Closure Tests and Uncertainties in Global PDF fits

Lucian Harland-Lang, University College London

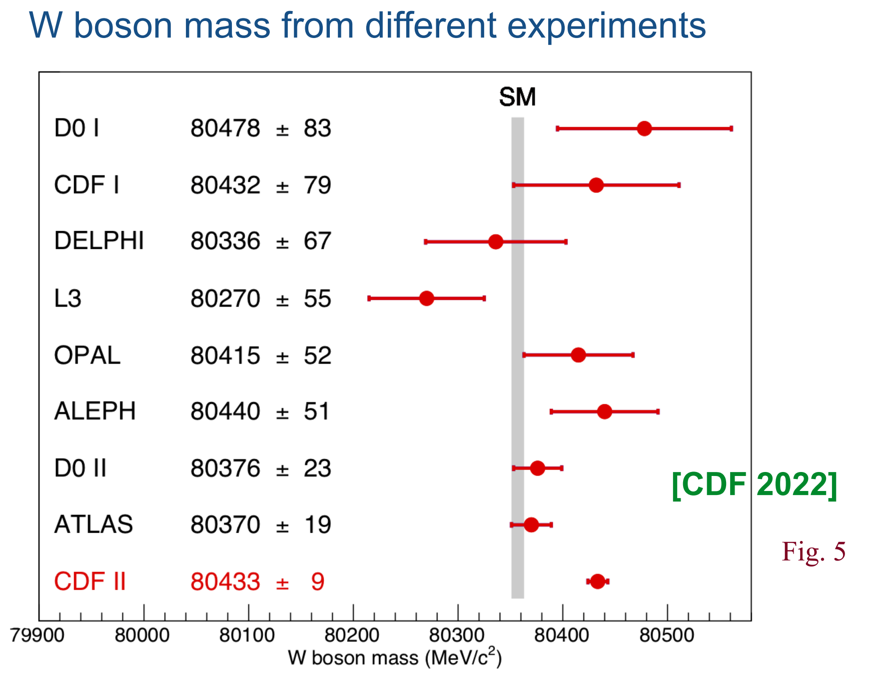
CTEQ Spring Meeting 2024, June 5

With Robert Thorne and Tom Cridge

Introduction

- Parton distribution functions (PDFs): a key ingredient in hadron collider physics.
- Knowledge of **PDFs** and their **uncertainties** a limiting factor in LHC precision and BSM searches.

W mass



[CDF 2022]

Fig. 5

M expectation: $M_W = 80{,}357 \pm 4_{\rm inputs} \pm 4_{\rm theory}$ (PDG 2020)

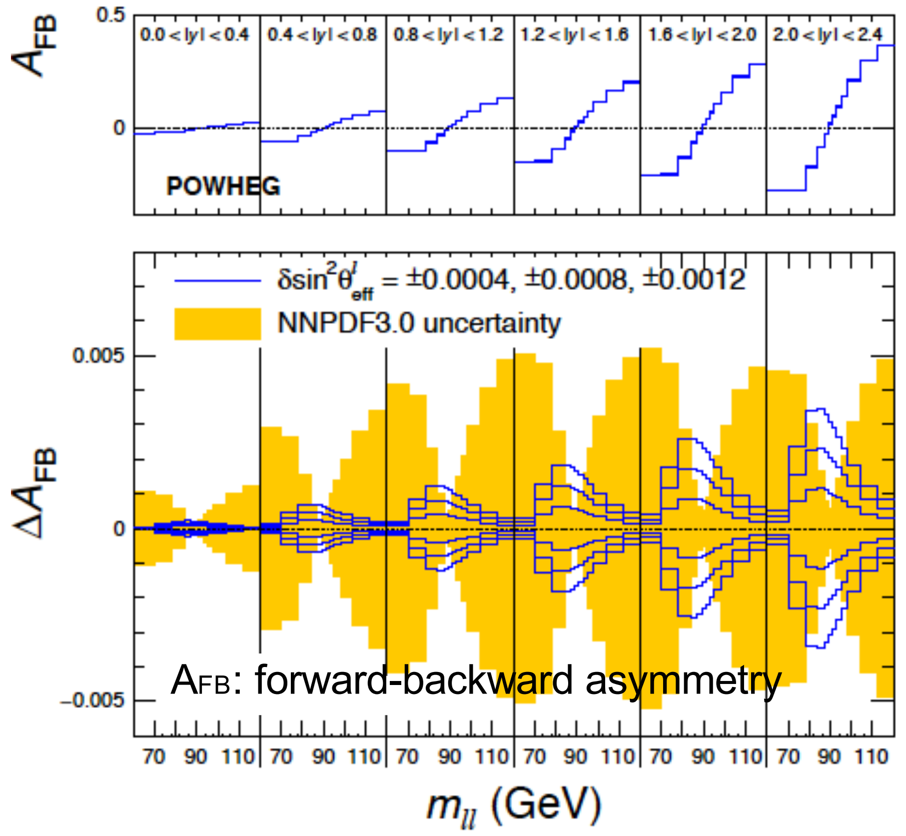
HCb measurement : $M_W = 80{,}354 \pm 23_{\rm stat} \pm 10_{\rm exp} \pm 17_{\rm theory} \pm 9_{\rm PDF}$ [JHEP 2022, 36 (2022)]

PDF unc. of CDF / ATLAS / LHCb: 3.9 / 8 / 9 MeV

$\sigma_{\rm PDF} \sim \sigma_{\rm tot}/2$

(up to)

$\sin^2\theta_W$

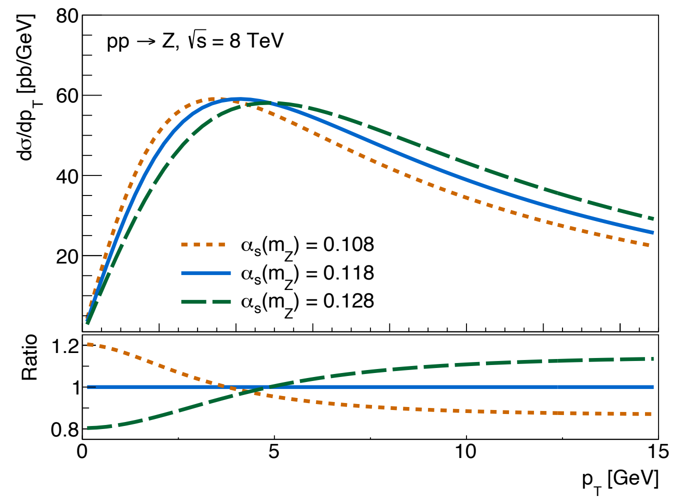


$\sin^2\theta^\ell_{\rm eff} = 0.23101 \pm 0.00036\,({\rm stat}) \pm 0.00018\,({\rm syst}) \pm 0.00016\,({\rm theo}) \pm 0.00031\,({\rm PDF})$

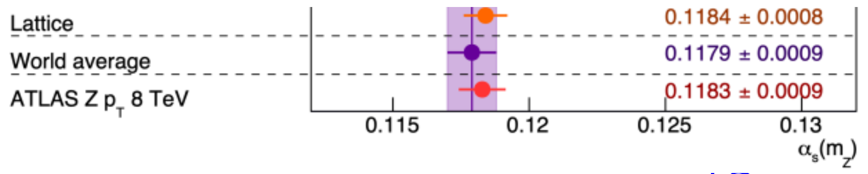
$\sigma_{\rm PDF} \sim \sigma_{\rm tot}/2$

4

α_S

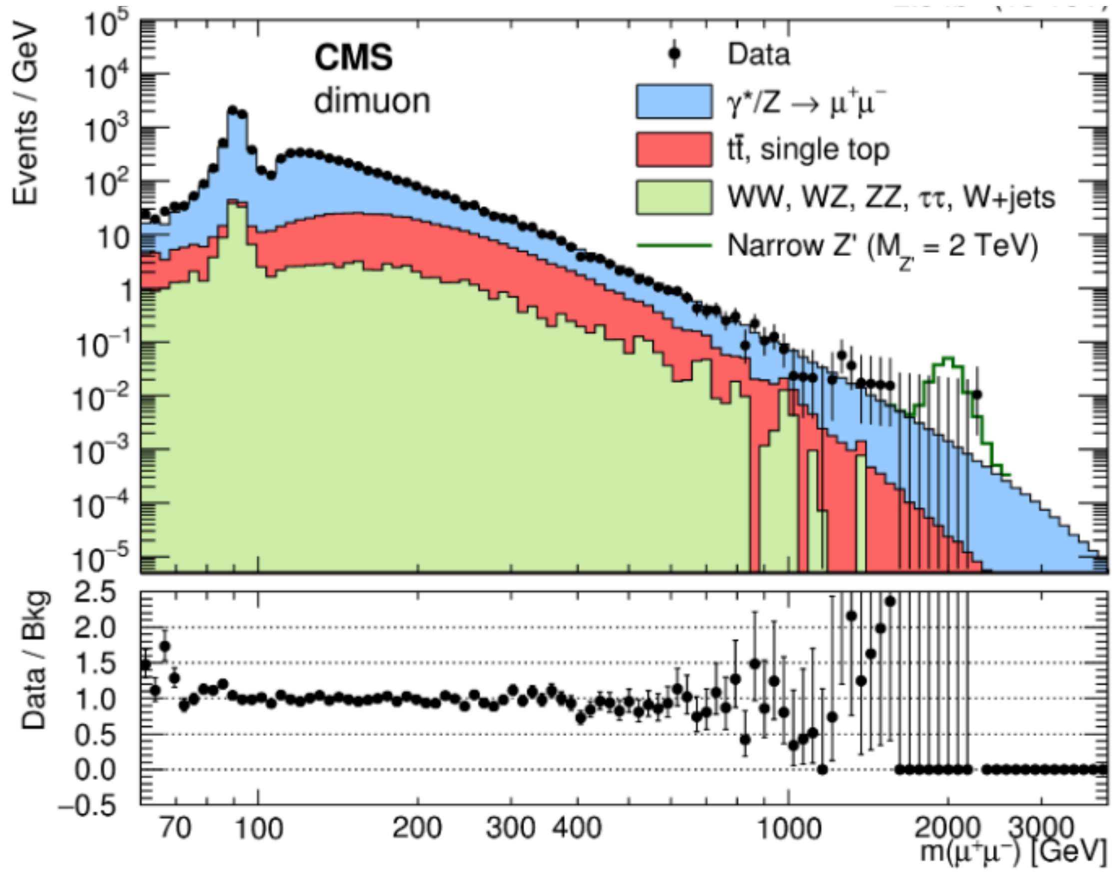


ATLAS, 2309.12986



$\sigma_{\rm PDF} \sim \sigma_{\rm tot}/2$

Disclaimer: will generally refer to papers by their arxiv number, even if published.



- The LHC is a BSM search machine. Often need PDFs here.
- High mass = high x , where PDFs are less well known. Key when looking for small/smooth deviations.

- Accuracy and precision in PDF determinations essential.

Global PDF fits: parameterisation

- Two distinct methodologies on the market to parameterising PDFs: **Neural Nets** (NNPDF) or **Explicit Parameterisation** (CT, MSHT).

✦ MSHT20: **52** free parameters in terms of Chebyshev polynomials.

$$xf(x,Q_0) = Ax^{\delta}(1-x)^{\eta}\left(1+\sum_{i=1}^{6} a_i T_i(y(x))\right) ,$$

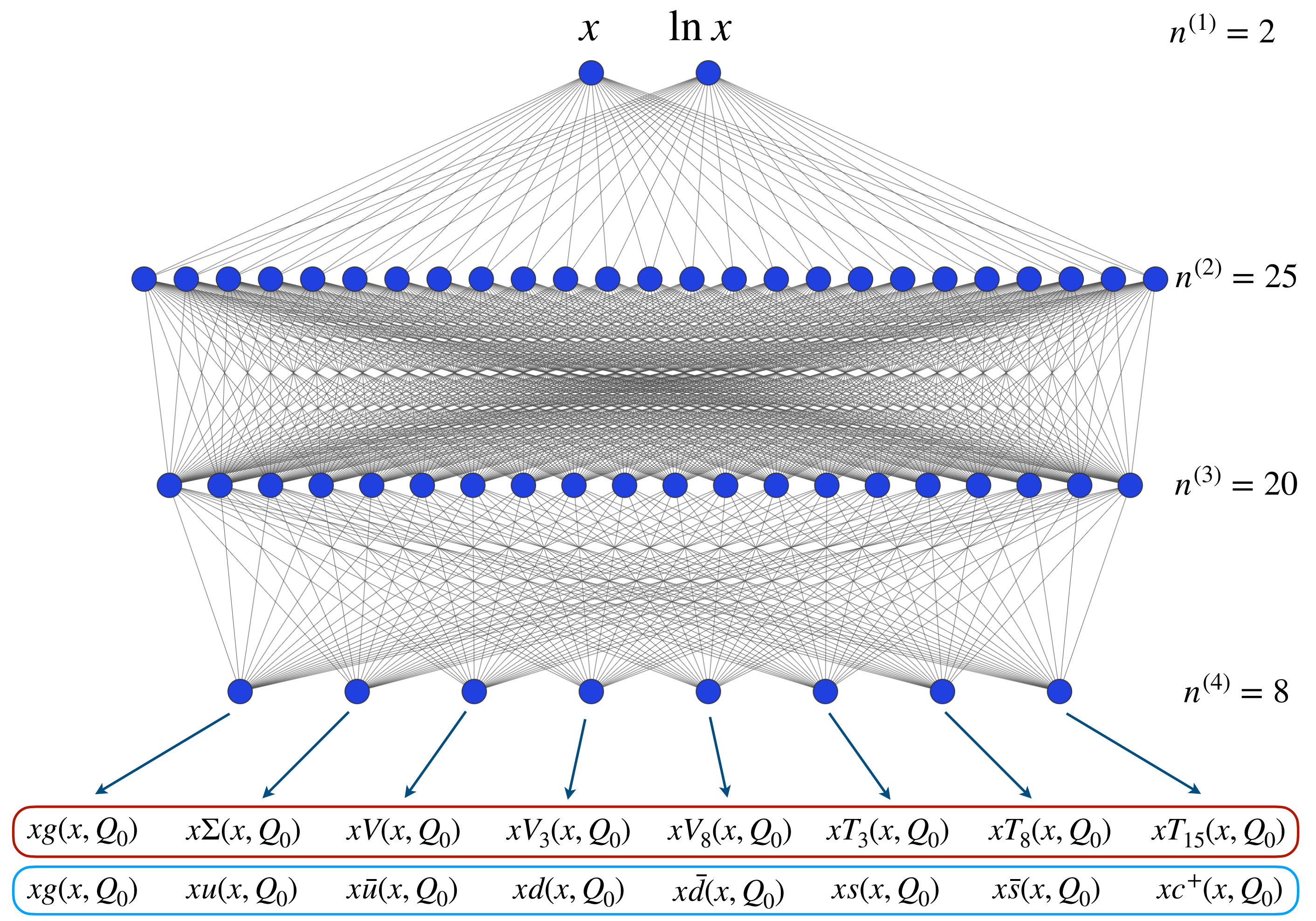
✦ CT18: **29** free parameters in term of Bernstein polynomials. Variations considered and PDF uncertainty expanded to account for these.

