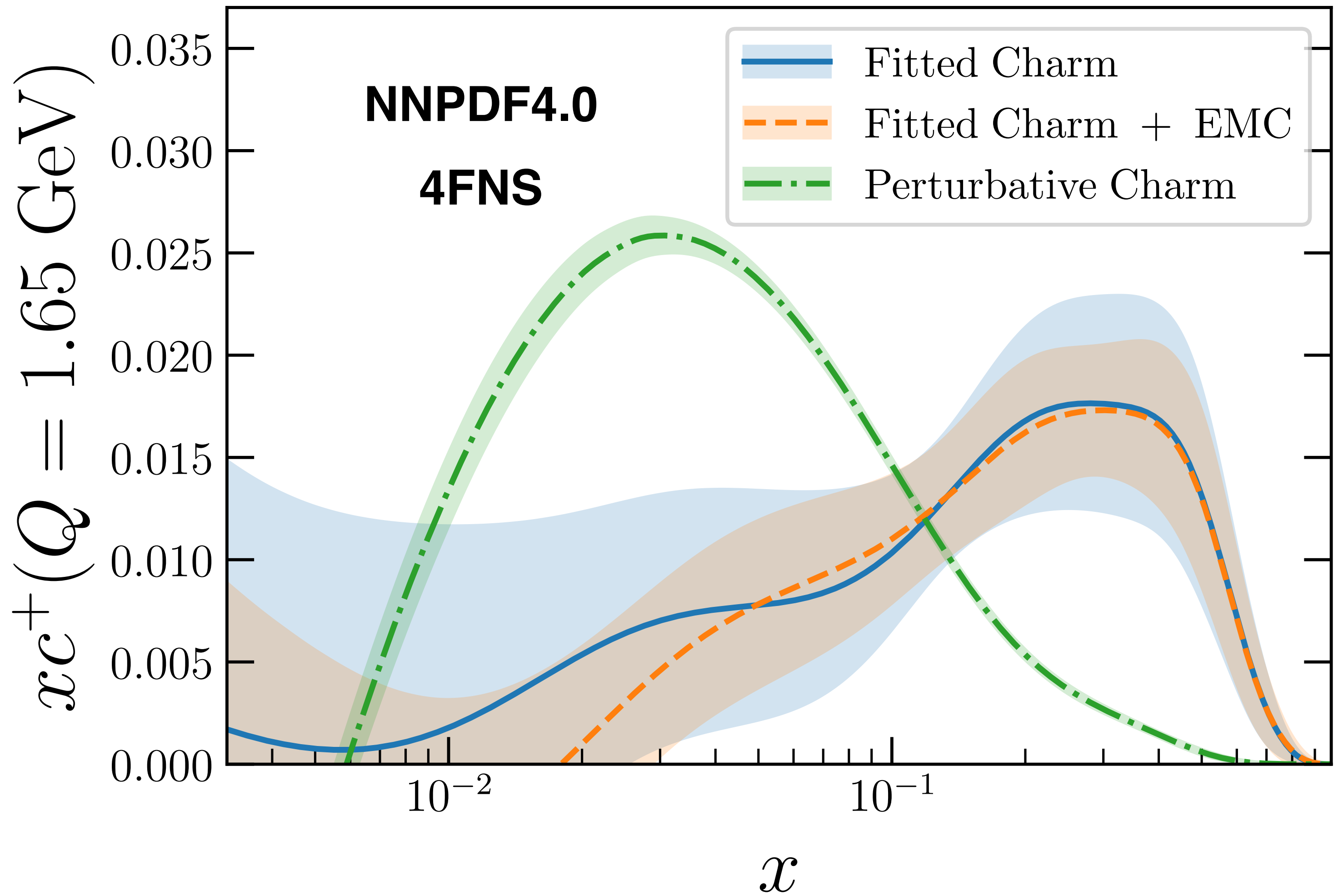


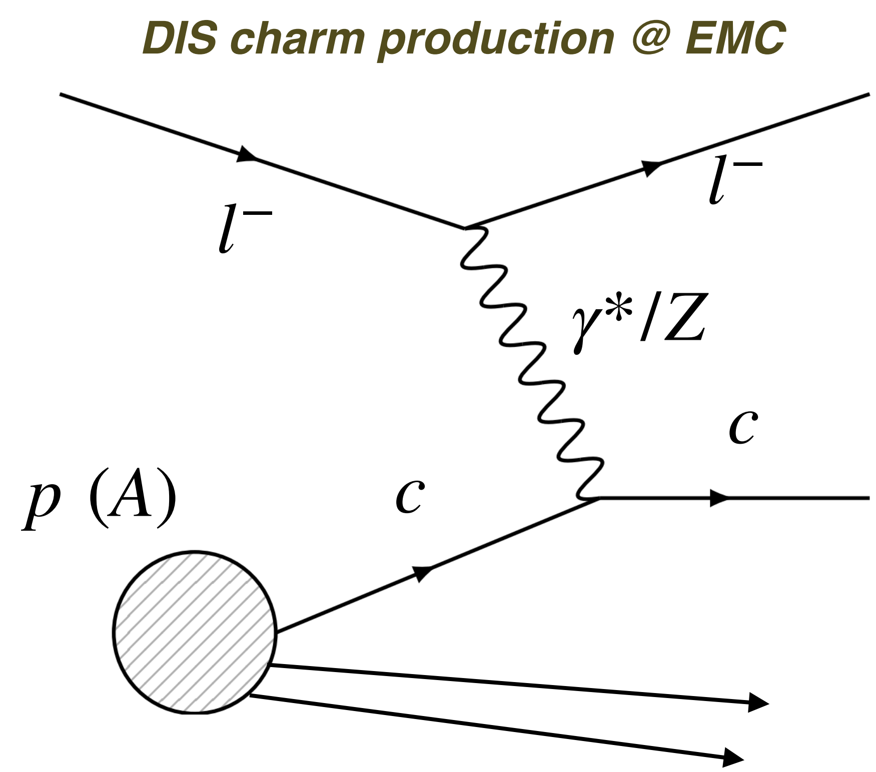
$$xc^+(x,Q_0) \propto \frac{1}{4}\left(\mathrm{NN}_\Sigma(x) - \mathrm{NN}_{T_{15}}(x)\right)$$

$$xc^+(x,Q_0) \propto \mathrm{NN}_{c^+}(x)$$

The fitted charm PDF is stable upon change of input **parameterisation basis:** the **large-*x* enhancement** (intrinsic component?) is genuine feature (no artefact) of the fit

Charm in NNPDF4.0





- In NNPDF4.0, **LHC data dominates** the constraints on the charm PDF, with EMC structure functions having a milder impact now

- Marked differences between fitted and perturbative approaches for the charm PDF, can we **distinguish them** phenomenologically?

- The perturbative calculation depends sensitively on the choice of **charm mass value**

charm mass dependence stabilised with a fitted charm PDF

Charm in NNPDF4.0

To disentangle the perturbative from the intrinsic component of the charm PDF, we need to **evolve backwards** and transform to the 3FNS via **matching conditions**

4FNS

Extract from global PDF fit

$$c^{(n_f=4)}(x, Q_0) \simeq \left(c^{(n_f=4)}_{\text{(pert)}}(x, Q_0) + c^{(n_f=4)}_{\text{(intr)}}(x, Q_0) \right) , \quad Q_0 = 1.65 \text{ GeV}$$

$$\mu_c = m_c = 1.51 \text{ GeV}$$

Backwards DGLAP + matching conditions

3FNS

Intrinsic component of the charm PDF

$$c^{(n_f=3)}(x, Q) \simeq c^{(n_f=3)}_{\text{(intr)}}(x) , \quad Q = 1.5 \text{ GeV}$$

the matching conditions implement the scheme change at the **heavy quark threshold**

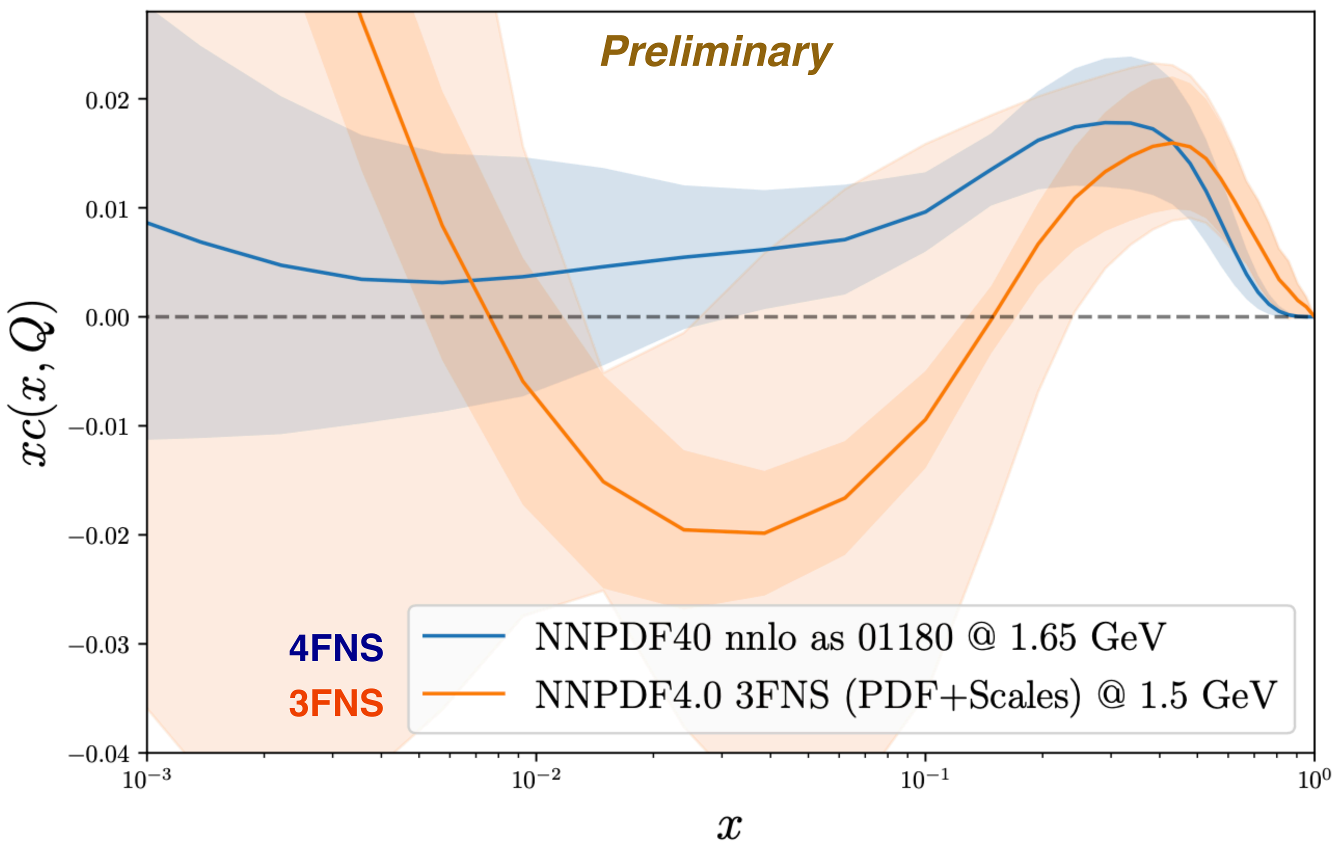
$$f_i^{(n_f+1)} = A_{ij}^{(n_f)} \otimes f_j^{(n_f)} , i, j = u, d, s, c, g, \dots$$

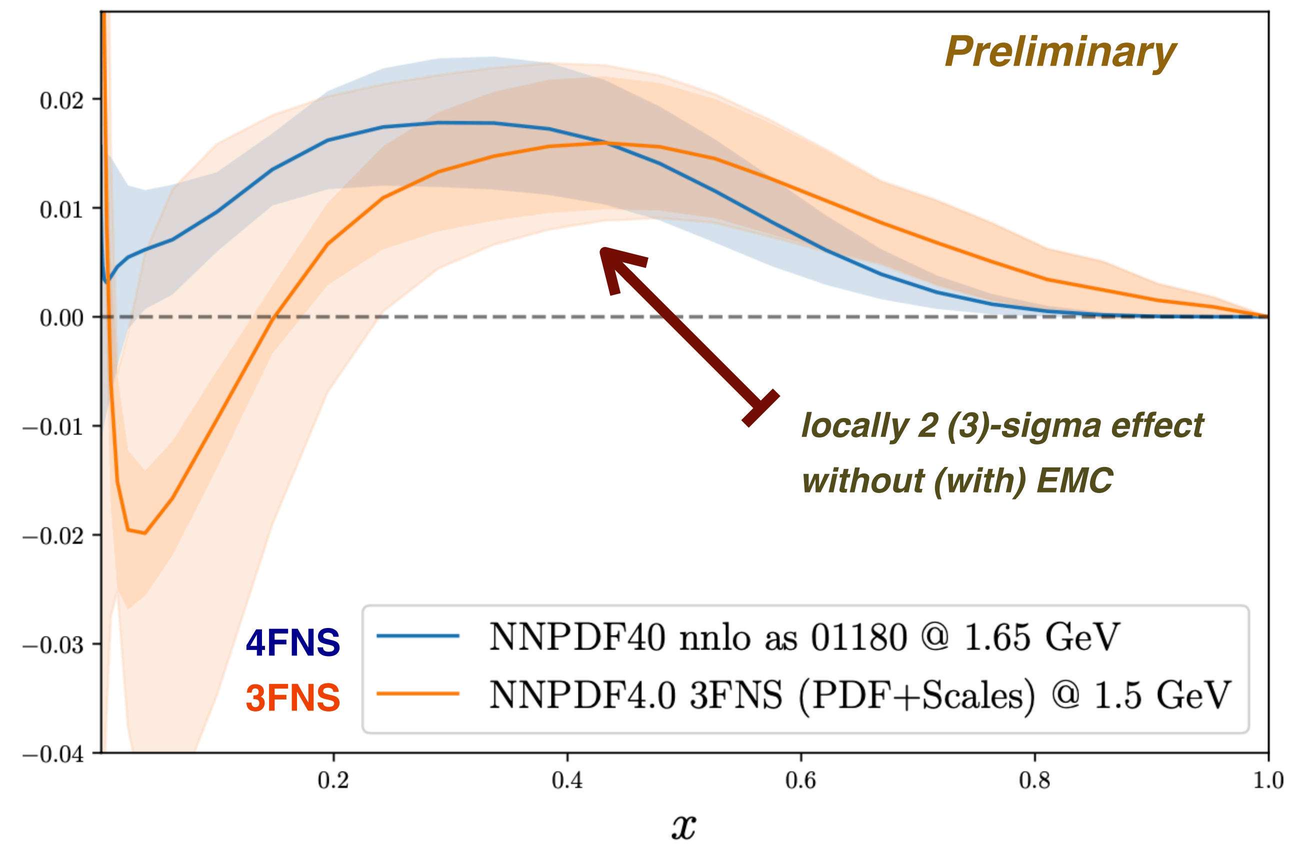
NNLO: Buza et al 92

N3LO: Bluemlein et al.