✦ Less flexible in general - need to be sure flexible enough! Allows direct handle on uncertainties in Hessian framework.

$f_i(x,Q_0):\ A_f x^{a_f}(1-x)^{b_f}\times$ ⟶ $\sum_{i=1}^{n}\alpha_{f,i}P_i(y(x))$, **CT, MSHT…**

⟶ $\mathrm{NN}_i(x)$ **NNPDF**

★ NNPDF4.0: **763** free parameter Neural Net.



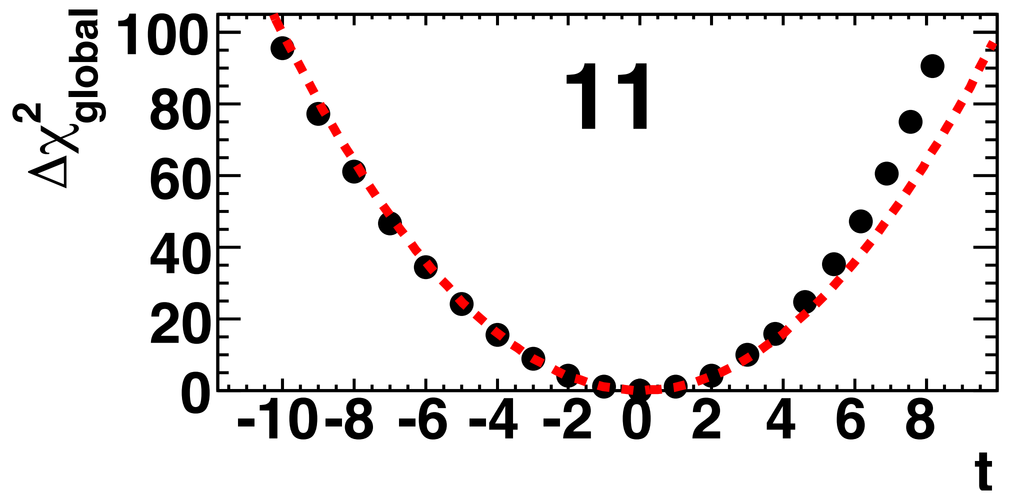
★ Increased flexibility, but needs robust optimisation + stopping (avoid over and under fitting).

Fixed Parameterisation PDFs

- Fixed parameterisation approach:

$$\chi^2_{\rm global} \sim \frac{(D_{\rm ata} - T_{\rm heory})^2}{\sigma^2} \qquad H_{ij} = \frac{1}{2} \frac{\partial^2 \chi^2_{\rm global}}{\partial a_i \partial a_j}\Big|_{\rm min}$$

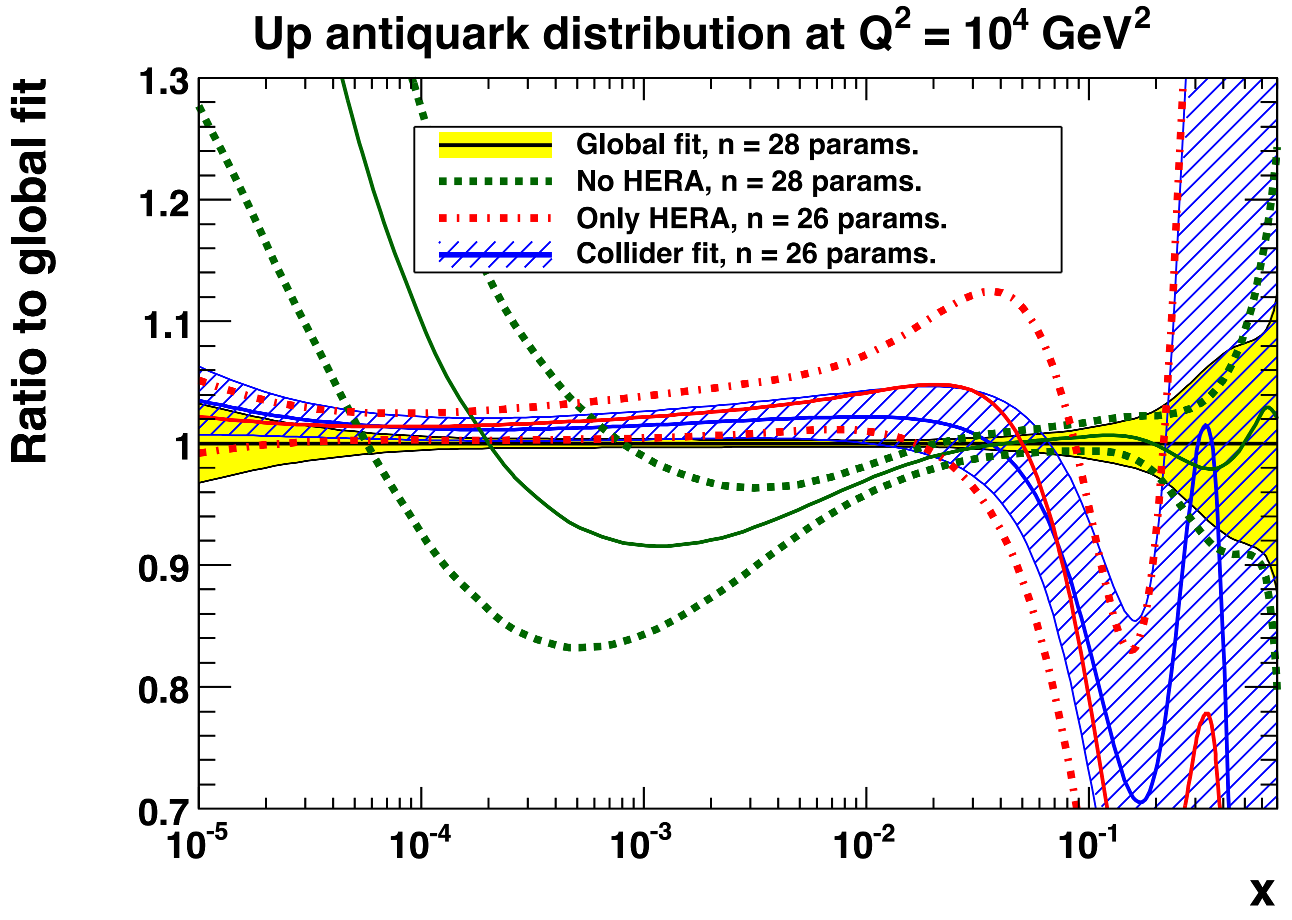
  - ✦ Find global minimum of χ^2 and evaluate eigenvectors of Hessian matrix at this point.
  - ✦ Parameter shifts corresponding to given $\Delta\chi^2$ criteria given in terms of these



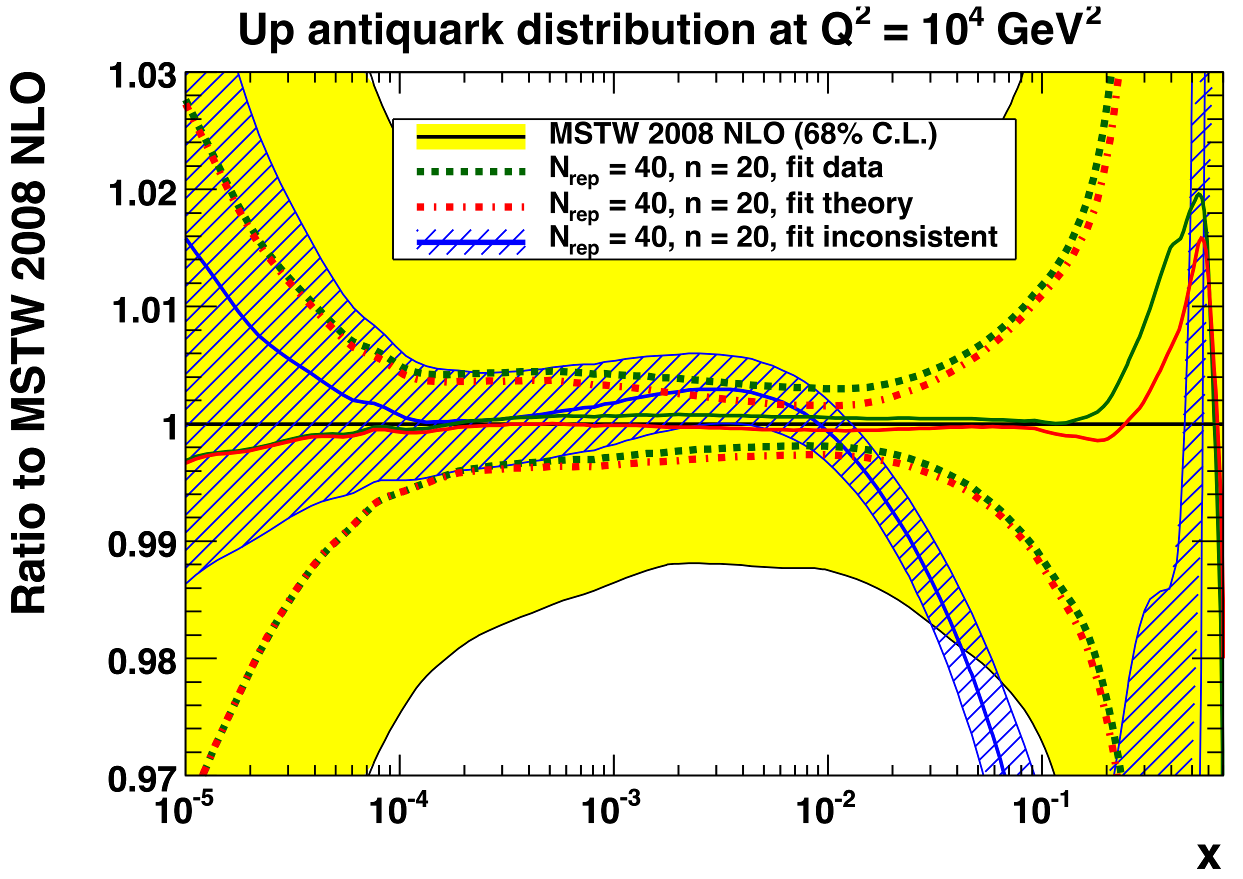
$$a_i(S_k^{\pm}) = a_i^0 \pm t\, e_{ik}, \qquad \text{with } t \text{ adjusted to give desired } T = \Delta\chi^2_{\rm global}$$

  - ✦ $T = 1$: \`textbook' criteria for case where data matches theory perfectly up to known determined (Gaussian) errors, across entire global dataset ($N_{\rm pts} \sim 4000 - 5000$).

  - ✦ Expect to not be sufficient: fit quality poor by textbook standard, dataset tensions, theory incomplete…
  - ✦ Backed up by evidence of e.g. fits to restricted datasets, or pseudodata with inconsistencies injected in.



$$\frac{\chi^2}{N_{\rm pts}} \gg 1 + \sigma(N_{\rm pts}) \sim 1.02$$



$\longrightarrow$ Motivates an enlarged $T > 1$, either fixed or 'dynamic'.