Charm in NNPDF4.0

To disentangle the perturbative from the intrinsic component of the charm PDF, we need to **evolve backwards** and transform to the 3FNS via **matching conditions**





Estimate MHOUs from the shift between NNLO and N3LO matching

while perturbative uncertainties from the matching are large, one can clearly appreciate a **non-zero component peaked at large-*x*: intrinsic charm**

Calculation carried out with new DGLAP evolution framework: **EKO**

https://eko.readthedocs.io/

EKO
Evolution Kernel Operators

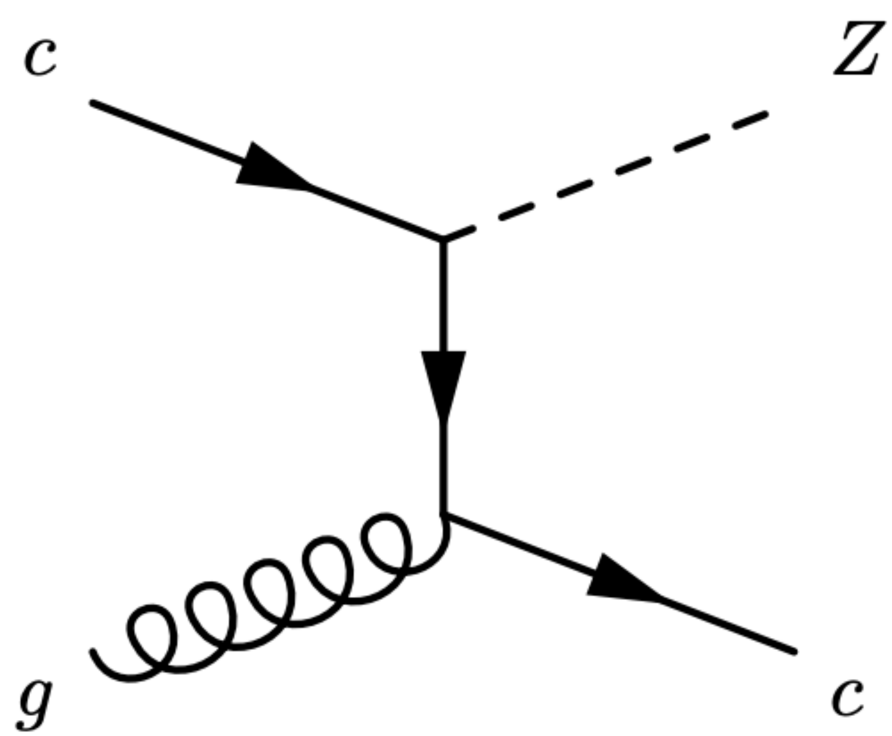
NNPDF Collaboration, in preparation

Candido, Hekhorn, Magni, in preparation

Z+charm at the LHC

Z+charm in PDF fits

Hence what are we learning about **nucleon structure** from the Z+charm process?



if the charm PDF is entirely generated by perturbative QCD evolution

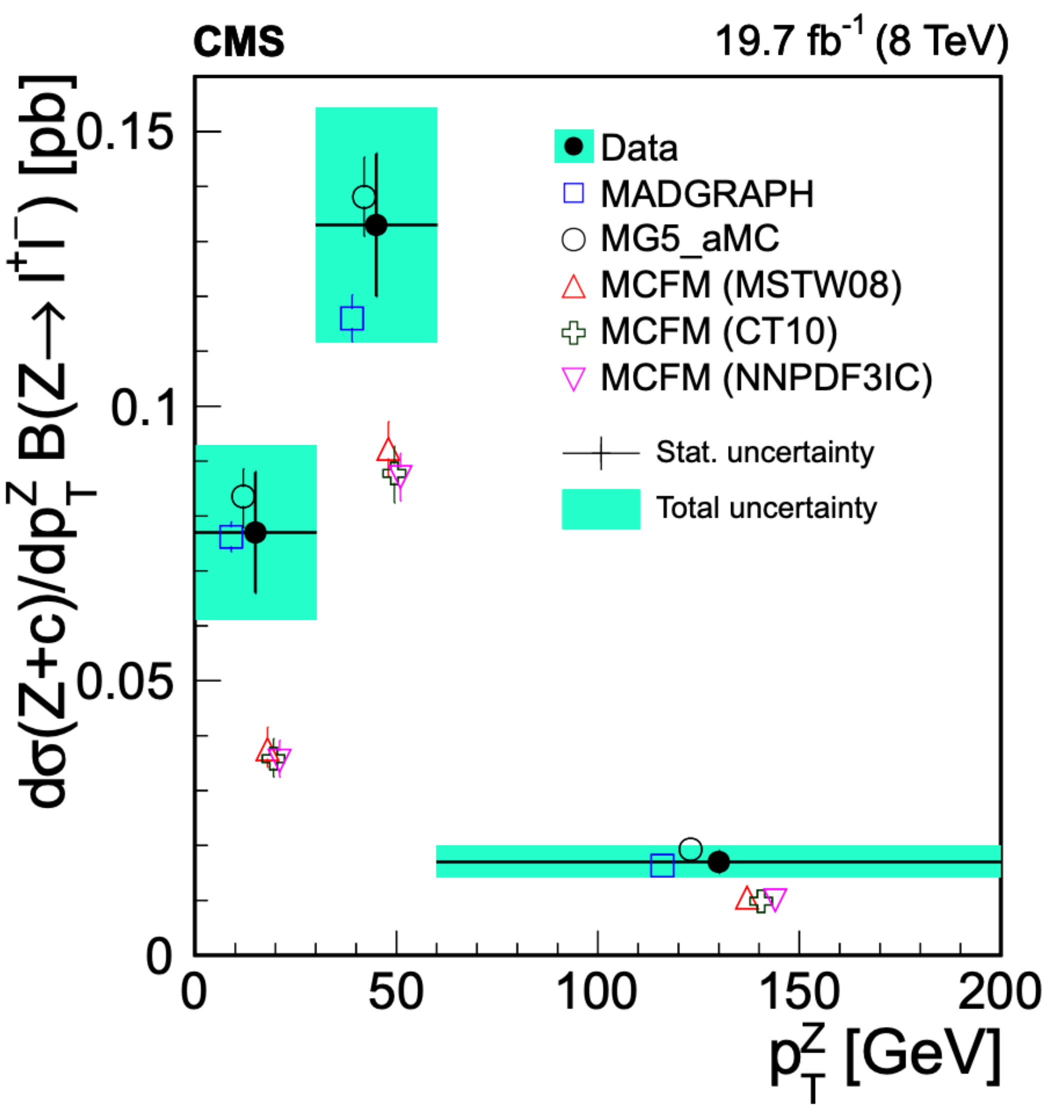
then one is probing the gluon and light quarks, which are already better constrained by other processes

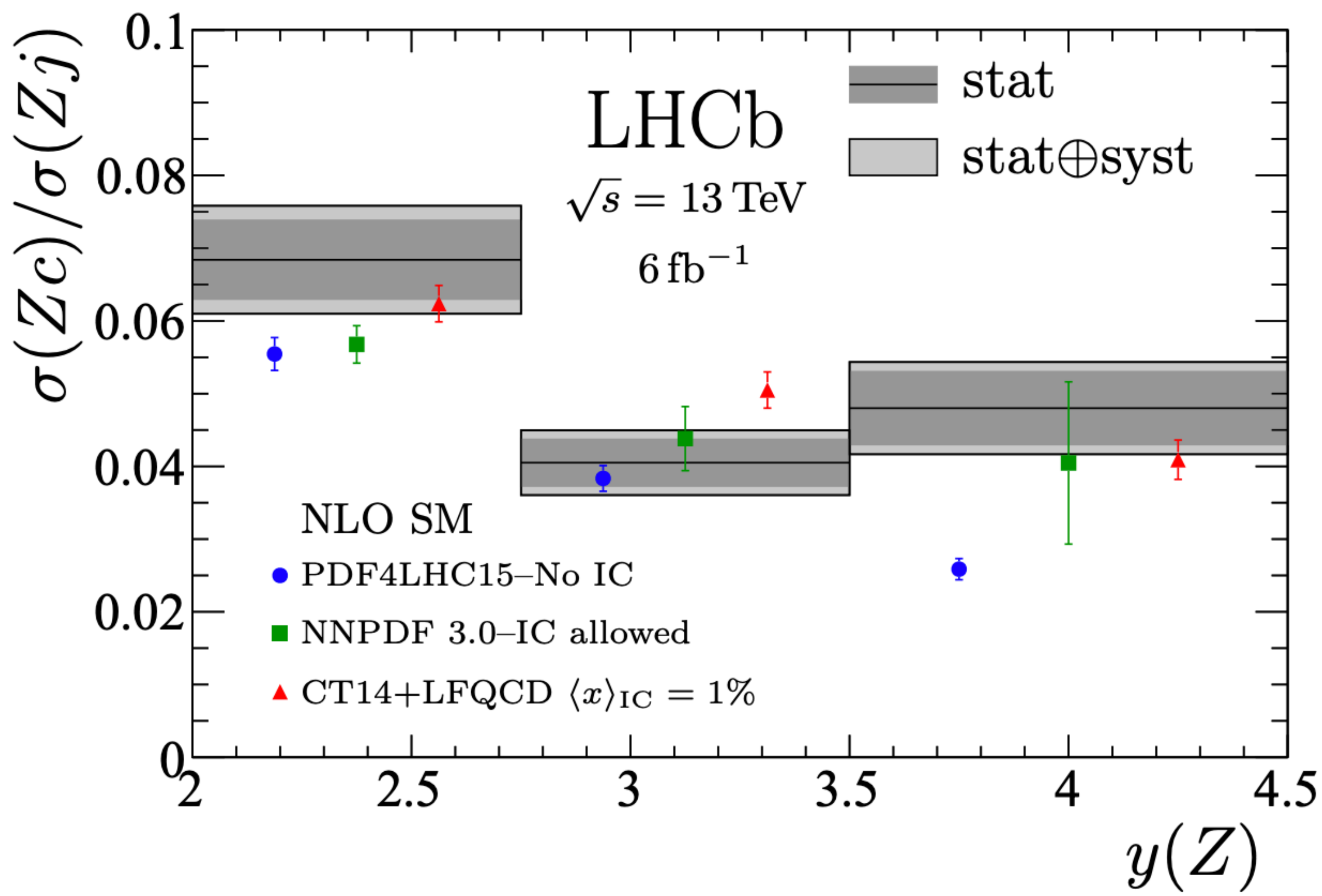
if the charm PDF receives contribution from an intrinsic component

then one is measuring whether or not the nucleon contains an intrinsic charm component!

Z+D @ LHC: direct probe of intrinsic charm

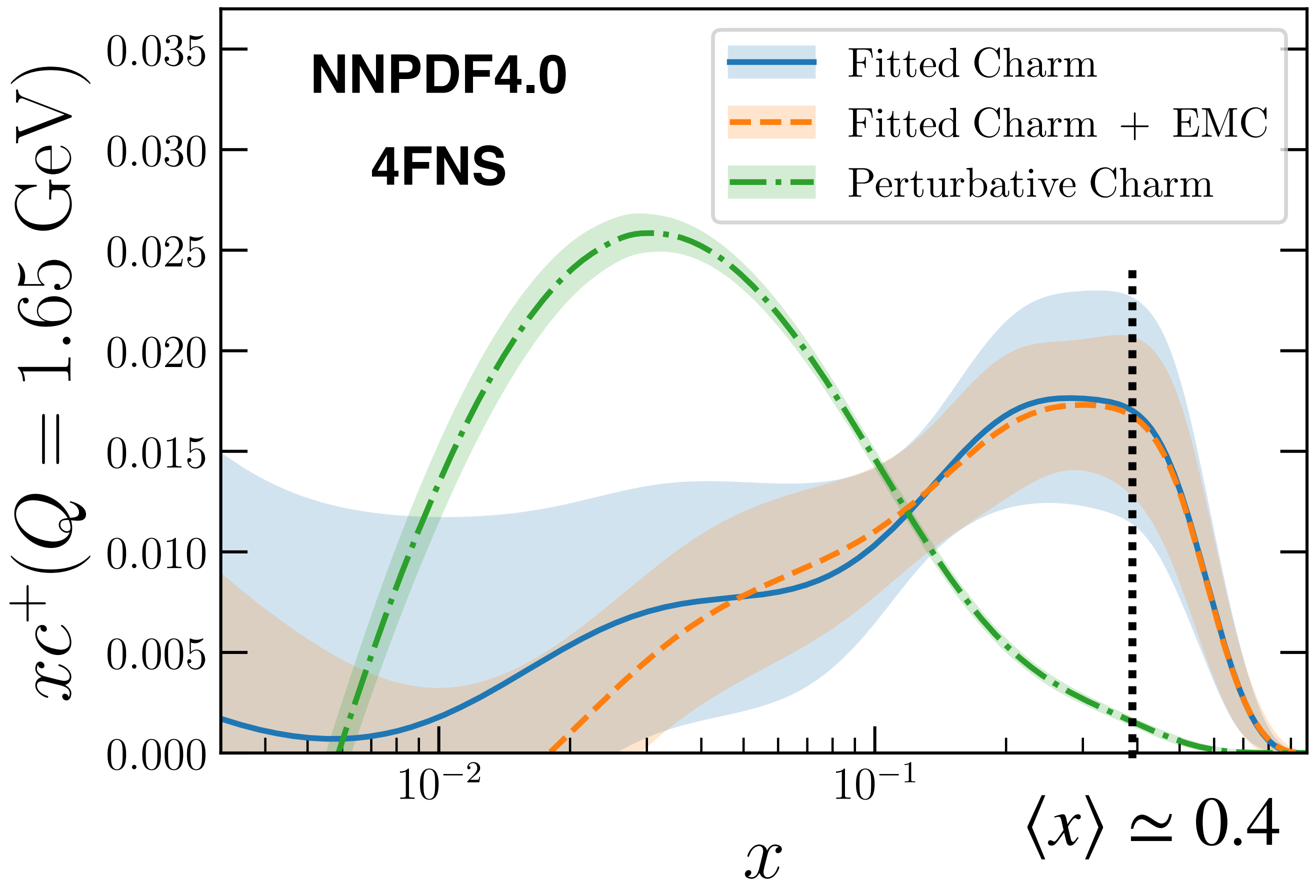
Z+charm @ LHC

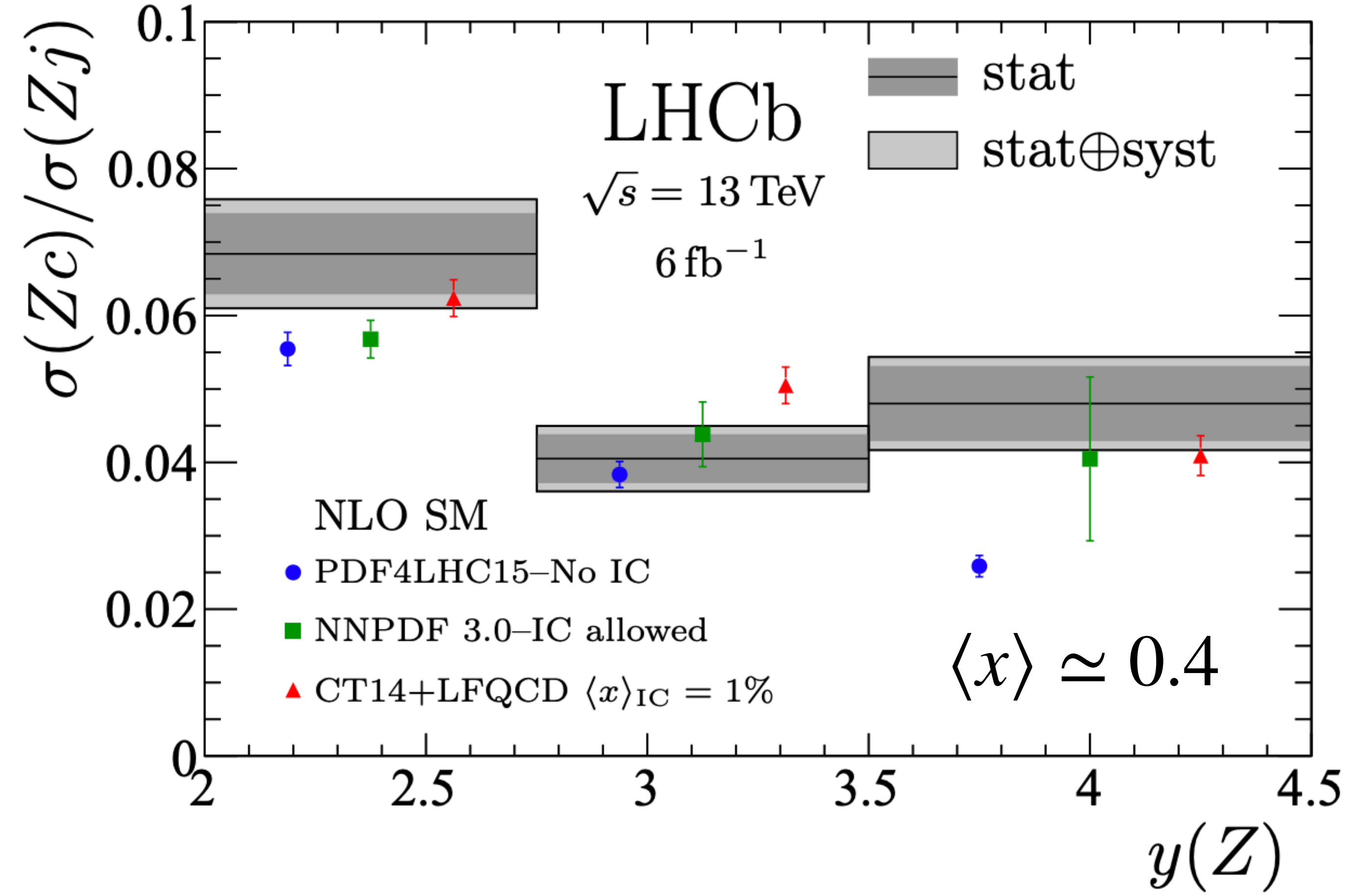




- ATLAS/CMS probe **medium-*x* charm**, where perturbative and fitted approaches yield similar results. Sensitivity to intrinsic charm requires reaching **higher values of p_T^Z**
- LHCb probes **forward kinematics**, where IC is enhanced (*<x>=0.4* in most forward bin)
- LHCb measurements **favour the intrinsic charm hypothesis**, though more data required