G. Watt and R. Thorne, arXiv:1205.4024

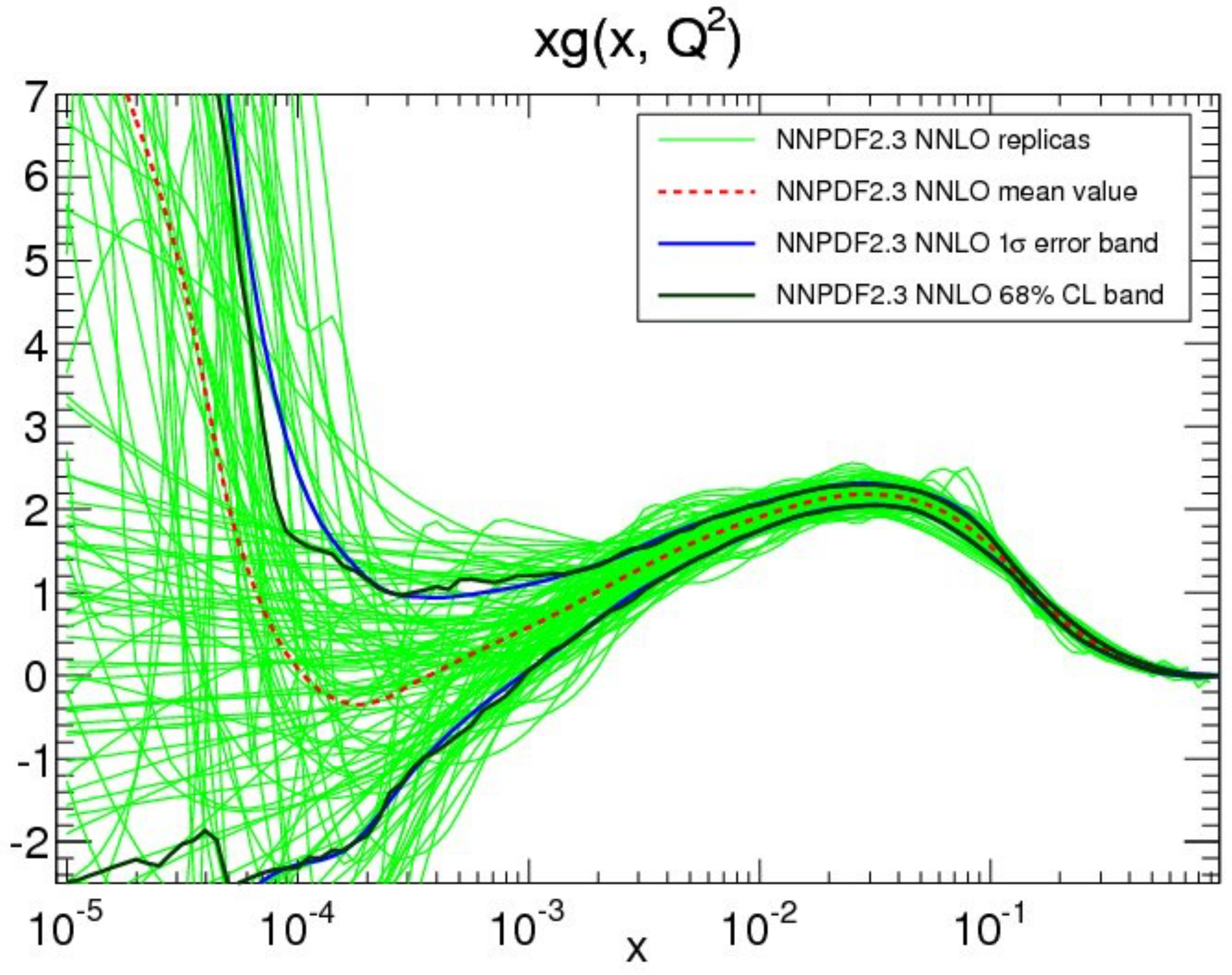
See also, J. Pumplin, arXiv:0909.0268

And arXiv.2406.01664 Today!

Neural Network PDFs

- Neural network approach:
  - ✦ Generate set of MC `replicas' by shifting data by errors.

    Each D_i gives f_i and from $\{f_i\} \Rightarrow$ PDF errors



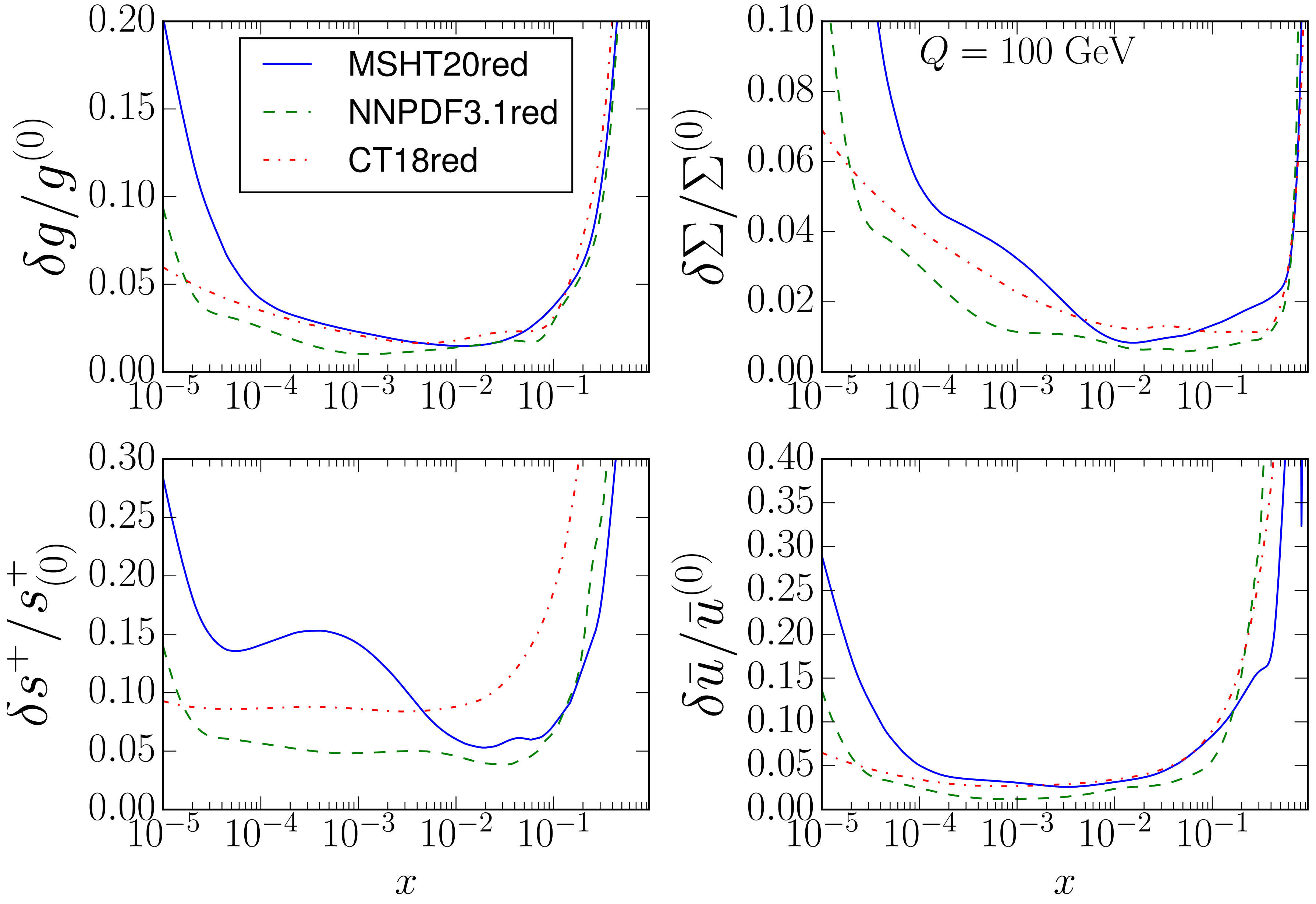
G. Watt and R. Thorne, arXiv:1205.4024

✦ Note not specific to NNs: can apply in fixed parameterisation as well: shown to be ~ equivalent to Hessian $\Delta\chi^2 = 1$ in that case.

✦ However, in NN approach direct correspondence is lost as Hessian approach does not apply.

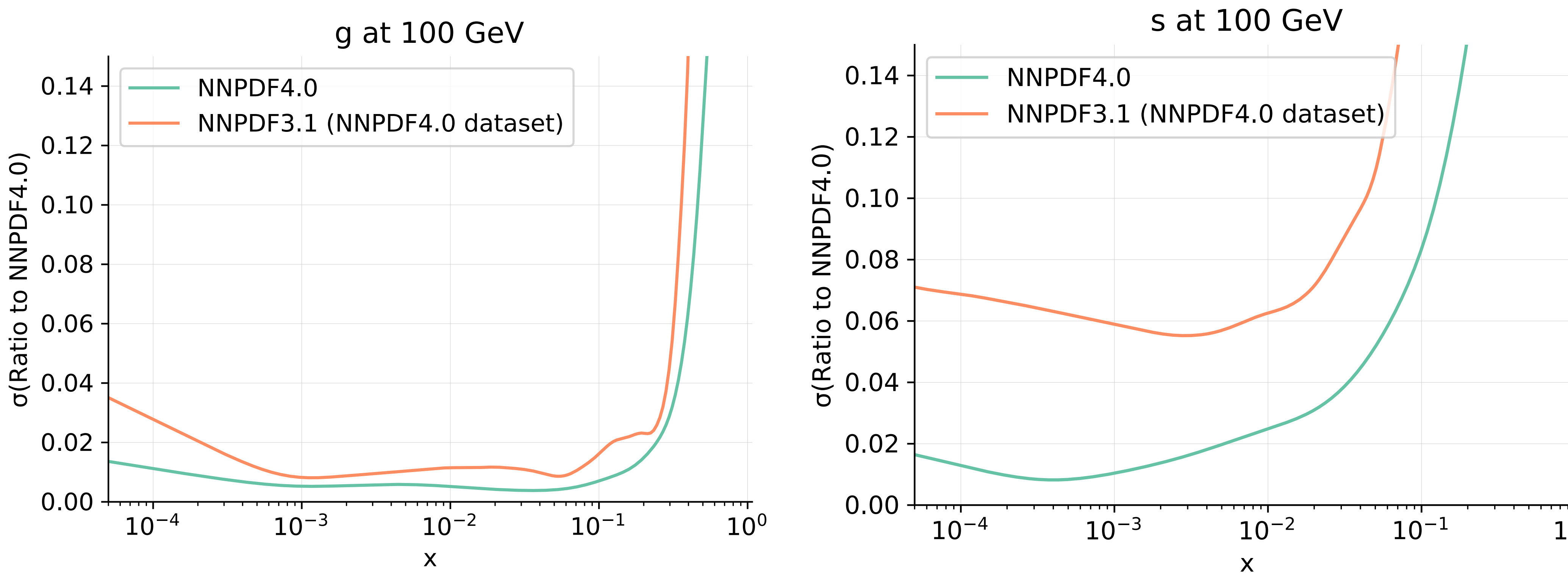
✦ Global fits give different errors in PDF4LHC21 benchmarking. NNPDF3.1 in general **smaller errors.**

Benchmark = similar data/settings



PDF4LHC21, arXiv:2203.05506

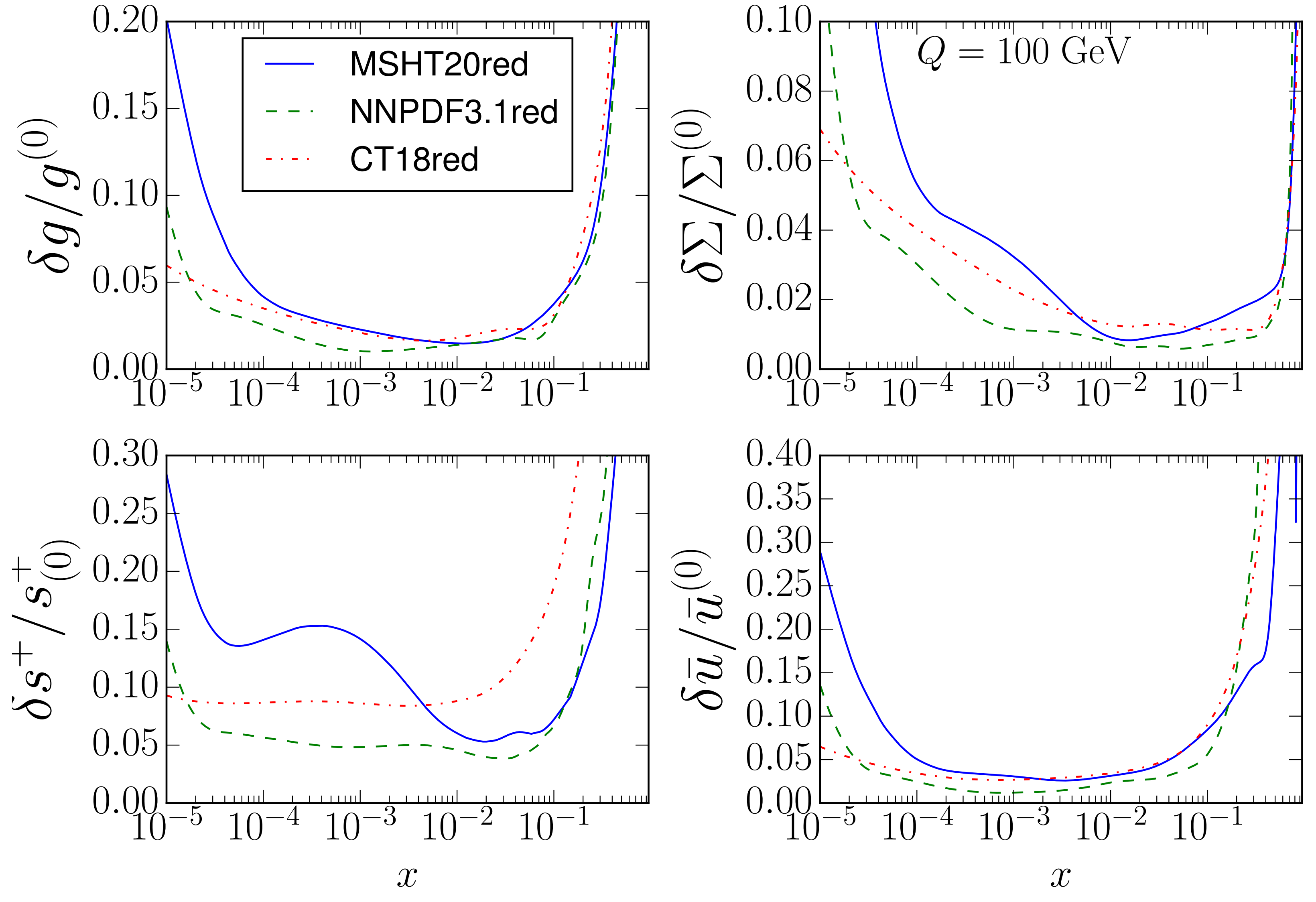
✦ And **4.0 methodology** gives further errors reduction.