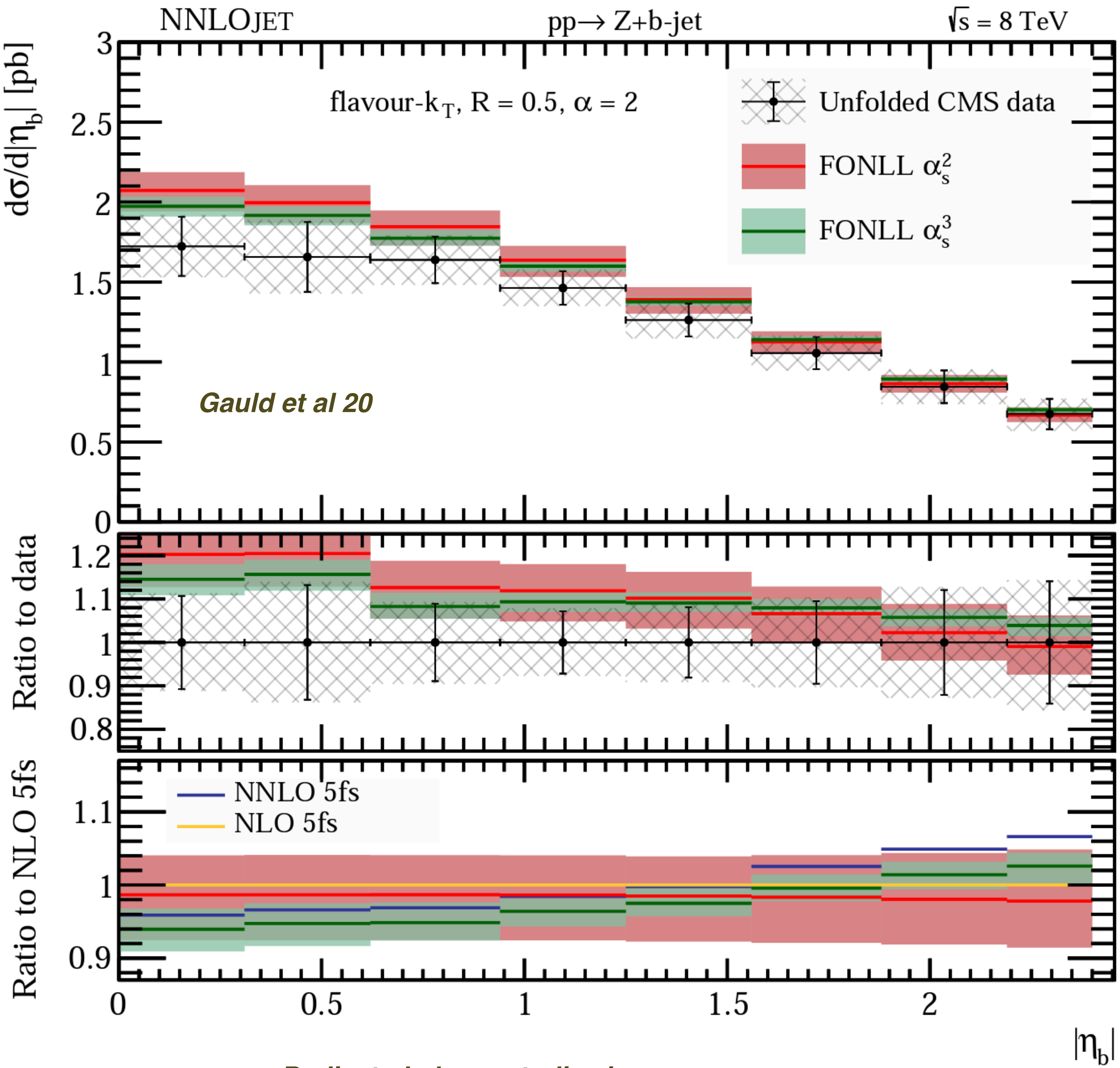
Z+charm @ LHC





- ATLAS/CMS probe **medium-*x* charm**, where perturbative and fitted approaches yield similar results. Sensitivity to intrinsic charm requires reaching **higher values of p_T^Z**
- LHCb probes **forward kinematics**, where IC is enhanced (*<x>=0.4* in most forward bin)
- LHCb measurements **favour the intrinsic charm hypothesis**, though more data required

Z+charm @ LHC



Dedicated pheno studies in progress

Theory subtleties in the interpretation of **V+Q data**

- Charm-jets definition require an **IR safe flavour algorithm**
- Modelling of heavy meson production involves charm/ bottom **fragmentation**
- Heavy quark mass effects may not be negligible, need **general-mass VFN scheme** e.g. FONLL
- **NNLO QCD corrections** required for precise calculations, though likely to partially cancel in ratios a la LHCb

Summary and outlook

- The associated production of Z bosons with charm quarks represents a direct probe of the **charm content of the proton**

- Recent progress in global PDF fits identifies an **intrinsic component** of the charm PDF with a local significance at the 3-sigma level

- Such intrinsic component is consistent with recent LHCb measurements of **forward Z+charm production**, though more theory and experimental work is required

``The simple hydrogen atom nucleus appears to be surprisingly charming''

Extra Material

Improved fitting methodology

- **Stochastic Gradient Descent** via TensorFlow for NN training
- Automated model **hyperparameter optimisation:** NN architecture, minimiser, learning rates …
- Validation with **future tests** (forecasting new datasets) and **closure tests** (data based on known PDFs)

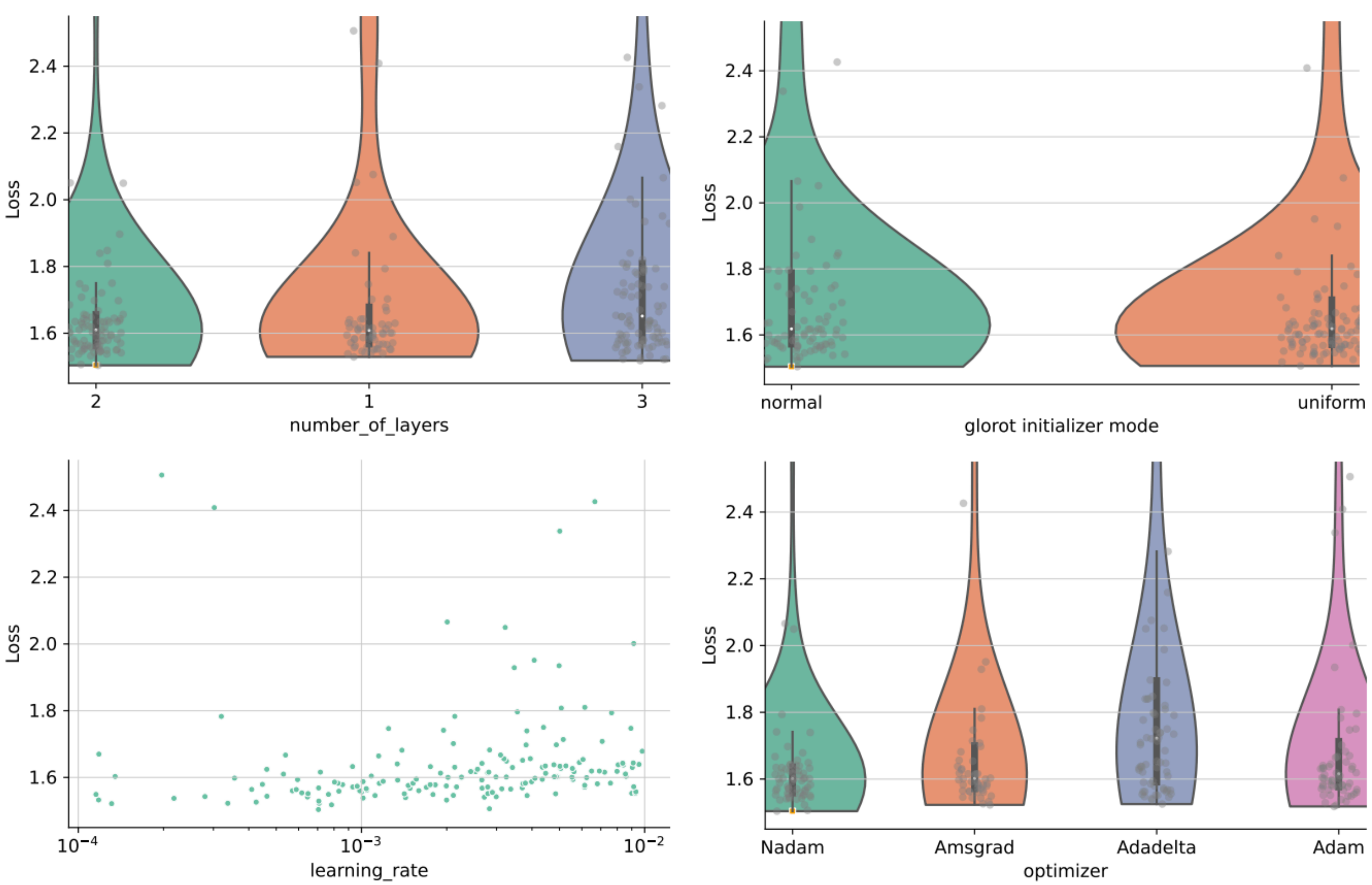
Loss (``average")

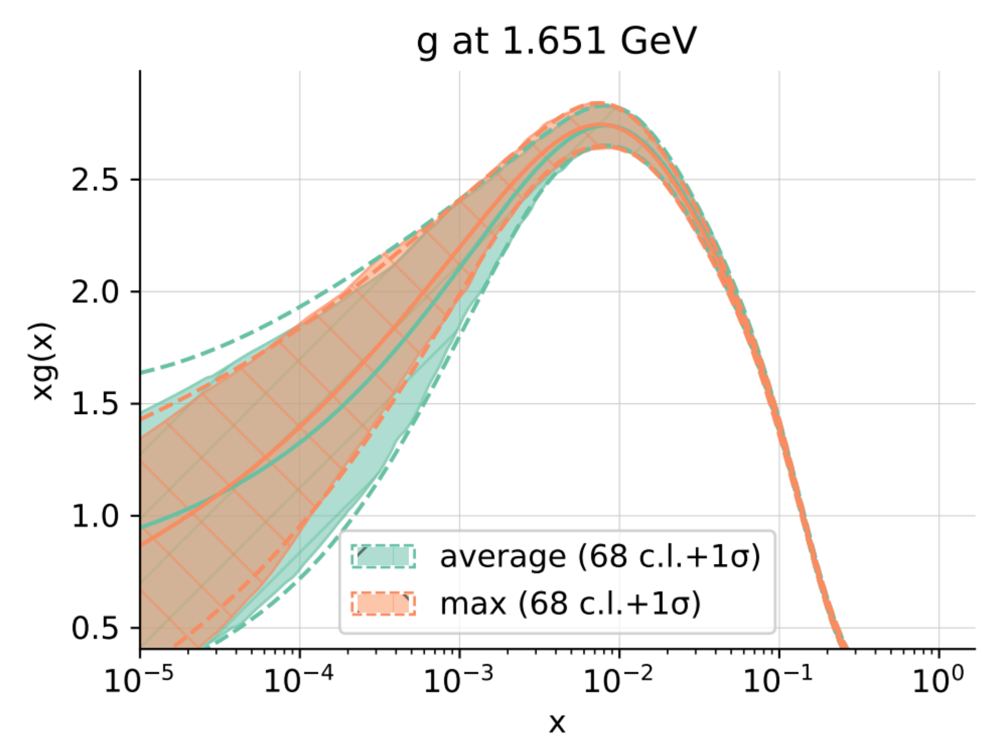
ML model hyperparams

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}\in\boldsymbol{\Theta}}{\arg\min}\left(\frac{1}{n_{\rm fold}}\sum_{k=1}^{n_{\rm fold}}\chi^2_k(\boldsymbol{\theta})\right)$$

Loss (``max")

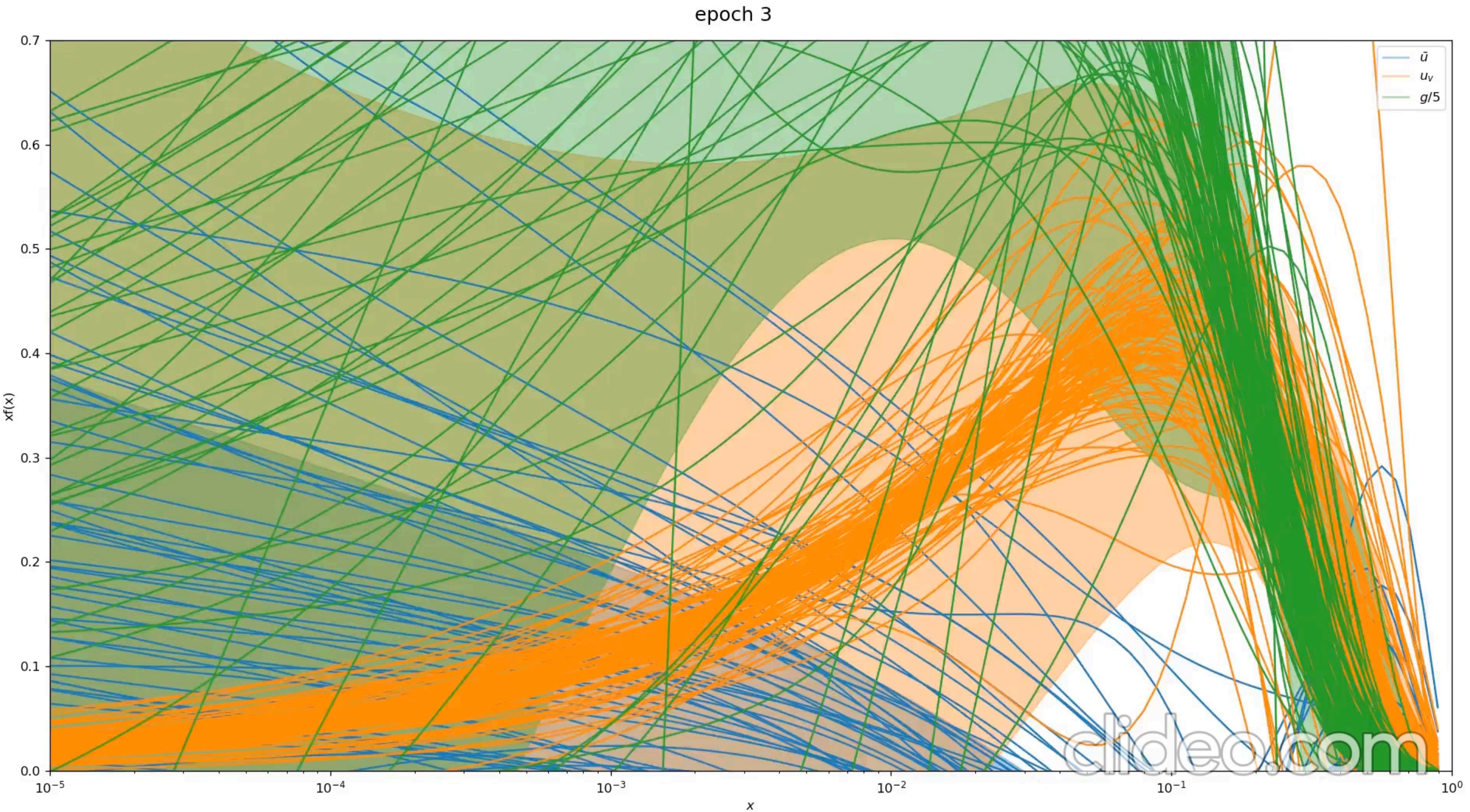
$$L = \max\left(\chi^2_1, \chi^2_2, \chi^2_3, \ldots, \chi^2_{n_{\rm fold}}\right)$$





Stability wrt hyperopt loss function

Improved fitting methodology



Illustrating the outcome of **SGD minimisation** (band: standard deviation over the MC replicas)

Closure and future tests

Closure tests

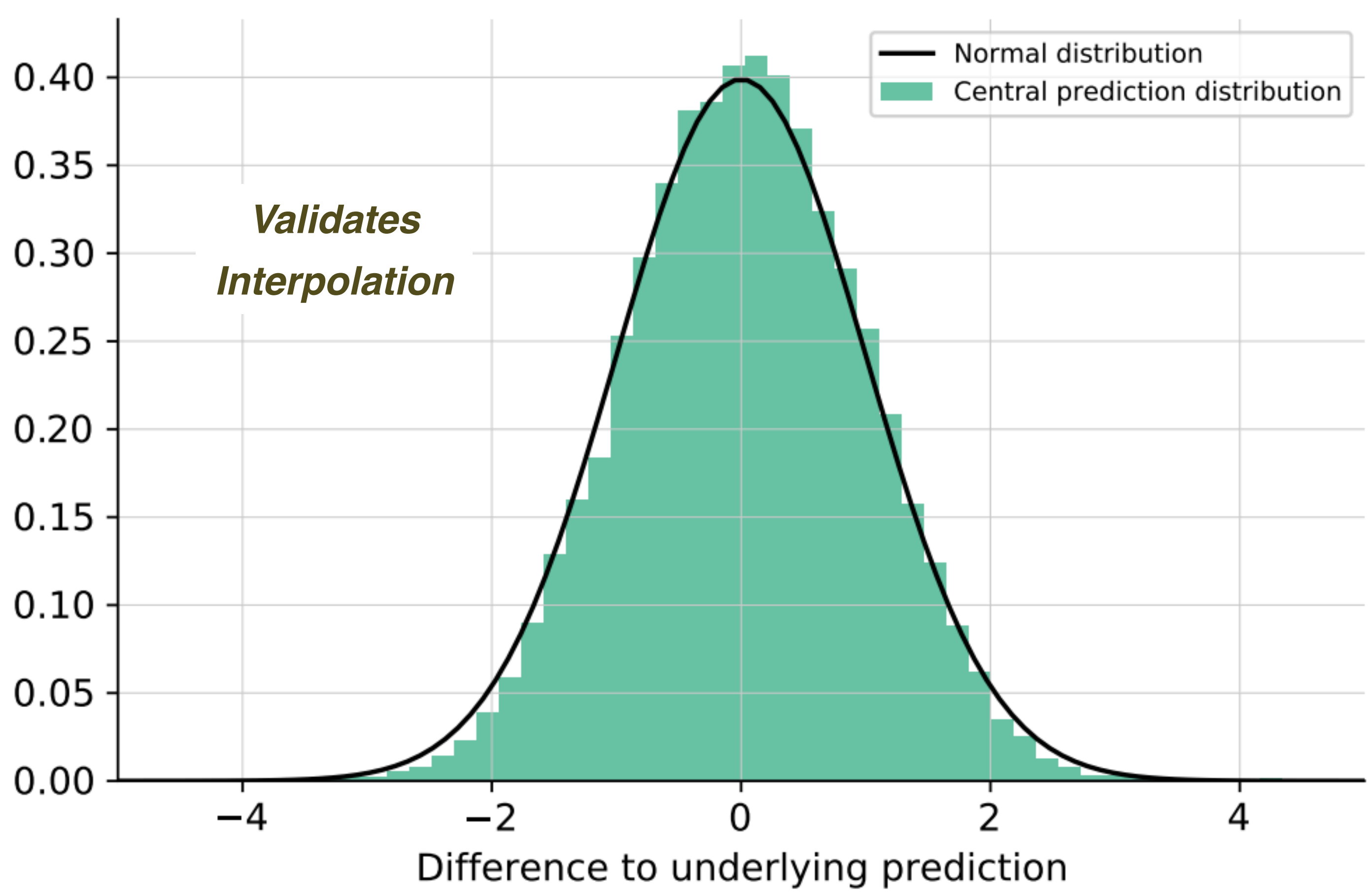
Generate **toy data** based on some known PDF, check *a posteriori* that the **true underlying law is reproduced** within errors

mean NN prediction *true central value*

generate many toys

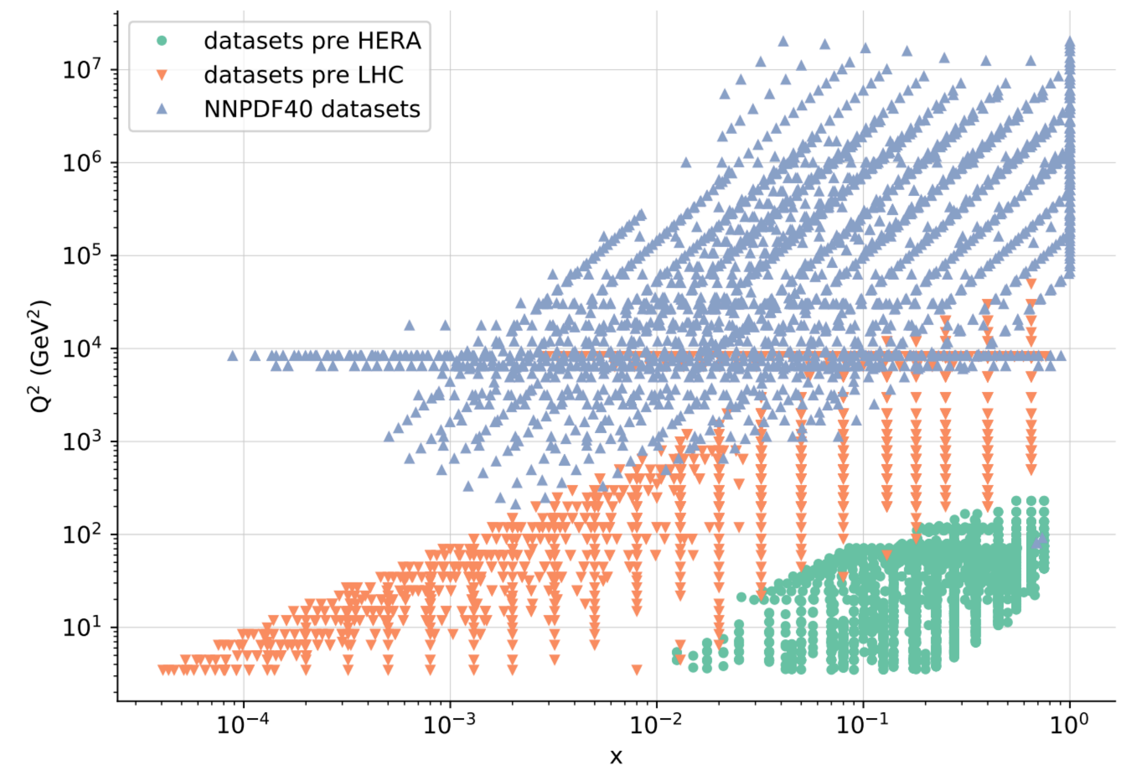
$$\delta_i^{(l)} \equiv \frac{\left(\mathbf{E}_\epsilon\left[g_i\right]^{(l)} - f_i\right)}{\sigma_i^{(l)}}$$

data index *PDF uncertainty*



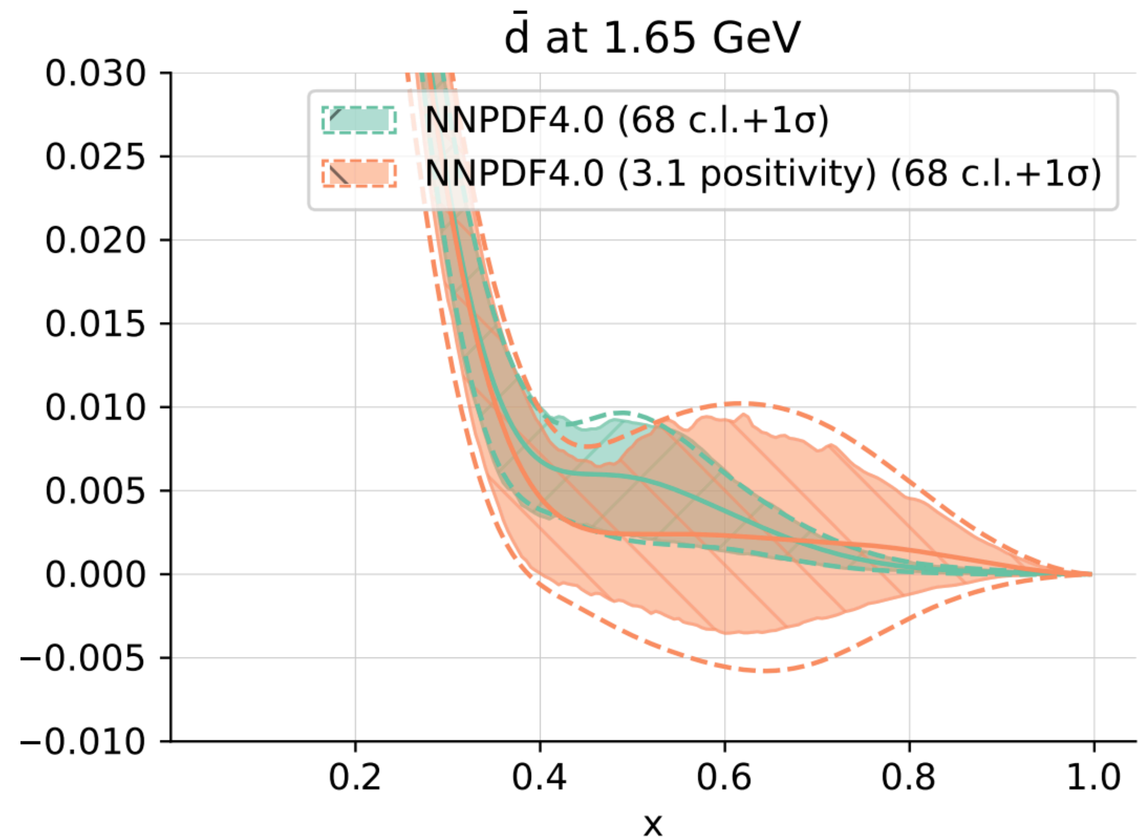
Future tests

Fit data restricted to specific kinematic regions, then verify **succesful extrapolation**



Process	χ^2 pre-HERA	χ^2 pre-LHC	χ^2 Global
Fixed target NC DIS	1.05	1.18	1.23
Fixed target CC DIS	0.80	0.85	0.87
Fixed target Drell-Yan	0.92	1.27	1.59
HERA	*27.20* (**1.23**)	1.22	1.20
Collider Drell-Yan (Tevatron)	*5.52* (**1.02**)	0.99	1.11
Collider Drell-Yan (LHC)	*18.91* (**1.31**)	*2.63* (**1.58**)	1.53
Top quark production	*20.01* (**1.06**)	*1.30* (**0.87**)	1.01
Jet production	*2.69* (**0.98**)	*2.12* (**1.10**)	1.26