Aims of Talk

★ Suggests three possibilities:

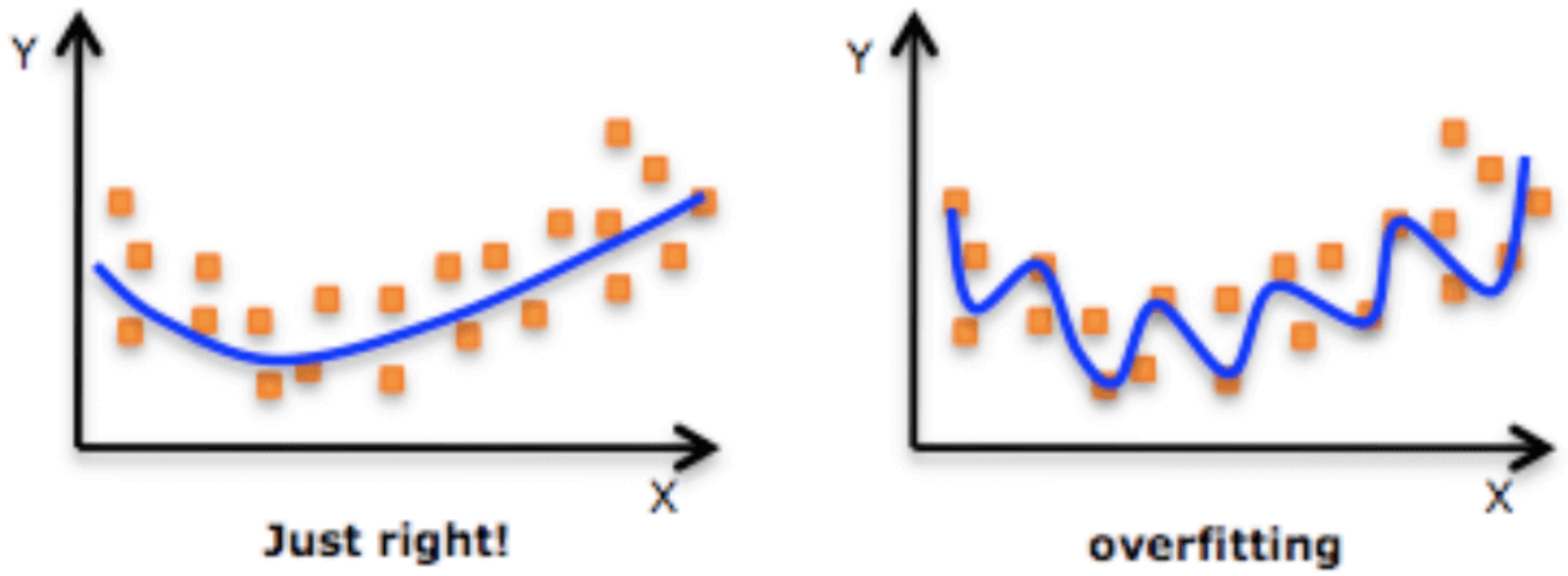
1. **NNPDF4.0** uncertainty not conservative enough (**too small**).
2. **MSHT (CT)** uncertainty too conservative (**too large**).
3. **MSHT (CT)** fit less accurate, due to parameterisation inflexibility, and hence enlarged errors needed (**less precise**).



- Or some combination of the three. Finding out which clearly important for LHC precision.

★ In addition, CT18 approach of fitting with more restricted parameterisation (29 vs. 52 free parameters) motivated by fit stability/avoiding overfitting.

- Do we see evidence of this in MSHT20 fit? What happens if we restrict parameterisation?



Aims of Talk

- In this talk I will present results that aim to address these issues. In particular will show:

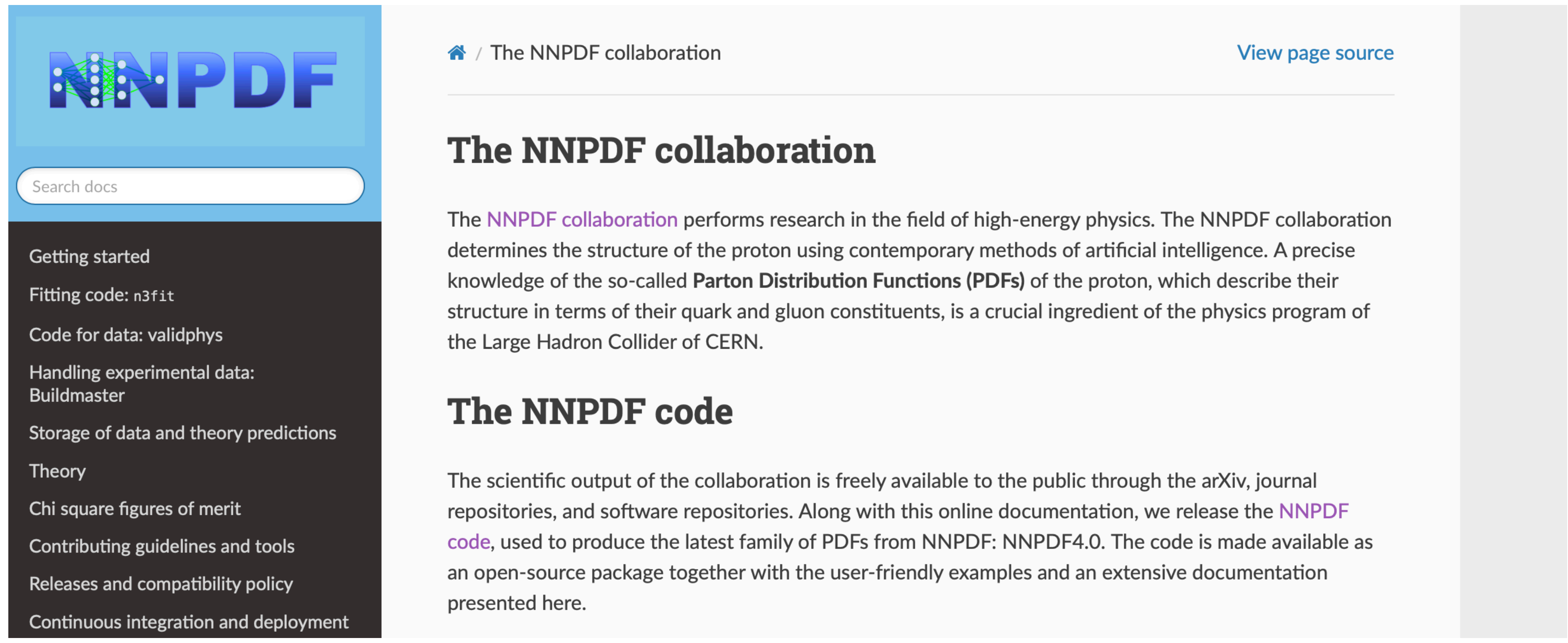
  ★ **First** global **closure test** of fixed parameterisation (MSHT) approach: is parameterisation flexible enough to give faithful description of global pseudodata? What if we restrict it?

  ★ **First** completely direct **comparison** between fixed parameterisation (MSHT) and NN approaches. How do these compare in full global fit?

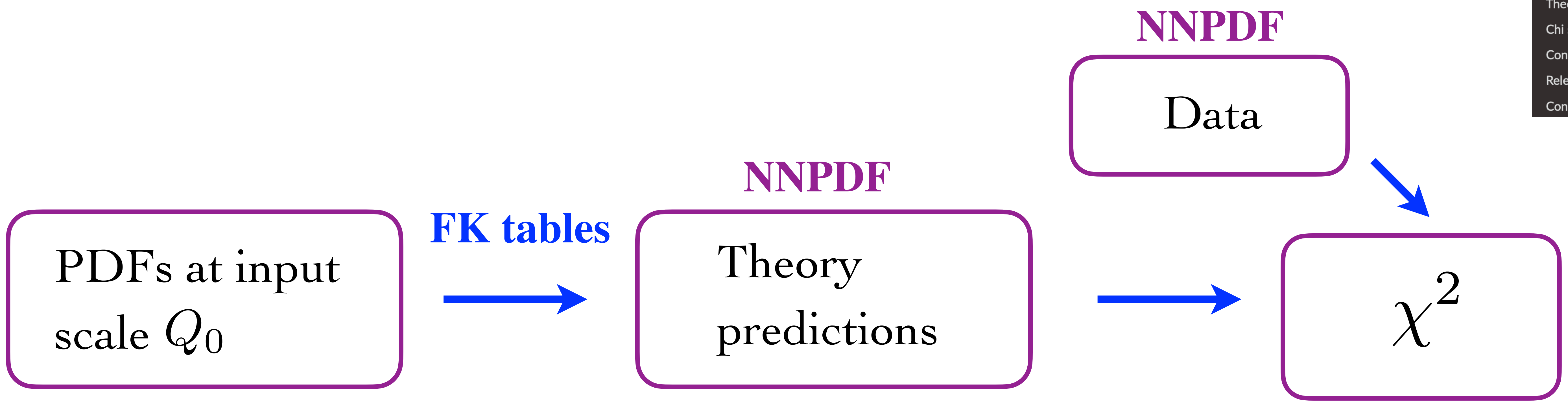
- Study is ongoing, so all slides can be viewed as if they have a `**preliminary**' label on them! Given time + audience, will focus more on first bullet point here.

Global Closure - set up

- How best to set up a global closure test? Will make use of publicly available NNPDF fitting code.
- Provides python libraries to load NNPDF dataset and theory predictions, given PDF set. More precisely gives:



https://docs.nnpdf.science/



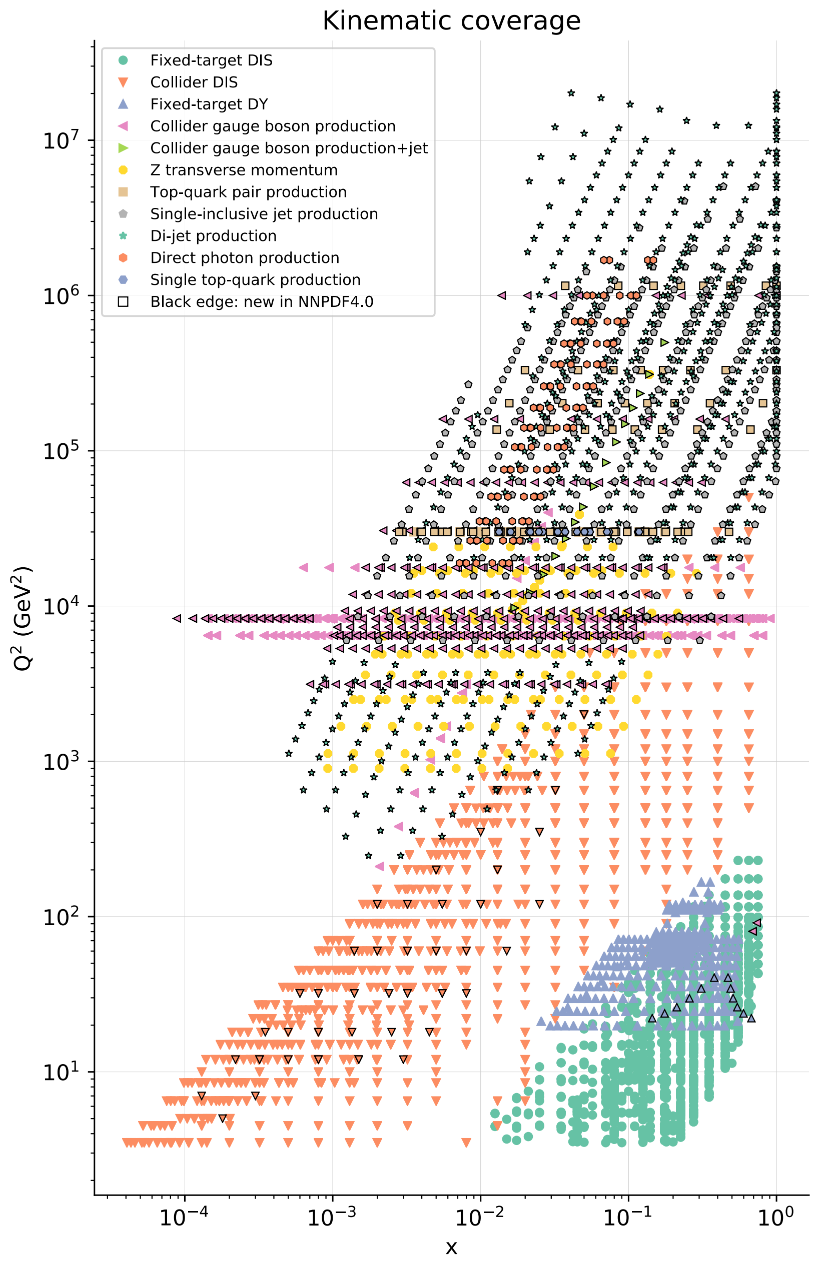
- Given arbitrary PDF set (grid of $\{f_i\}$ at $\{x_i\}$ and Q_0) can evaluate theory predictions + fit quality.
- This allows us to evaluate corresponding fit quality with a (MSHT) fixed parameterisation, but to NNPDF data/theory - **only difference** is **input parameterisation.** From above module can also build up optimizer in usual way to give best fit, Hessian errors etc.
- Will use for closure tests (though not essential) - but setting things up in this way will allow direct comparison at level of full fit.