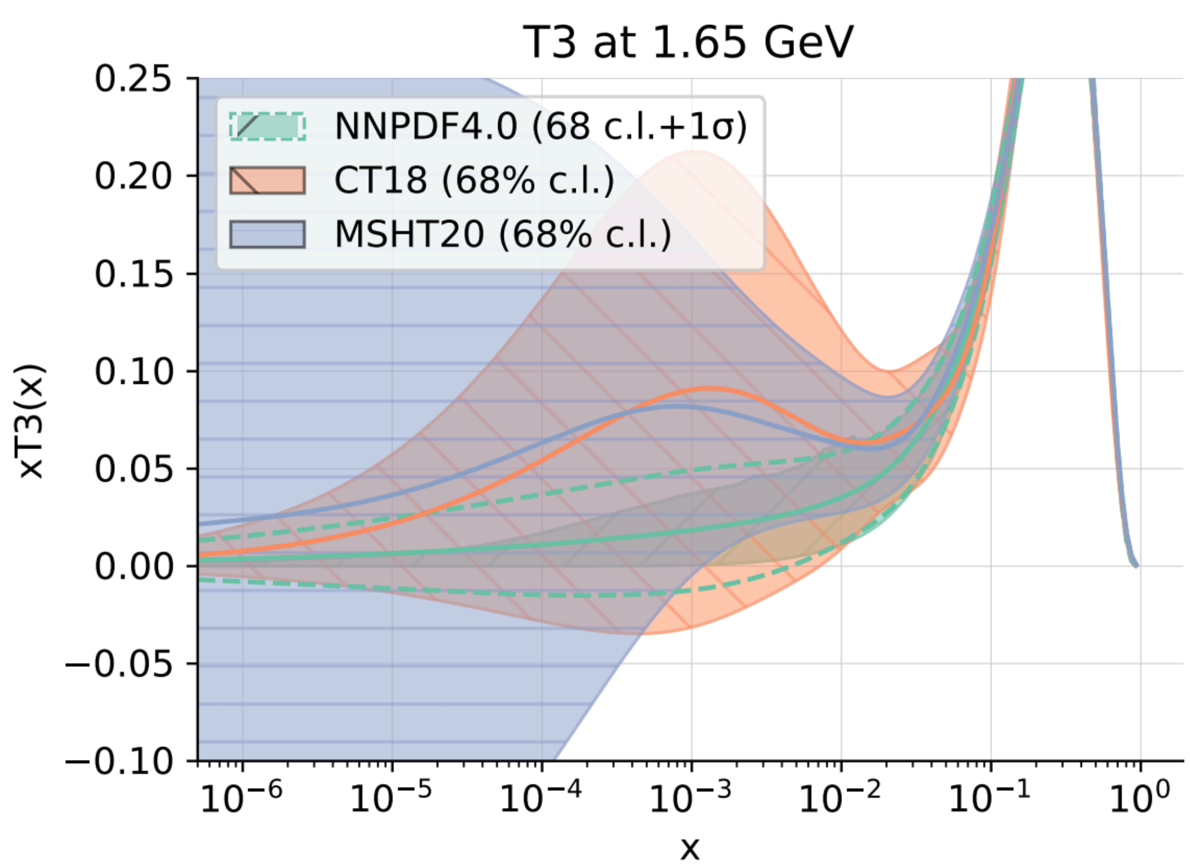
Positivity and integrability

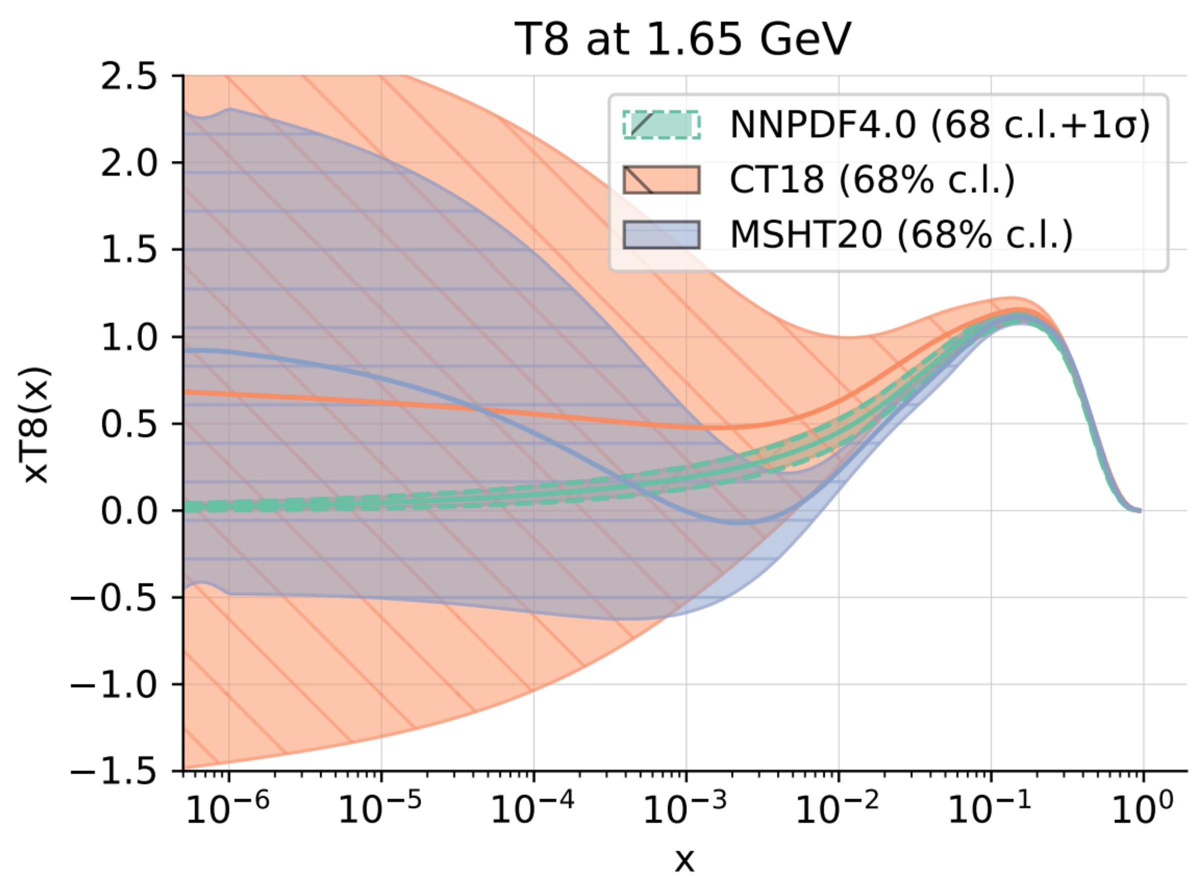


- MSbar PDFs have been shown to satisfy **positivity** requirements at all orders: **reduce large-*x* uncertainties**

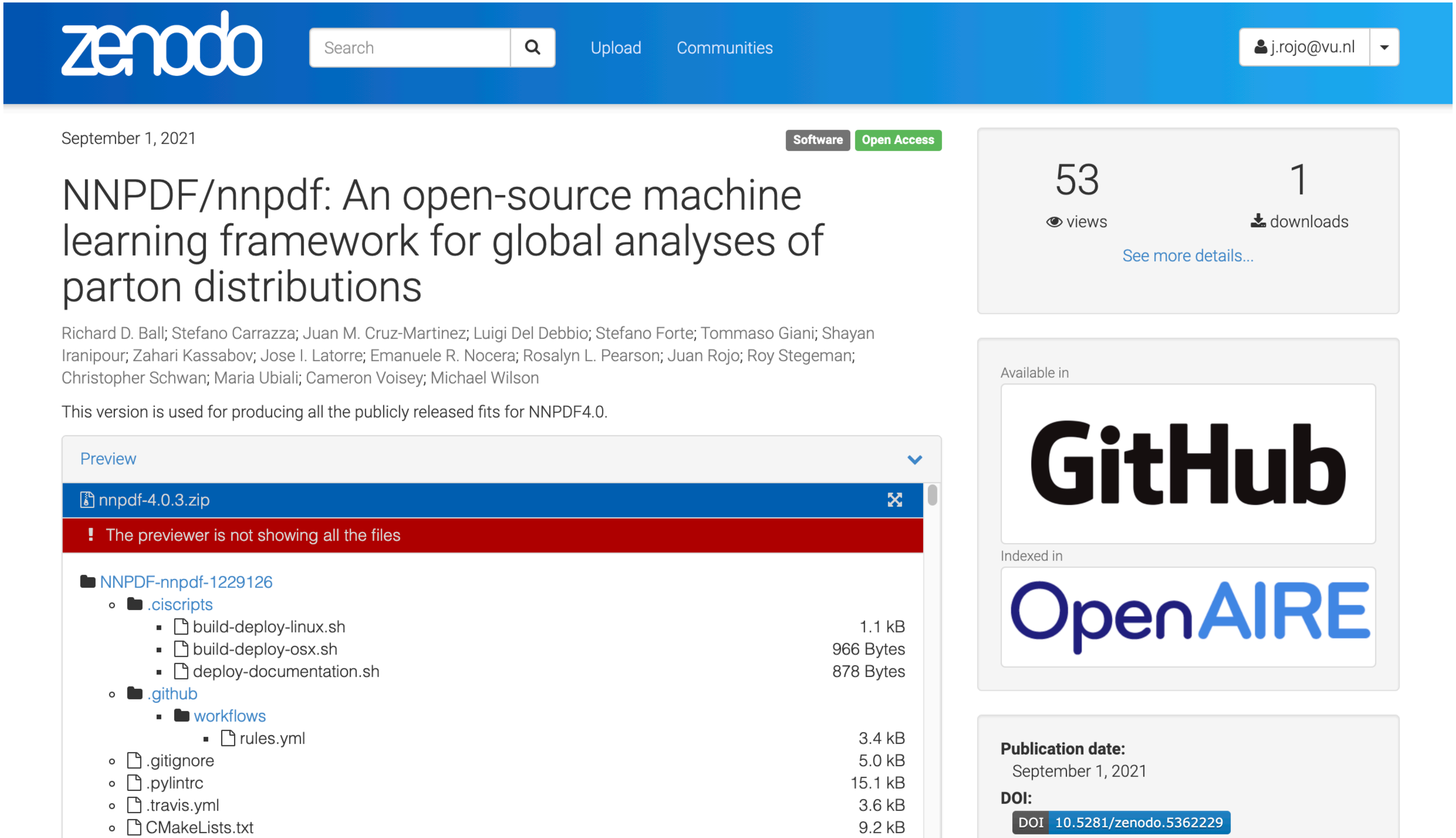
- The non-singlet quark triplet and octet should be *integrable* (e.g. Gottfried sum rule): **reduce small-*x* uncertainties**

$$T_8 = (u + \bar{u}) + \left(d + \bar{d}\right) - 2\,(s + \bar{s})$$



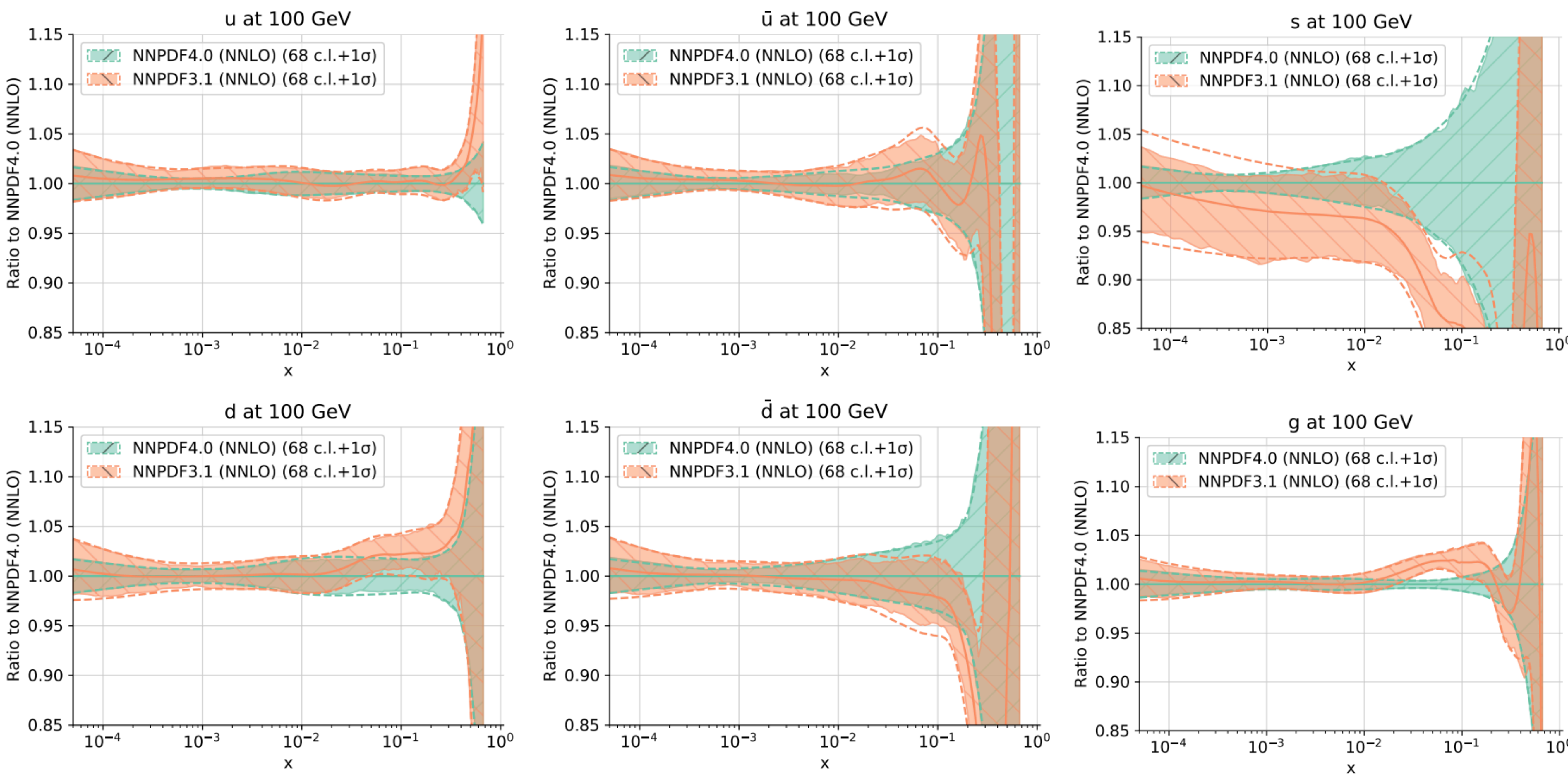


A ML open-source QCD fitting framework



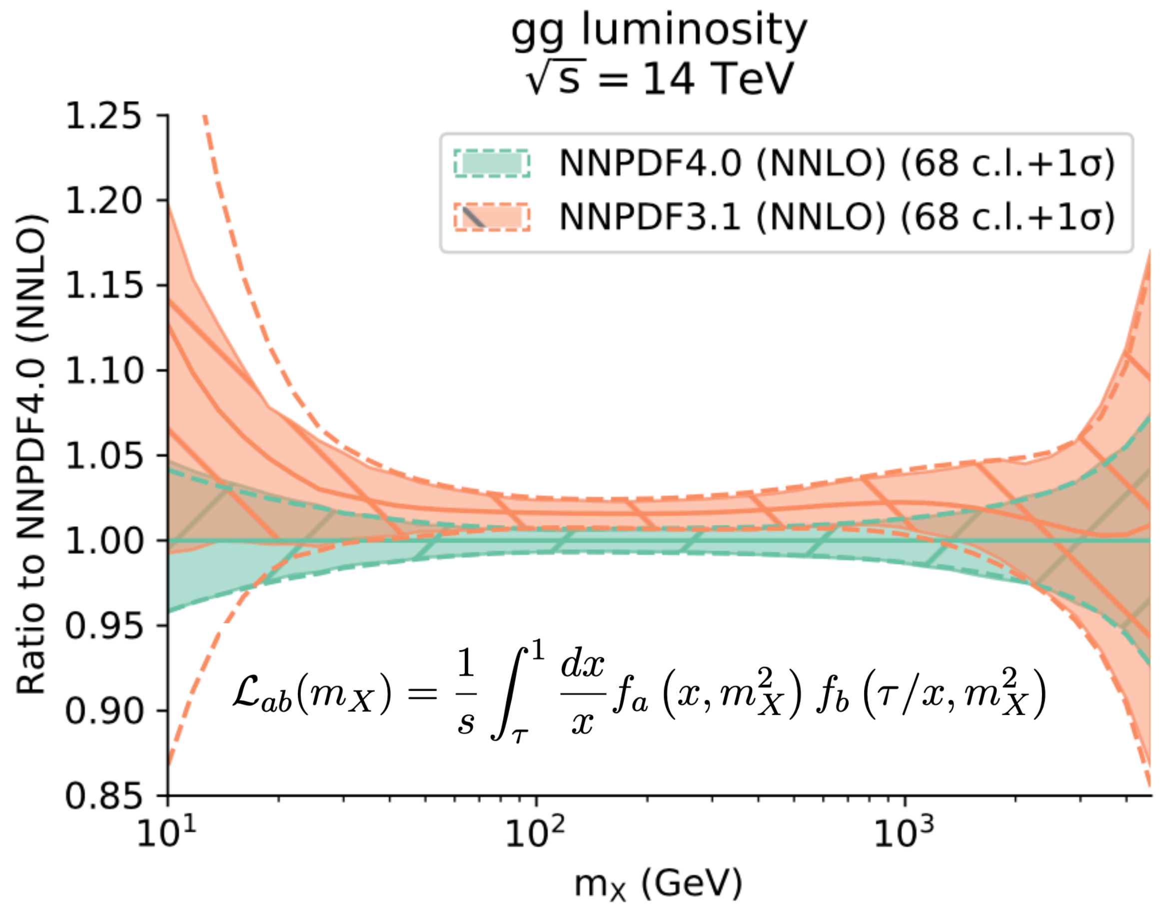
The full **NNPDF machine learning fitting framework** has been publicly released open source, together with extensive documentation and user-friendly examples