Global Unfluctuated Closure Test

Always NNLO

- For direct comparison will consider perturbative charm - NNPDF4.0pch set as input.
- Then generate unshifted pseudodata for 4.0 global dataset ($N_{\rm pts} = 4627$). In principle exact agreement possible, with $\chi^2 = 0$. Then perform fit with default MSHT parameterisation. What do we find?

$$\chi^2_{\rm global} \sim \frac{(D_{\rm ata} - T_{\rm heory})^2}{\sigma^2}$$



NNPDF, arXiv:2109.02653

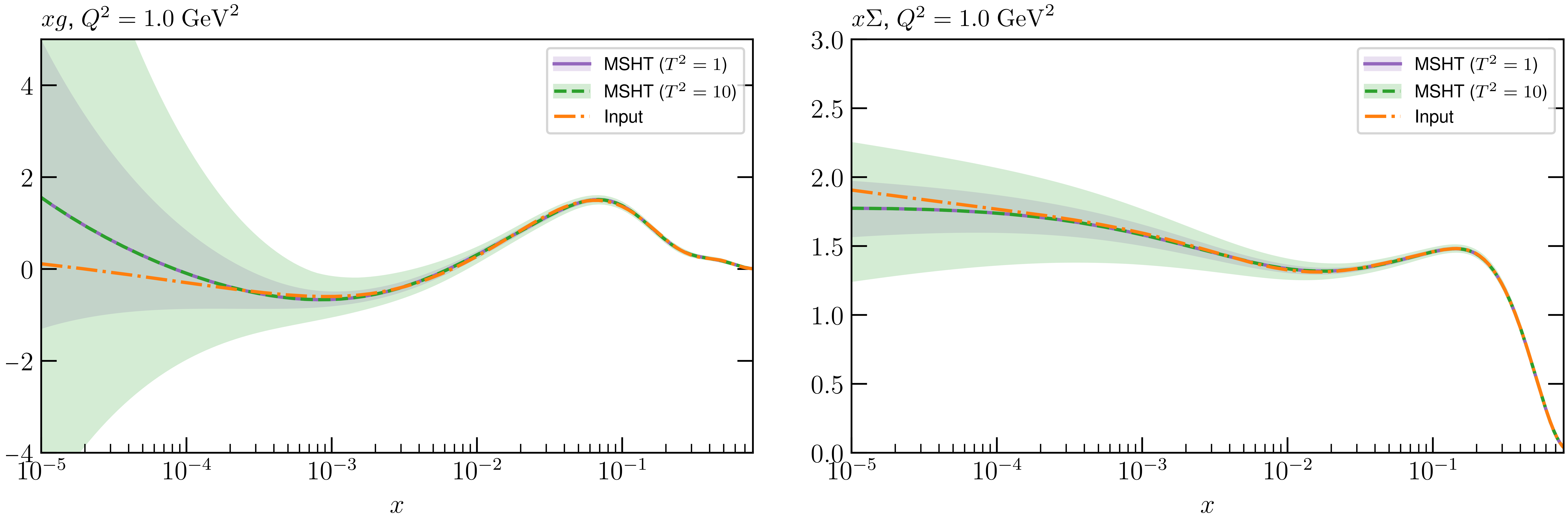
	χ^2	$\chi^2/N_{\rm pts}$
Fit quality:	2.4	0.0005

- **Remarkably good!** In fact lower than reported result of NNPDF L0 closure test.

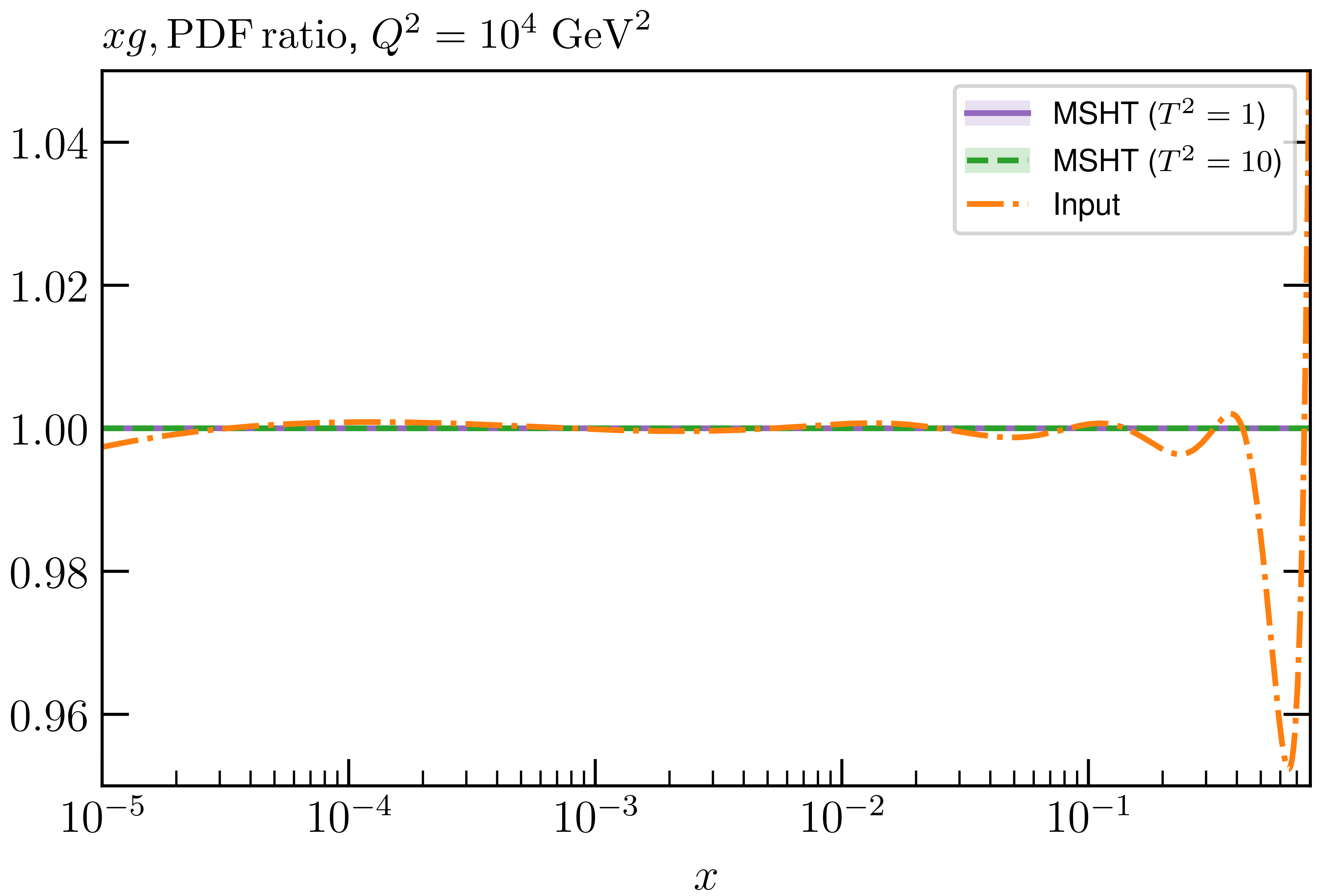
L. Del Debbio, T. Giani and M. Wilson, arXiv:2111.05787

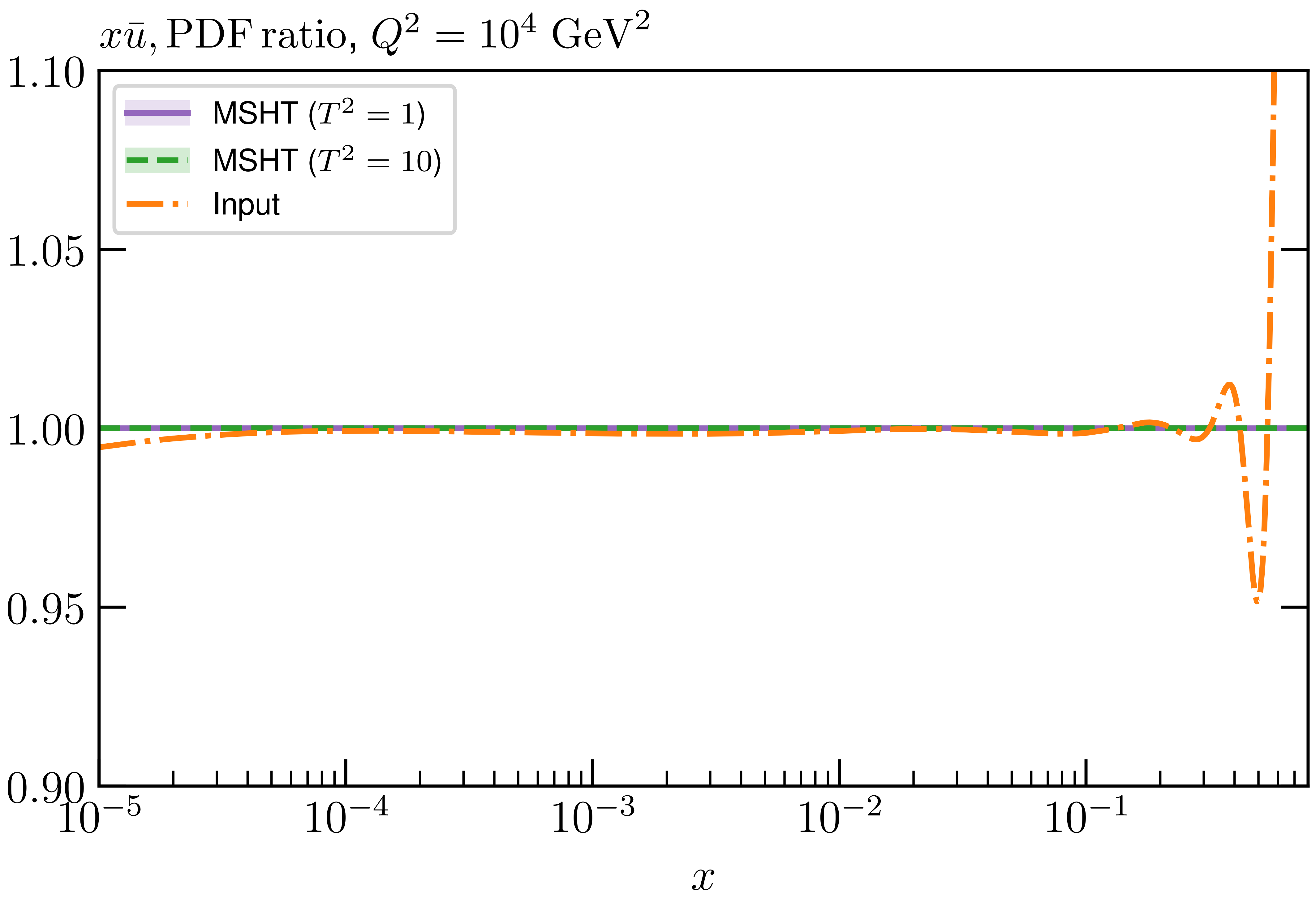
	3.1 meth.	**4.0 meth.**
$\chi^2/N_{\rm pts}$	0.012	0.002

- **Caveat**: only one input set, may well be different (not quite as good) for others. Trend should be similar.
- But apparently no issue with parameterisation inflexibility in this case. But what about PDFs?



- First look: encouraging results! In more detail…





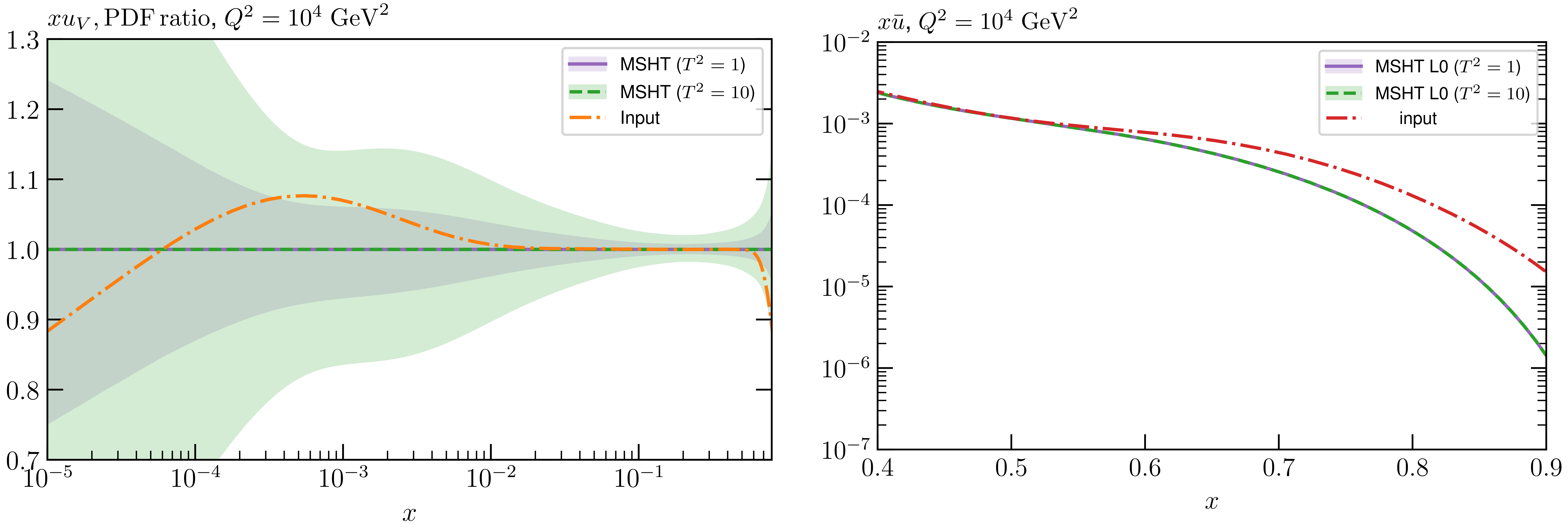
- Ratio of (NNPDF4.0pch) input to fit result, including PDF uncertainties with $T^2 = 1$ and 10 that come from the closure test fit. Latter is ~ result of dynamic tolerance used in MSHT20 (checked here).

Similar results for other quarks - see backup

★ **Deviation** in general (in data region) **per mille** level and well within the $T^2 = 1$ uncertainties.

★ More precisely, deviation is ~ 10% or less of $T^2 = 1$ uncertainty, and a factor of $\sim 2-5$ lower a a fraction of the $T^2 = 10$ uncertainty.

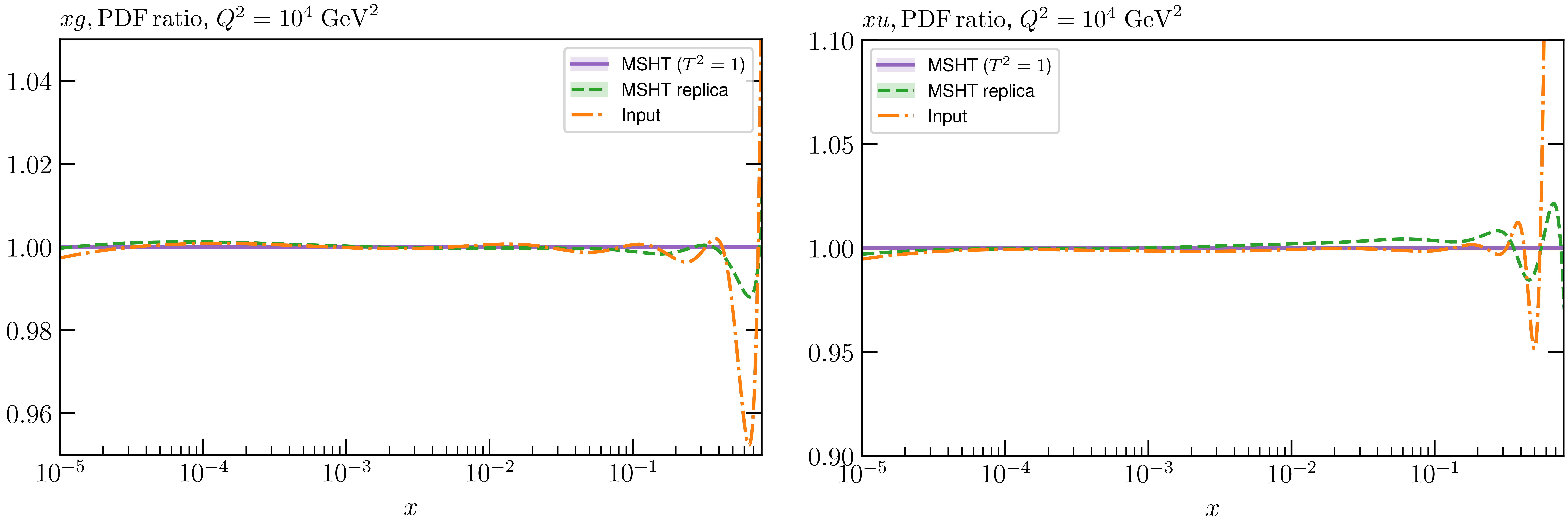
$\longrightarrow$ In **data region** input PDF matched very well, and much better than $T^2 = 1$ uncertainties. **No evidence** that the increased tolerance is driven by **parameterisation inflexibility** for MSHT.



- In less well constrained regions deviation larger, e.g for u_V, d_V at low and high x and the $\overline{u}, \overline{d}$ at high x .
- Hence in extrapolation region input not always consistent within uncertainties
- As ~ outside data region not inconsistent (errors driven by data), but indicates more conservative error definition in these regions may be desirable (as tends to happen in NN approach).
- Though arguably no 'right' answer in true extrapolation region (too conservative vs. over-conservative).

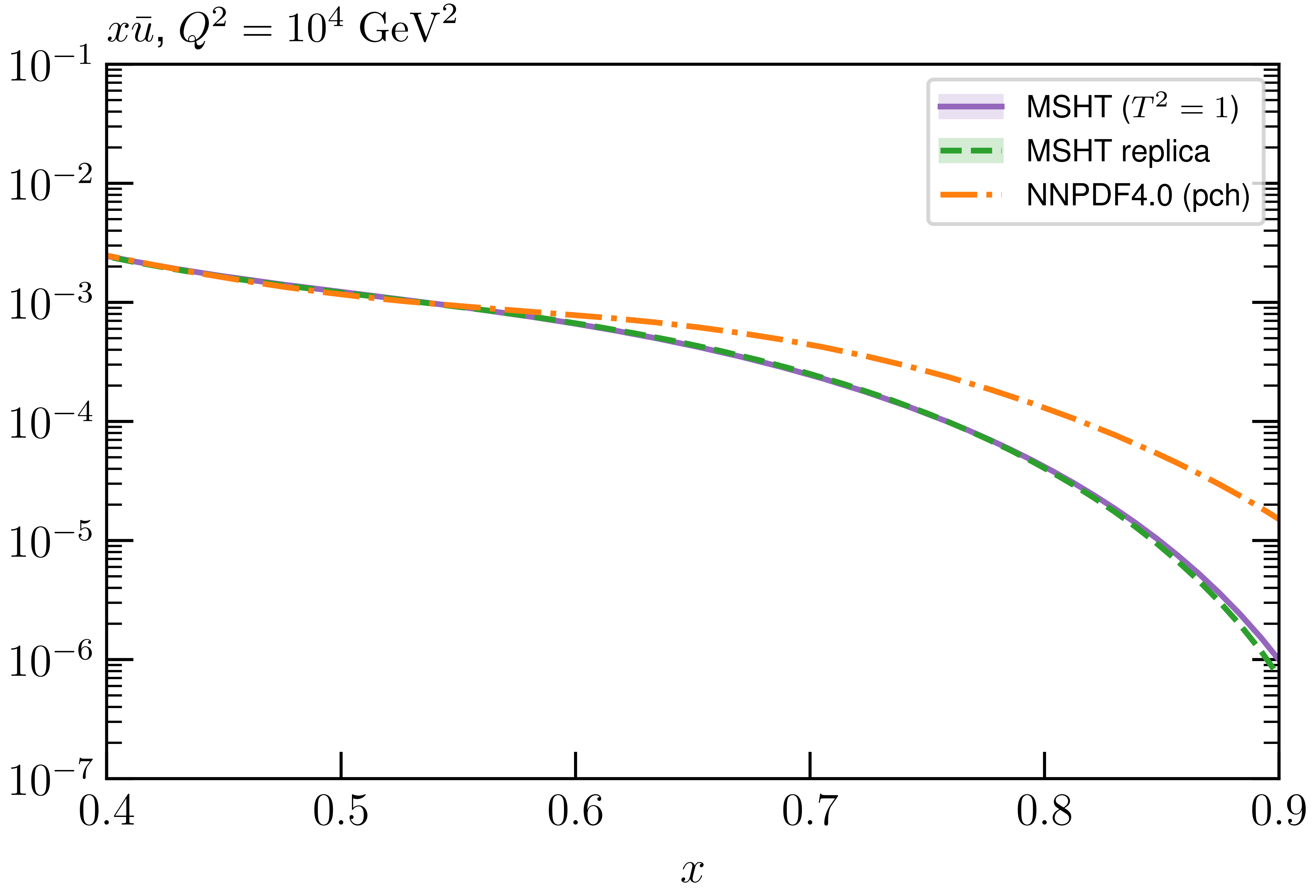
Global Fluctuated Closure Test

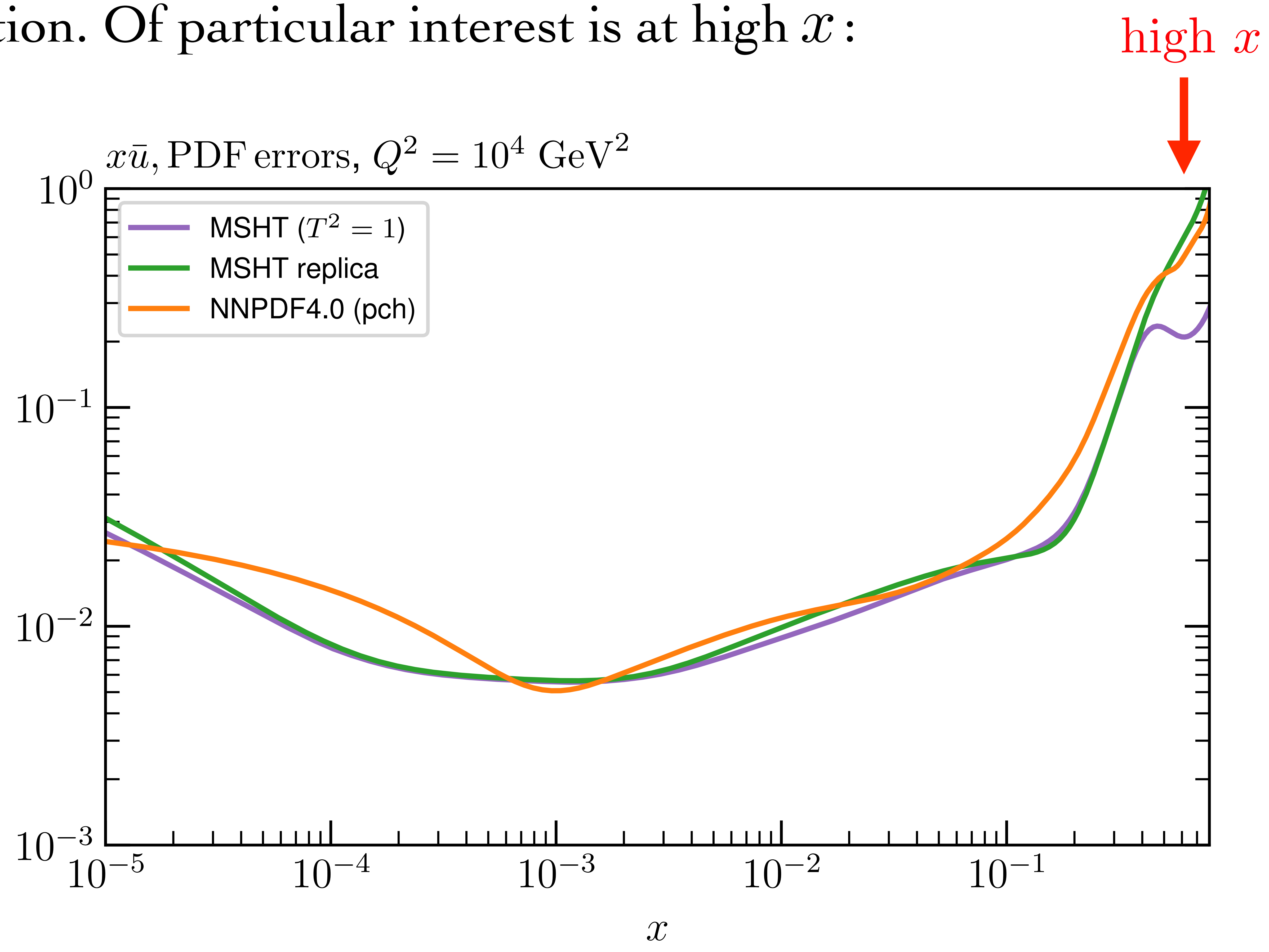
- Exactly the same closure test settings, but fluctuate pseudodata according to experimental uncertainties. Fit quality $\chi^2 \sim N_{\rm dat} \pm \sqrt{2N_{\rm dat}}$ expected (and found).
- Test faithfulness of MSHT parameterisation by producing MC replica set - perform 100 replica fits.



- Find very encouraging agreement between MC replica and Hessian uncertainties. Would not necessarily expect if issues with parameterisation inflexibility (or overfitting).

- However exact agreement between Hessian and MC replica approach only expected in exact Gaussian approximation. Away from this can see some deviation. Of particular interest is at high x :





- MC replica uncertainty much larger here - helps improve matching with input set.
- Much more in line with NNPDF uncertainty. Perhaps MC replica propagation (rather than NN) playing (most?) significant role here?

Comparison to NNPDF uncertainties

- Can do first comparison of **MSHT** vs. **NNPDF** PDF **uncertainties**. Not completely direct as MSHT is in closure test, and NNPDF result of full fit. But theory and underlying datasets same. Find:

★ Quark flavour decomposition: $\sigma(\text{NNPDF}) \sim \sigma(\text{MSHT}, T^2 = 1)$

★ Gluon (singlet at intermediate x): $\sigma(\text{MSHT}, T^2 = 1) \lesssim \sigma(\text{NNPDF}) \lesssim \sigma(\text{MSHT}, T^2 = 10)$

MSHT, $T^2 = 1$
MSHT, $T^2 = 10$
NNPDF4.0pch

- With rather similar overall trends with x .
- Exception at high x where NNPDF uncertainty becomes larger.