Comparison with NNPDF3.1

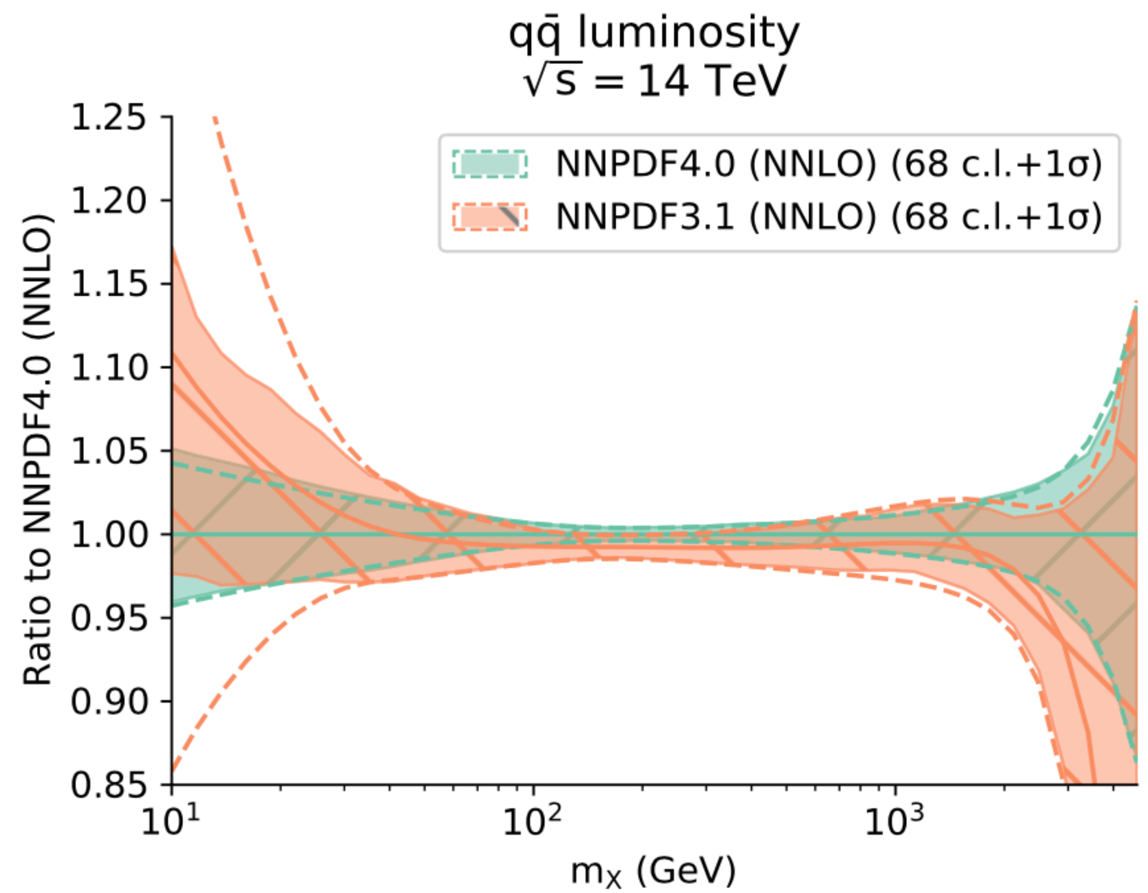


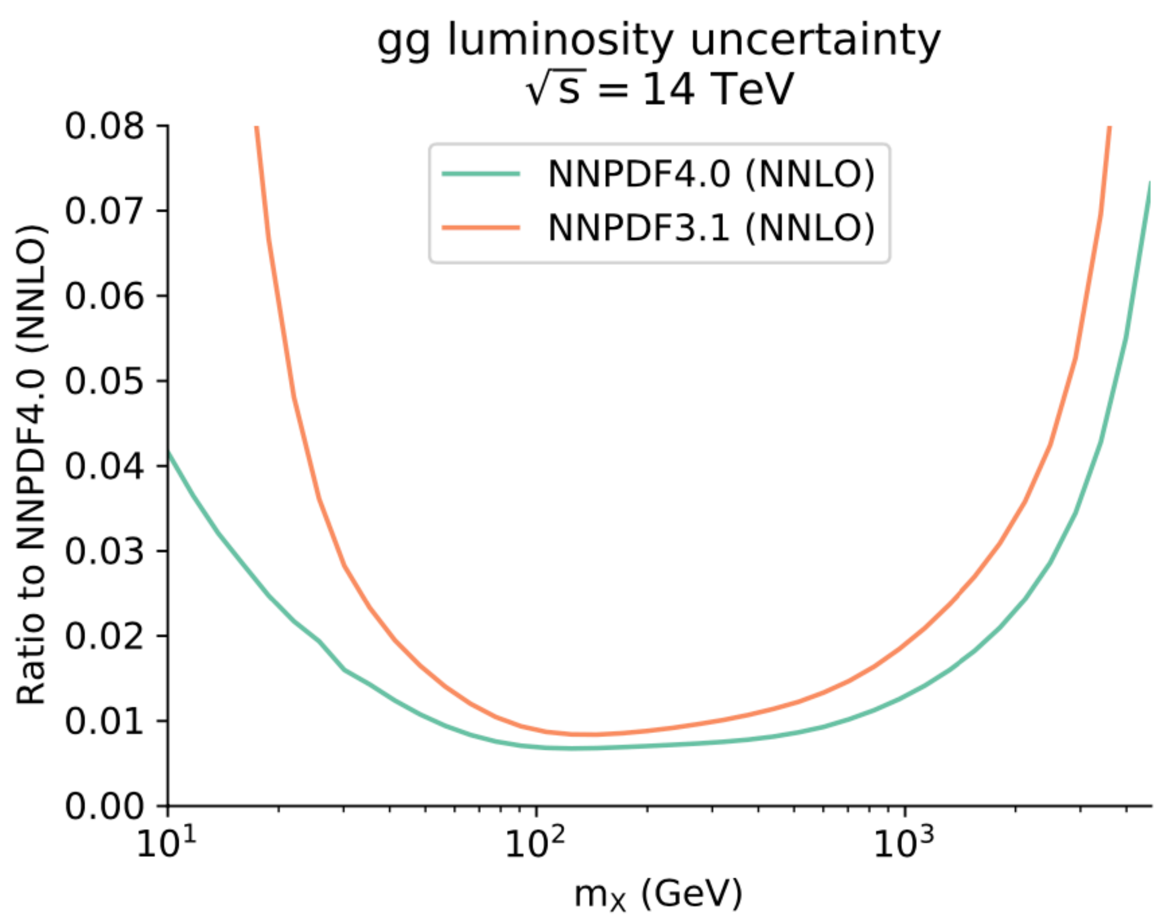
- **Good agreement** with NNPDF3.1 within uncertainties, with NNPDF4.0 being more precise
- Differences can be traced back to the **impact of specific datasets** (e.g. dijets for large-*x* gluon) or improvements in **theory calculations** (e.g. NNLO corrections in dimuon DIS for strangeness)

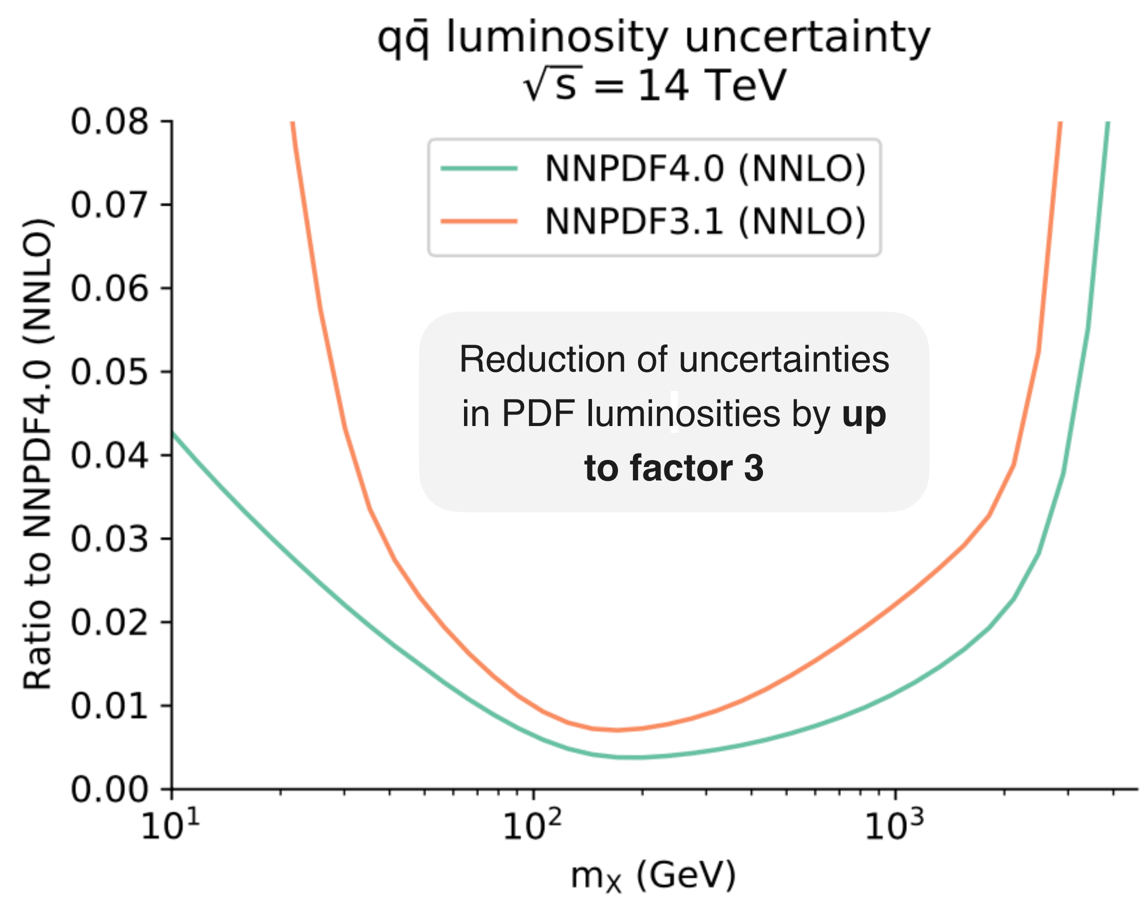
Comparison with NNPDF3.1



$$\mathcal{L}_{ab}(m_X) = \frac{1}{s}\int_\tau^1 \frac{dx}{x} f_a\left(x, m_X^2\right) f_b\left(\tau/x, m_X^2\right)$$

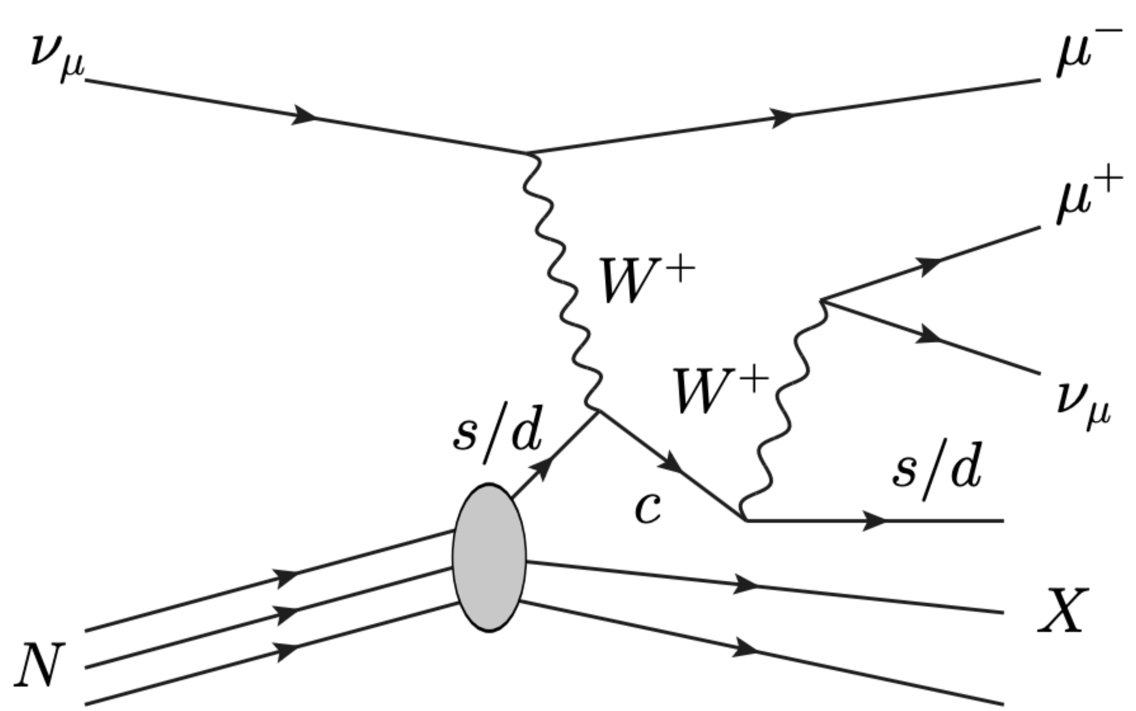






Reduction of uncertainties in PDF luminosities by **up to factor 3**

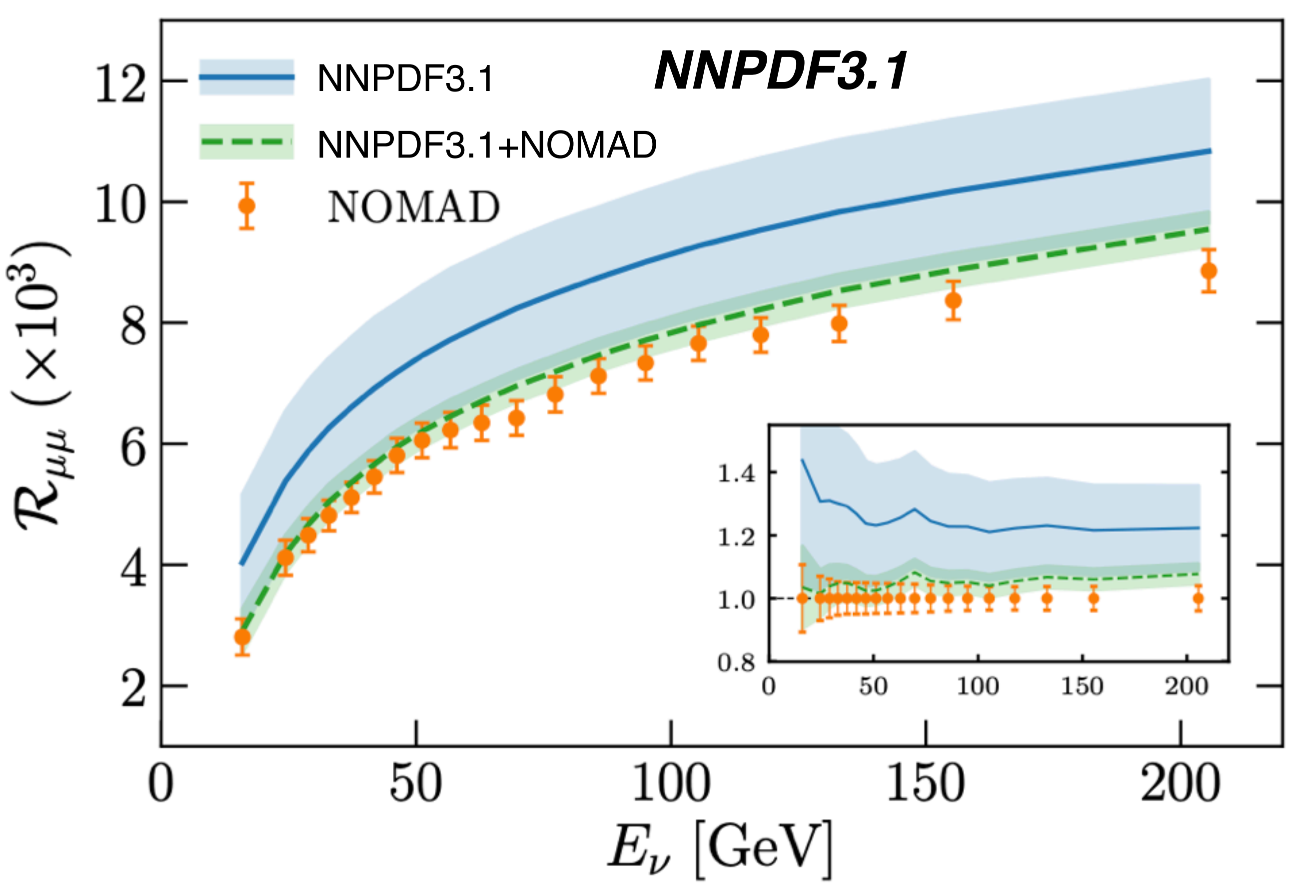
The strangest proton

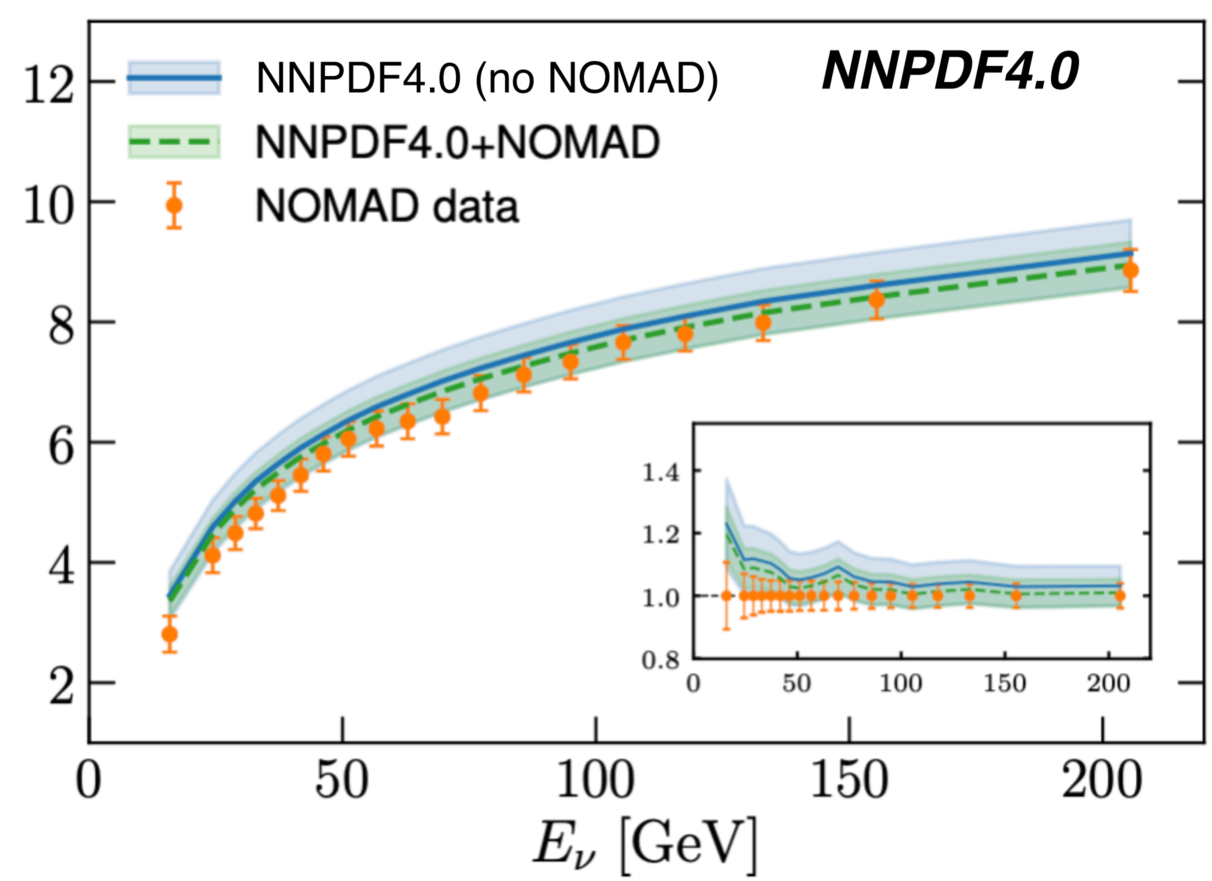


- NOMAD dimuon DIS data sensitive to **strangeness** via charged-current scattering

- Fitting NOMAD had large impact on the strangeness in NNPDF3.1, now in NNPDF4.0 the **no-NOMAD fit** is already spot on the data

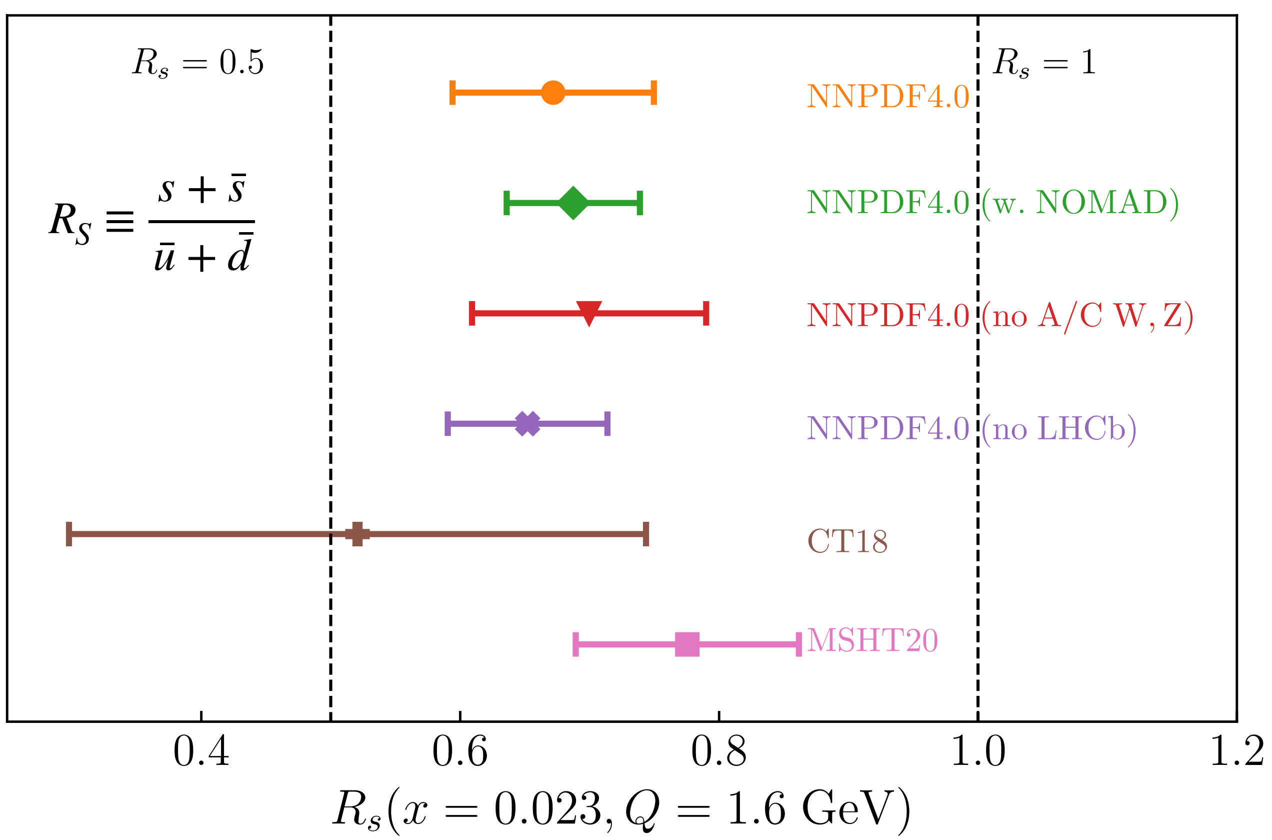
Excellent consistency of global dataset

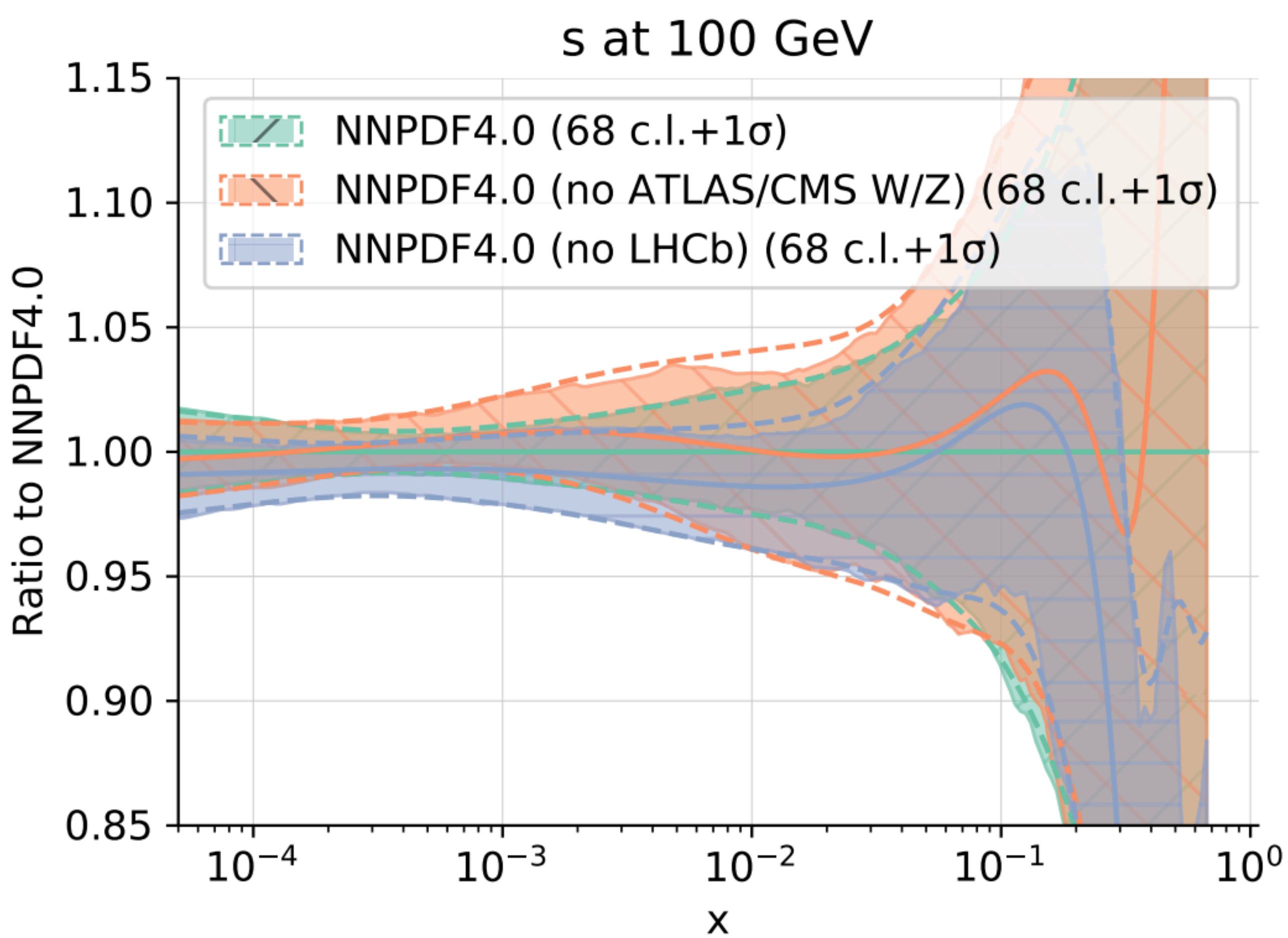


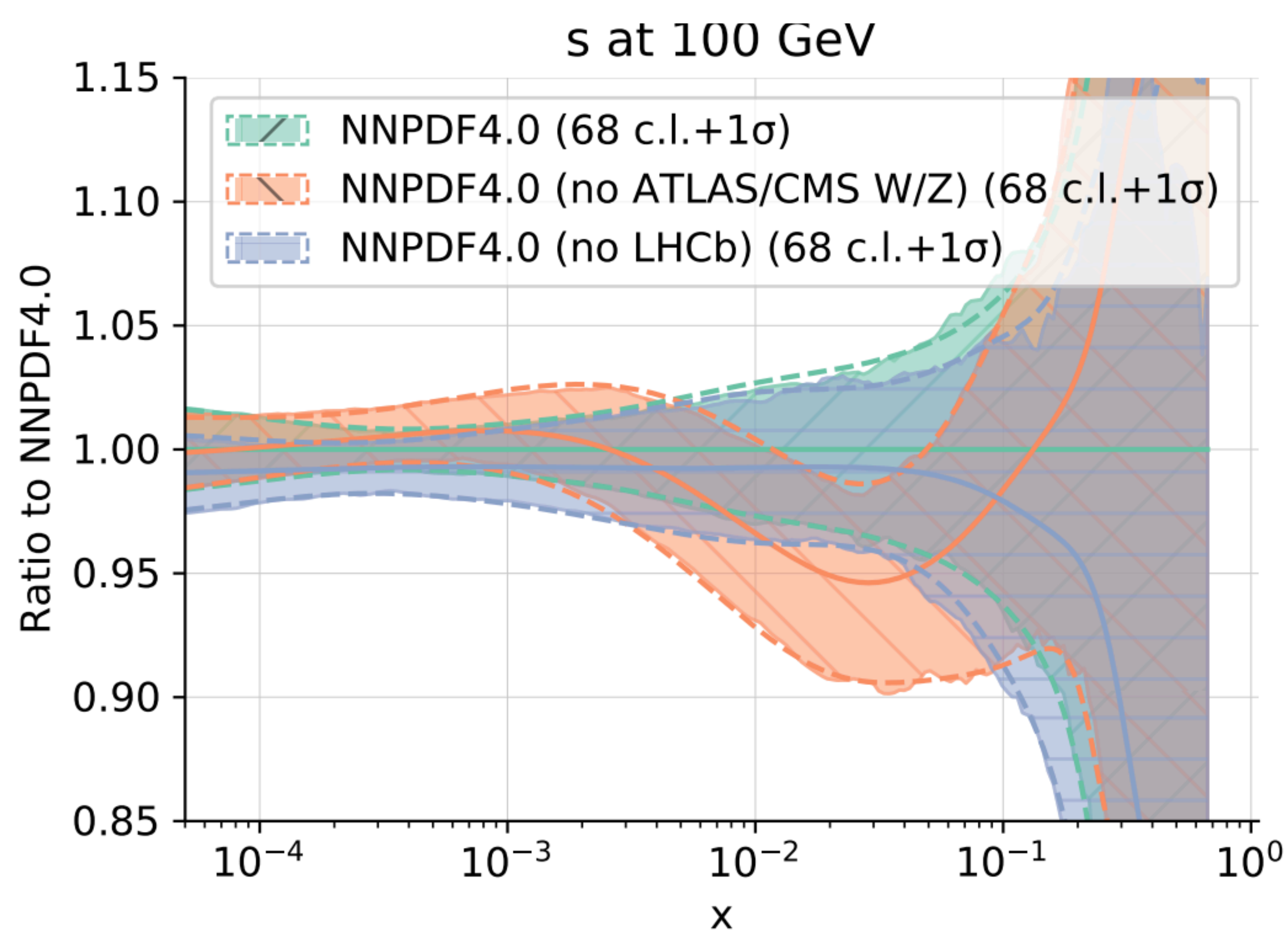


The strangest proton

- ☑ The LHC **inclusive *W, Z* production** data are also sensitive probes of the proton strangeness
- ☑ Fit results stable, within uncertainties, when either **ATLAS/CMS** or **LHCb *W, Z*** data are removed
- ☑ **No tension** between LHC and DIS neutrino data observed