Restricted Parameterisations

- Returning to unfluctuated closure test, can explore impact of restricting/extending parameterisation:

$$xf(x,Q_0) = Ax^{\delta}(1-x)^{\eta}\left(1+\sum_{i=1}^{n} a_i T_i(y(x))\right) ,$$

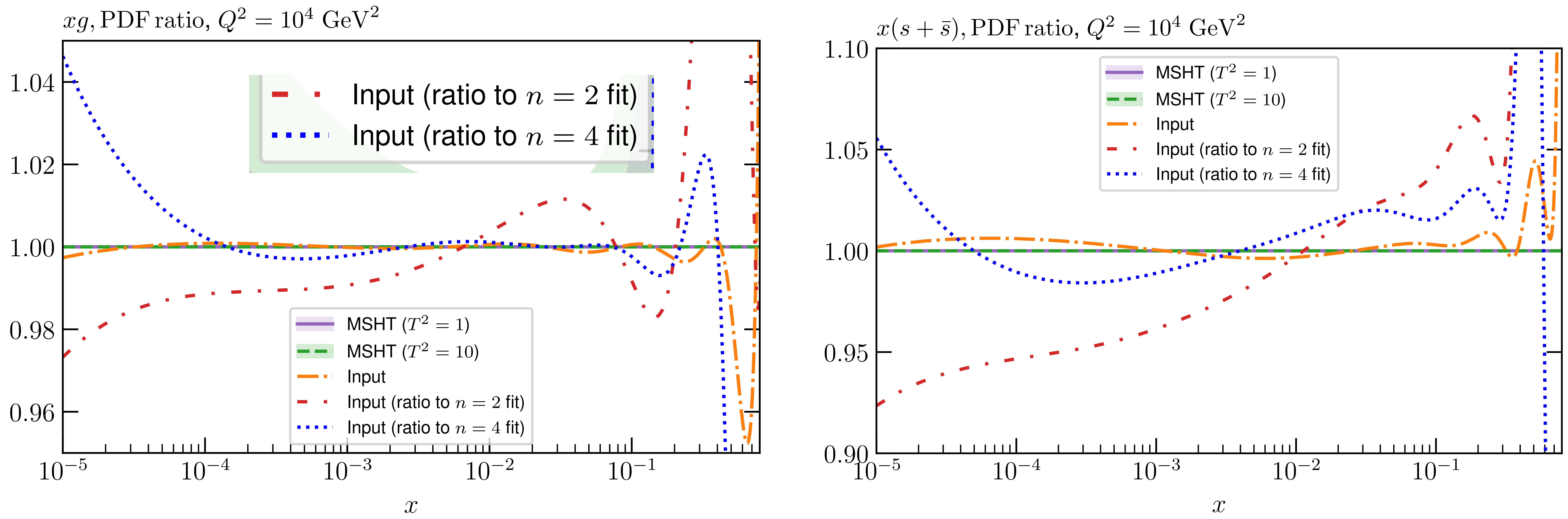
$n = 2, 4 \Rightarrow n_{\rm par} = 28, 40$

$n = 6 \Rightarrow n_{\rm par} = 52$

- Recall for CT18, have 29 free parameters.
- Impact on fit quality: $n = 2, 4 \Rightarrow \chi^2 = 103.8, 69.0$

i.e. significant deterioration with respect to default ($n = 6$) case, where $\chi^2 = 2.4$.

- Impact on PDFs?



- Matching between input and fit significantly worse then in the default case.
- For $n = 4$ case already matching on edge of (or beyond) $T^2 = 1$ uncertainties. For $n = 2$ (~ CT18) matching beyond $T^2 = 1$ and even up to $T^2 = 10$.
- $T^2 = 1$ uncertainty is representative of experimental uncertainty $\Rightarrow$ $n = 2, 4$ parameterisations insufficient to match precision in state-of-the-art global fits. Can enlarge tolerance to account for this but more flexible $n = 6$ parameterisation would allow more accurate result.
- Also indicates that $n = 2, 4$ parameterisations unlikely to lead to overfitting, given description poor here.

Full fit: comparison

- Can also consider result of fit to real data entering NNPDF4.0 fit. To restate: exactly **same data and theory**, with **only difference** from PDF input **parameterisation**.
- Will in addition consider case where positivity is imposed at PDF (and cross section) level, as in NNPDF fit, for $x_i \in \{5\cdot 10^{-7}, 0.9\}$. Not something that is done in MSHT fits!

	NNPDF4.0 pch	MSHT fit	MSHT fit (w positivity)
$\chi^2_{t_0}$	5928.3 (1.282)	5736.7 (1.240)	5837.8 (1.262)

$$\Delta\chi^2_{t_0}: \quad -191.6\,(-0.04) \quad -90.5\,(-0.02)$$

$\longrightarrow$ Fit quality with **MSHT** parameterisation is **significantly better** than result of central NNPDF set.

- **Positivity** clearly plays significant role - completely dominated by low x gluon. Indeed removing it from pure NNPDF fit gives O(100) improvement. But **not the only difference** here.

Integrability checked and not issue: Backup

- Improvement spread across fixed target, HERA DIS (without positivity) and LHC DY data.
- Do not expect central replica $\chi^2_{\mathrm{rep},0}$ to be absolute minimum of χ^2 but difference too large for this.

See Backup

PDFs

- Comparing PDFs, see clear effect of positivity at low x: driven by known fact that default result prefers gluon to be negative at low x, Q^2. Trend also seen if positivity removed from pure NNPDF fit.
- Imposing **positivity** gives much better agreement at low x, but clear **difference** in flavour decomposition **remains**.
- NNPDF vs. MSHT uncertainties v. similar to closure test:

★ Quark flavour decomposition: $\sigma(\text{NNPDF}) \sim \sigma(\text{MSHT}, T^2=1)$

★ Gluon (singlet at intermediate x): $\sigma(\text{MSHT}, T^2=1) \lesssim \sigma(\text{NNPDF}) \lesssim \sigma(\text{MSHT}, T^2=10)$

NNPDF4.0pch MSHT (pos) $T^2 = 10$ MSHT, $T^2 = 10$

Full fit with fitted charm

- Readily extend previous study to include **fitted charm**. NNPDF theory inputs change accordingly, while PDFs parametrised at $Q_0 = 1.65\,\mathrm{GeV}\,(> m_c)$ rather than 1 GeV, and parameterise charm:

$$xc_+(x,Q_0) = A_{c_+} x^{\delta_{c+}} (1-x)^{\eta_{c+}} \left(1 + \sum_{i=1}^{6} a_{c,i} T_i(y(x))\right) \qquad xc_-(x,Q_0) = 0$$

- Find:

	NNPDF4.0	MSHT fit	MSHT fit (w positivity)
$\chi^2_{t_0}$	5692.1 (1.233)	5645.2 (1.222)	5651.0 (1.224)

$$\Delta\chi^2_{t_0}: \qquad -46.9\,(0.011) \qquad -41.1\,(0.009)$$

$\longrightarrow$ Fit quality with **MSHT** parameterisation again **better** than result of central NNPDF set, albeit by less than in perturbative charm case.

- Difference in PDFs than in p. charm case, but **clear difference** in **flavour decomposition** remains.
- Role of **positivity** now **marginal** (confirmed with direct NNPDF fit).
- NNPDF vs. MSHT uncertainties similar to p. charm.

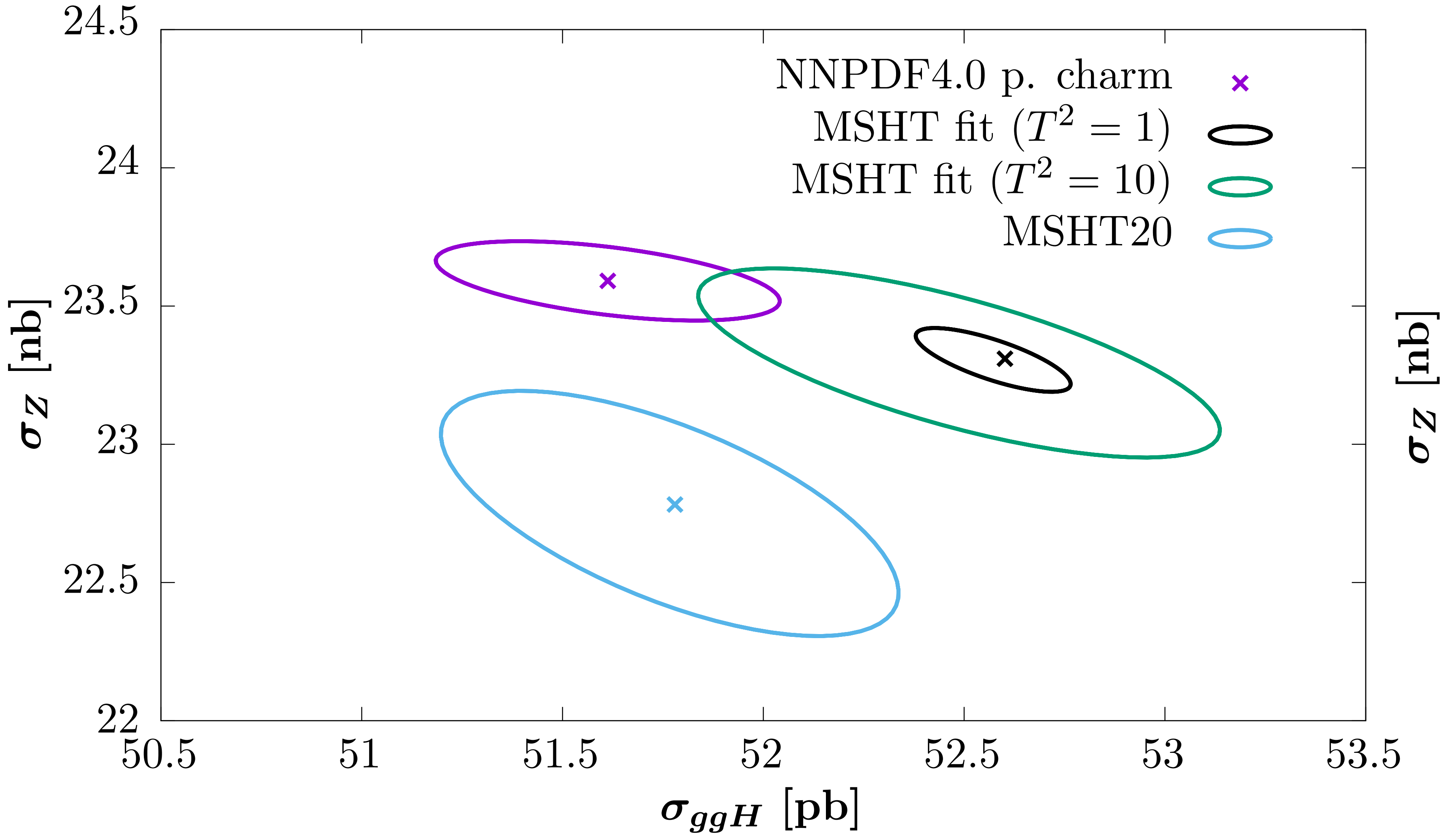
⋆ Quark flavour decomposition: $\sigma(\mathrm{NNPDF}) \sim \sigma(\mathrm{MSHT}, T^2 = 1)$

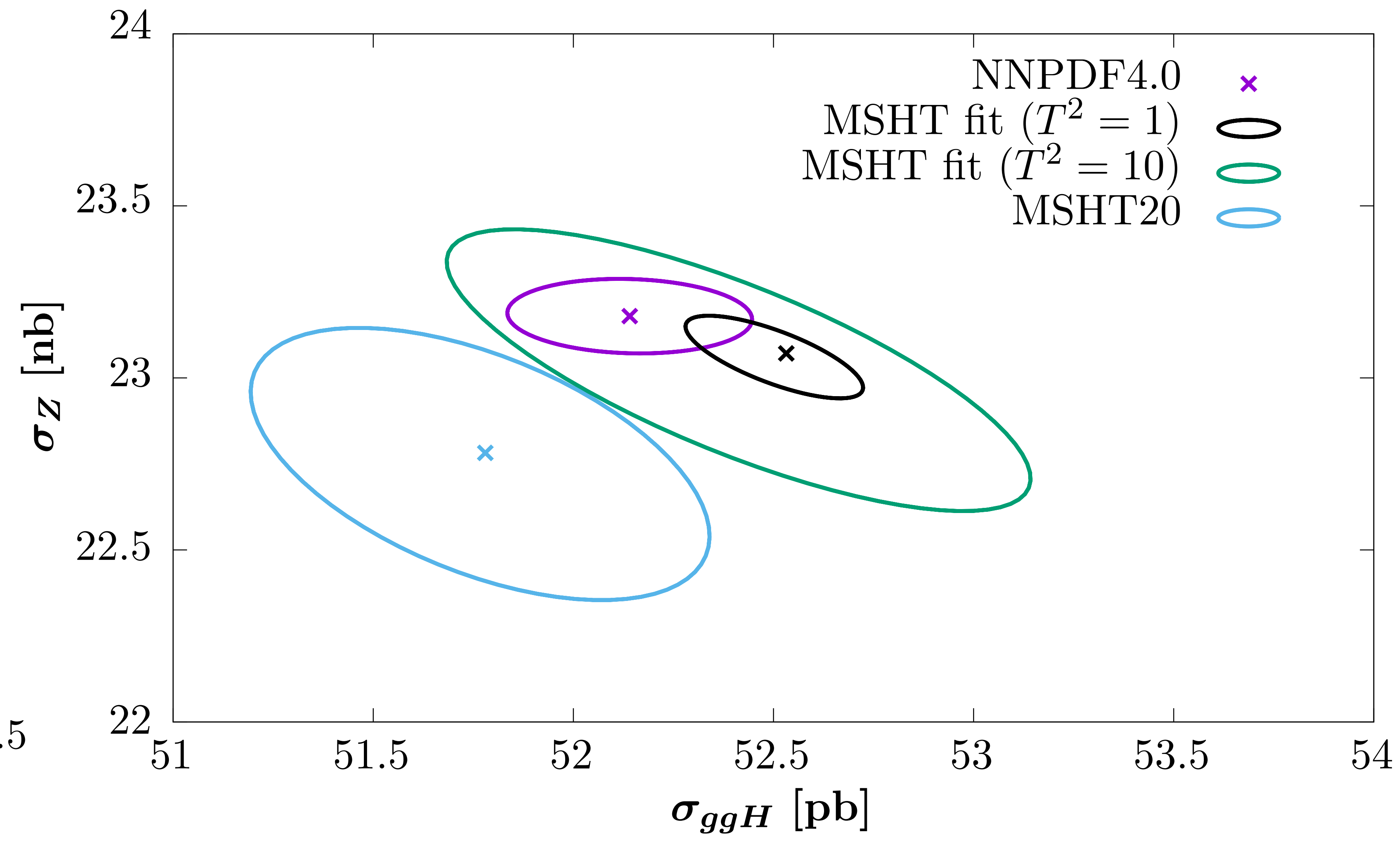
⋆ Gluon (singlet at intermediate x): $\sigma(\mathrm{MSHT}, T^2 = 1) \lesssim \sigma(\mathrm{NNPDF}) \lesssim \sigma(\mathrm{MSHT}, T^2 = 10)$