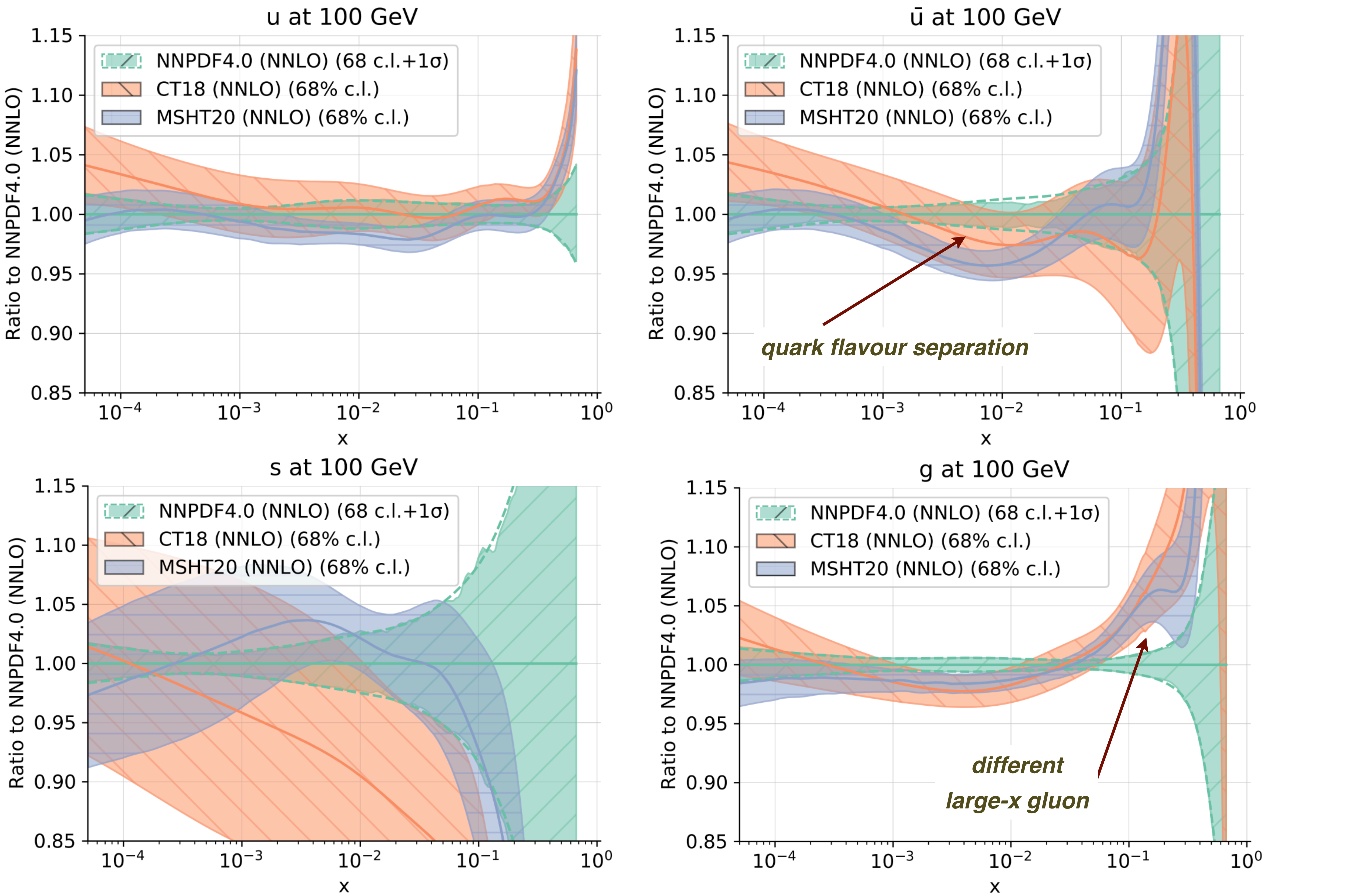






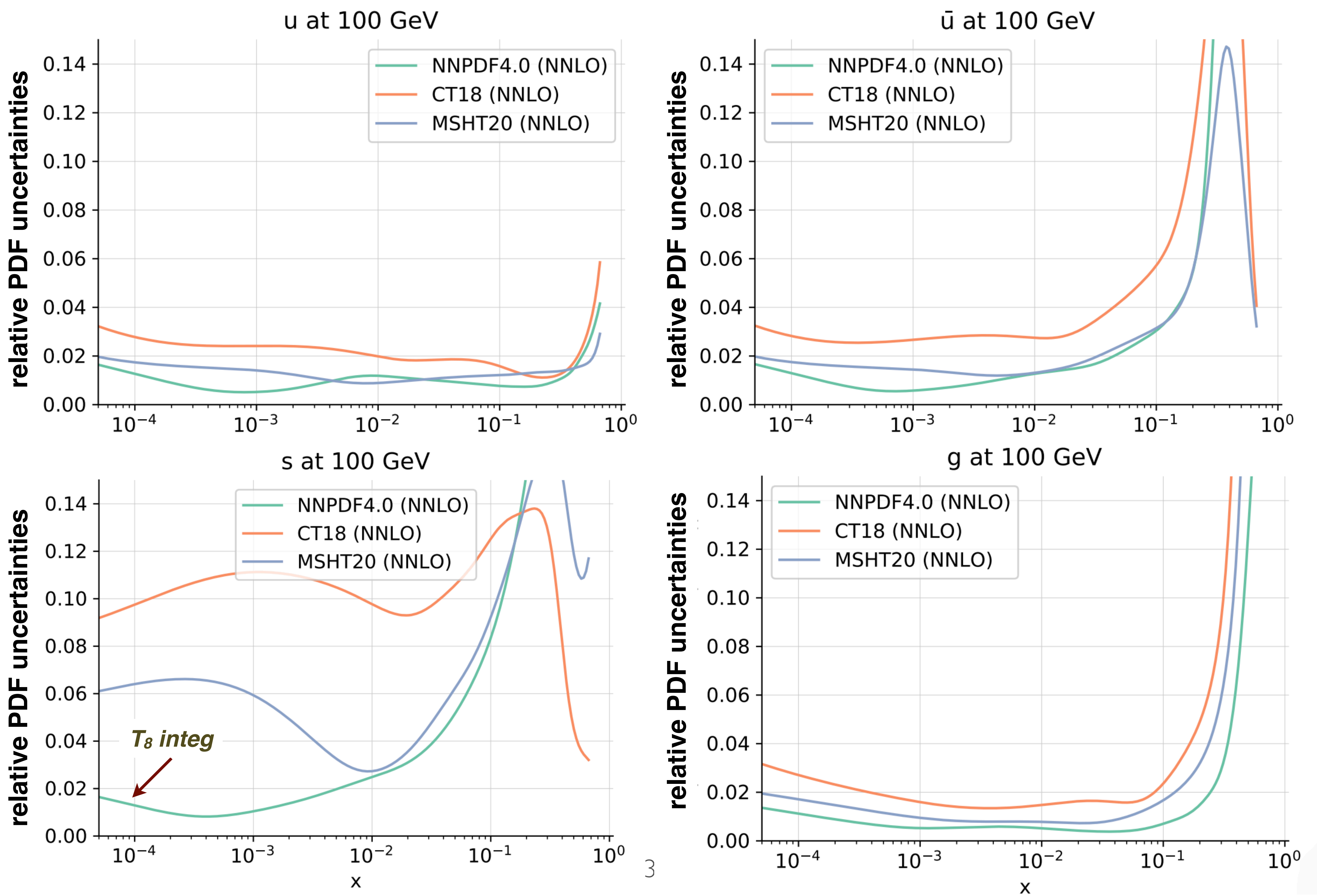
Comparison between global fits

reasonable agreement with CT18, and MSHT20, different pattern of PDF uncertainties



Comparison between global fits

different **pattern of PDF uncertainties** … $\delta_{\mathrm{PDF}}(\mathrm{CT}) \gtrsim \delta_{\mathrm{PDF}}(\mathrm{MSHT}) \gtrsim \delta_{\mathrm{PDF}}(\mathrm{NNPDF})$



Comparison between global fits

… follows pattern of **input datasets**

$$\delta_{\text{PDF}}(\text{CT}) \gtrsim \delta_{\text{PDF}}(\text{MSHT}) \gtrsim \delta_{\text{PDF}}(\text{NNPDF})$$

Data set	Ref.	NNPDF3.1	NNPDF4.0	ABMP16	CT18	MSHT20
ATLAS W,Z 7 TeV ($\mathcal{L}=35$ pb^{-1})	[51]	✓	✓	✓	✓	✓
ATLAS W,Z 7 TeV ($\mathcal{L}=4.6$ fb^{-1})	[52]	✓	✓	✗	(✓)	✓
ATLAS low-mass DY 7 TeV	[53]	✓	✓	✗	(✓)	✗
ATLAS high-mass DY 7 TeV	[54]	✓	✓	✗	(✓)	✓
ATLAS W 8 TeV	[79]	✗	(✓)	✗	✗	✓
ATLAS DY 2D 8 TeV	[78]	✗	✓	✗	✗	✓
ATLAS high-mass DY 2D 8 TeV	[77]	✗	✓	✗	(✓)	✓
ATLAS $\sigma_{W,Z}$ 13 TeV	[81]	✗	✓	✓	✗	✗
ATLAS W+jet 8 TeV	[93]	✗	✓	✗	✗	✓
ATLAS Z p_T 7 TeV	[259]	(✓)	✗	✗	(✓)	✗
ATLAS Z p_T 8 TeV	[63]	✓	✓	✗	✓	✓
ATLAS $W+c$ 7 TeV	[83]	✗	✓	✗	(✓)	✗
ATLAS σ_{tt}^{tot} 7, 8 TeV	[65]	✓	✓	✓	✗	✗
ATLAS σ_{tt}^{tot} 7, 8 TeV	[260–265]	✗	✗	✓	✗	✗
ATLAS σ_{tt}^{tot} 13 TeV ($\mathcal{L}=3.2$ fb^{-1})	[66]	✓	✗	✓	✗	✗
ATLAS σ_{tt}^{tot} 13 TeV ($\mathcal{L}=139$ fb^{-1})	[134]	✗	✓	✗	✗	✗
ATLAS σ_{tt}^{tot} and Z ratios	[266]	✗	✗	✗	✗	(✓)
ATLAS $t\bar{t}$ lepton+jets 8 TeV	[67]	✓	✓	✗	✓	✓
ATLAS $t\bar{t}$ dilepton 8 TeV	[89]	✗	✓	✗	✗	✓
ATLAS single-inclusive jets 7 TeV, R=0.6	[73]	✓	(✓)	✗	✓	✓
ATLAS single-inclusive jets 8 TeV, R=0.6	[86]	✗	✓	✗	✗	✗
ATLAS dijets 7 TeV, R=0.6	[148]	✗	✓	✗	✗	✗
ATLAS direct photon production 8 TeV	[100]	✗	(✓)	✗	✗	✗
ATLAS direct photon production 13 TeV	[101]	✗	✓	✗	✗	✗
ATLAS single top R_t 7, 8, 13 TeV	[94, 96, 98]	✗	✓	✓	✗	✗
ATLAS single top diff. 7 TeV	[94]	✗	✓	✗	✗	✗
ATLAS single top diff. 8 TeV	[96]	✗	✓	✗	✗	✗

Data set	Ref.	NNPDF3.1	NNPDF4.0	ABMP16	CT18	MSHT20
CMS W asym. 7 TeV ($\mathcal{L}=36$ pb^{-1})	[267]	✗	✗	✗	✗	✓
CMS Z 7 TeV ($\mathcal{L}=36$ pb^{-1})	[268]	✗	✗	✗	✗	✓
CMS W electron asymmetry 7 TeV	[55]	✓	✓	✗	✓	✓
CMS W muon asymmetry 7 TeV	[56]	✓	✓	✓	✓	✗
CMS Drell-Yan 2D 7 TeV	[57]	✓	✓	✗	(✓)	✓
CMS Drell-Yan 2D 8 TeV	[269]	(✓)	✗	✗	✗	✗
CMS W rapidity 8 TeV	[58]	✓	✓	✓	✓	✓
CMS W,Z p_T 8 TeV ($\mathcal{L}=18.4$ fb^{-1})	[270]	✗	✗	✗	(✓)	✗
CMS Z p_T 8 TeV	[64]	✓	✓	✗	(✓)	✗
CMS $W+c$ 7 TeV	[76]	✓	✓	✗	(✓)	✓
CMS $W+c$ 13 TeV	[84]	✗	✓	✗	✗	(✓)
CMS single-inclusive jets 2.76 TeV	[75]	✓	✗	✗	✗	✓
CMS single-inclusive jets 7 TeV	[147]	✓	(✓)	✗	✓	✓
CMS dijets 7 TeV	[74]	✗	✓	✗	✗	✗
CMS single-inclusive jets 8 TeV	[87]	✗	✓	✗	✓	✓
CMS 3D dijets 8 TeV	[149]	✗	(✓)	✗	✗	✗
CMS σ_{tt}^{tot} 5 TeV	[88]	✗	✓	✗	✗	✗
CMS σ_{tt}^{tot} 7, 8 TeV	[146]	✓	✓	✗	✗	✗
CMS σ_{tt}^{tot} 8 TeV	[271]	✗	✗	✗	✗	✓
CMS σ_{tt}^{tot} 5, 7, 8, 13 TeV	[68, 272–280]	✗	✗	✓	✗	✗
CMS σ_{tt}^{tot} 13 TeV	[69]	✓	✓	✓	✗	✗
CMS $t\bar{t}$ lepton+jets 8 TeV	[70]	✓	✓	✗	✗	✓
CMS $t\bar{t}$ 2D dilepton 8 TeV	[90]	✗	✓	✗	✓	✓
CMS $t\bar{t}$ lepton+jet 13 TeV	[91]	✗	✓	✗	✗	✗
CMS $t\bar{t}$ dilepton 13 TeV	[92]	✗	✓	✗	✗	✗
CMS single top $\sigma_t+\sigma_{\bar{t}}$ 7 TeV	[95]	✗	✓	✓	✗	✗
CMS single top R_t 8, 13 TeV	[97, 99]	✗	✓	✓	✗	✗
CMS single top 13 TeV	[281, 282]	✗	✗	✗	✗	(✓)

Data set	Ref.	NNPDF3.1	NNPDF4.0	ABMP16	CT18	MSHT20
LHCb Z 7 TeV ($\mathcal{L}=940$ pb^{-1})	[59]	✓	✓	✗	✗	✓
LHCb $Z\to ee$ 8 TeV ($\mathcal{L}=2$ fb^{-1})	[61]	✓	✓	✓	✓	✓
LHCb W 7 TeV ($\mathcal{L}=37$ pb^{-1})	[283]	✗	✗	✗	✗	✓
LHCb $W,Z\to\mu$ 7 TeV	[60]	✓	✓	✓	✓	✓
LHCb $W,Z\to\mu$ 8 TeV	[62]	✓	✓	✓	✓	✓
LHCb $W\to e$ 8 TeV	[80]	✗	(✓)	✗	✗	✗
LHCb $Z\to\mu\mu, ee$ 13 TeV	[82]	✗	✓	✗	✗	✗

✓ in baseline dataset

✗ not considered

(✓) impact assessed but excluded from baseline