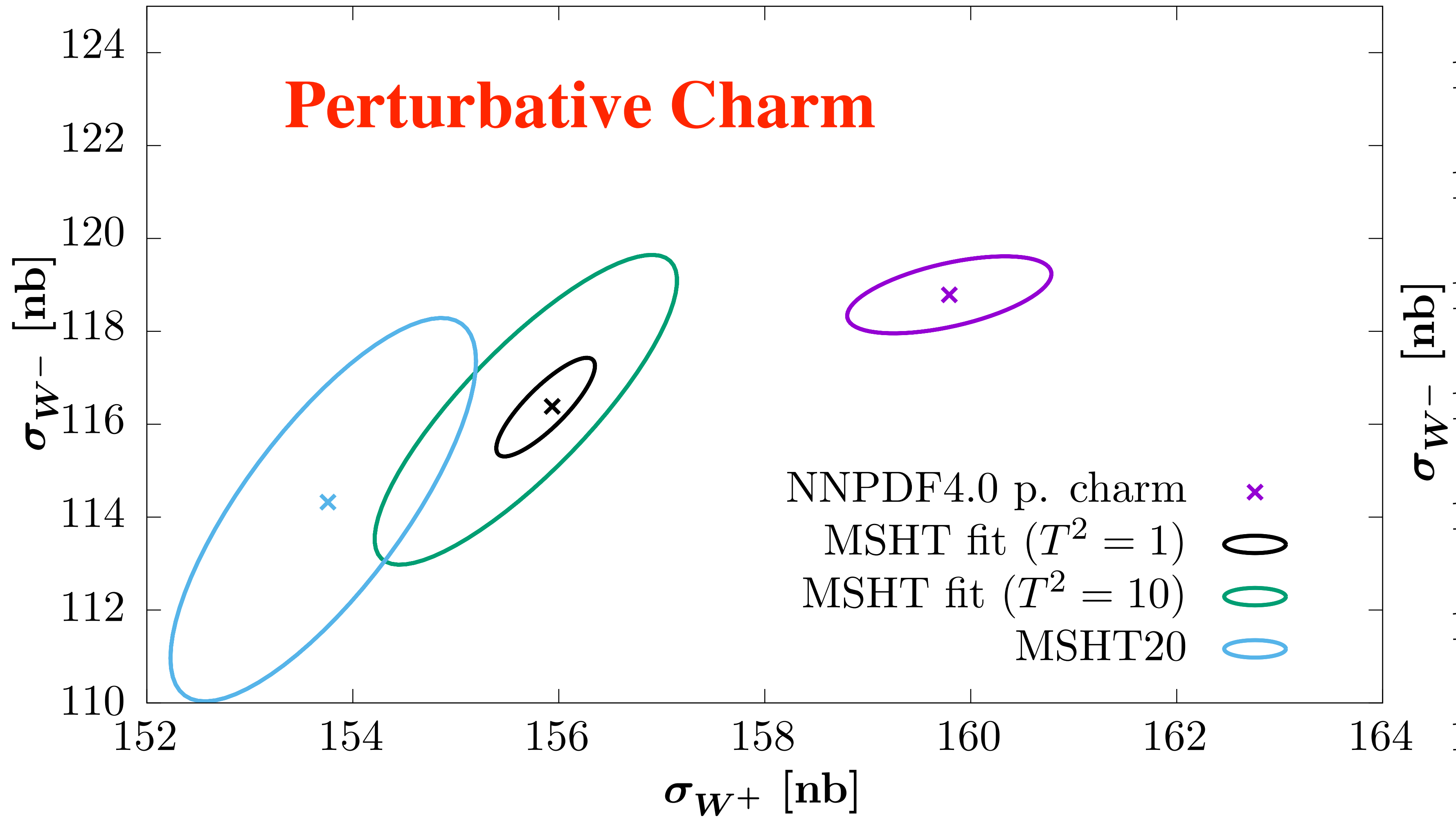
Cross Sections

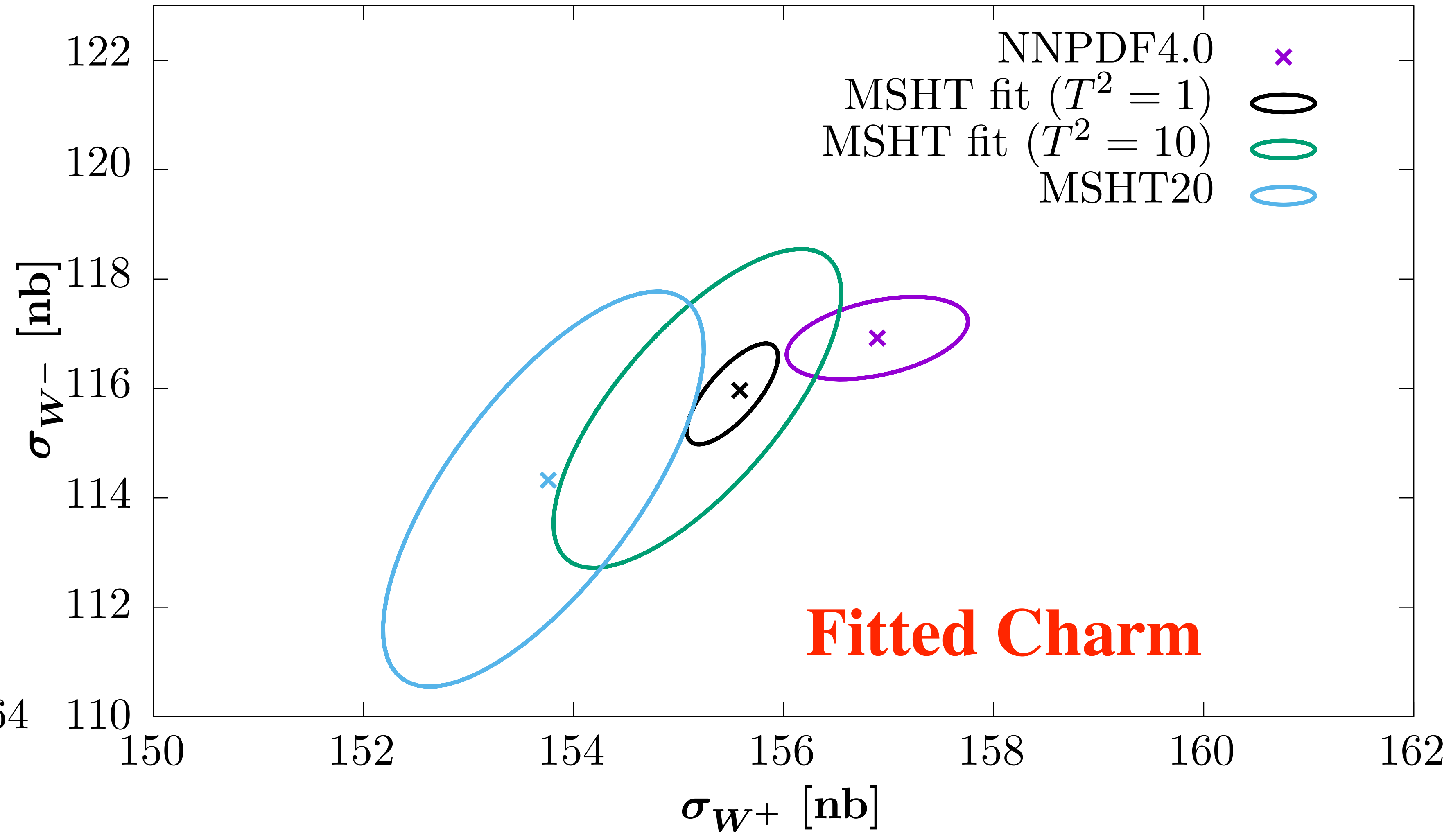
- Consider ggH, and W, Z cross sections (14 TeV): MSHT, $T^2 = 10$ MSHT $(T^2 = 1)$ NNPDF4.0

    - ★ NNPDF uncertainties ~ MSHT ($T^2 = 1$) but **significantly smaller** than $T^2 = 10$
    - ★ NNPDF and MSHT fit basically consistent within $T^2 = 1$ uncertainties but not relevant factor given fit qualities.
    - ★ Also shown is MSHT20 - impact of changing data/theory alone. Need for tolerance clear!









Restricted Parameterisations

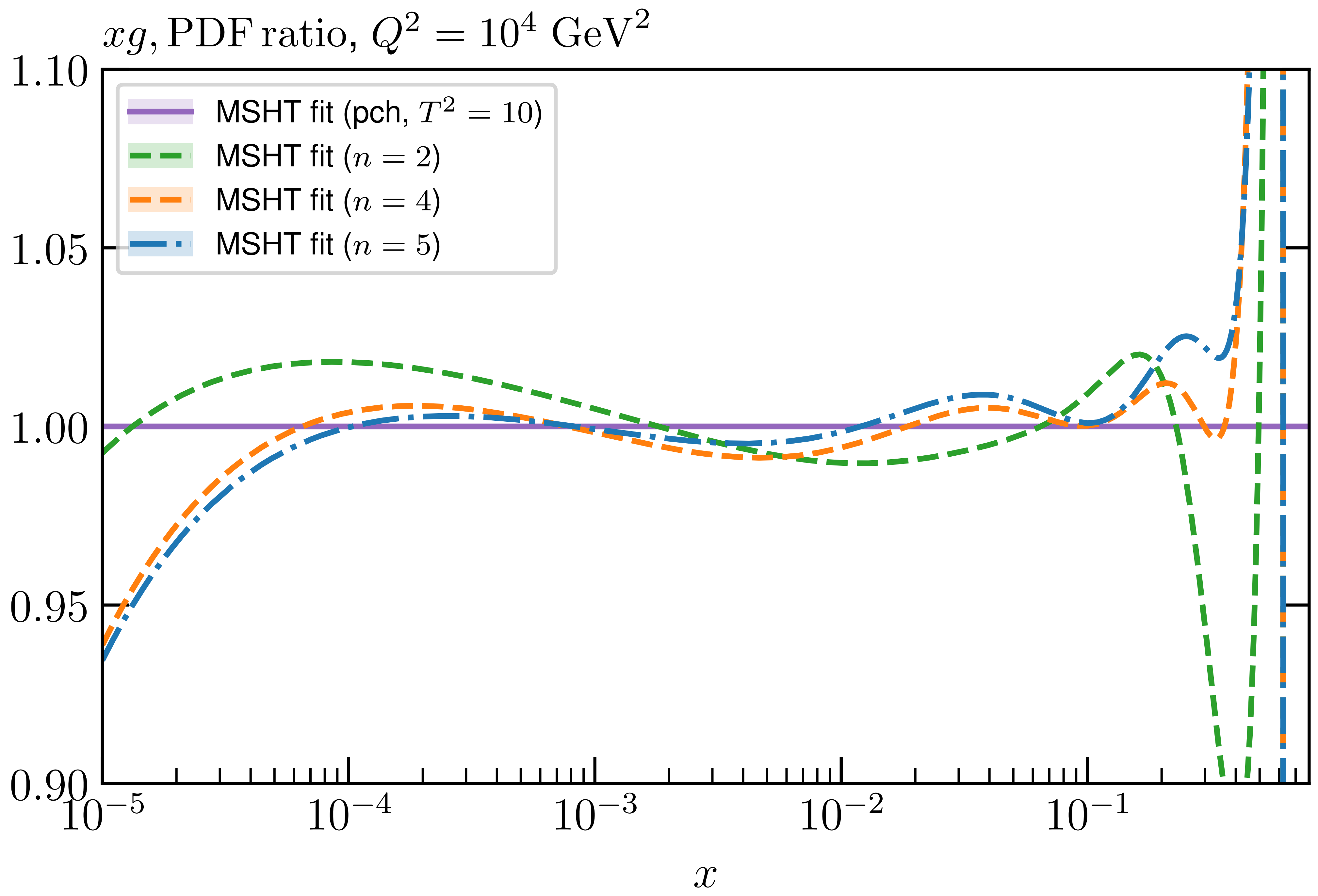
- As in closure tests, can examine impact of restricting parameterisation (in p. charm case).

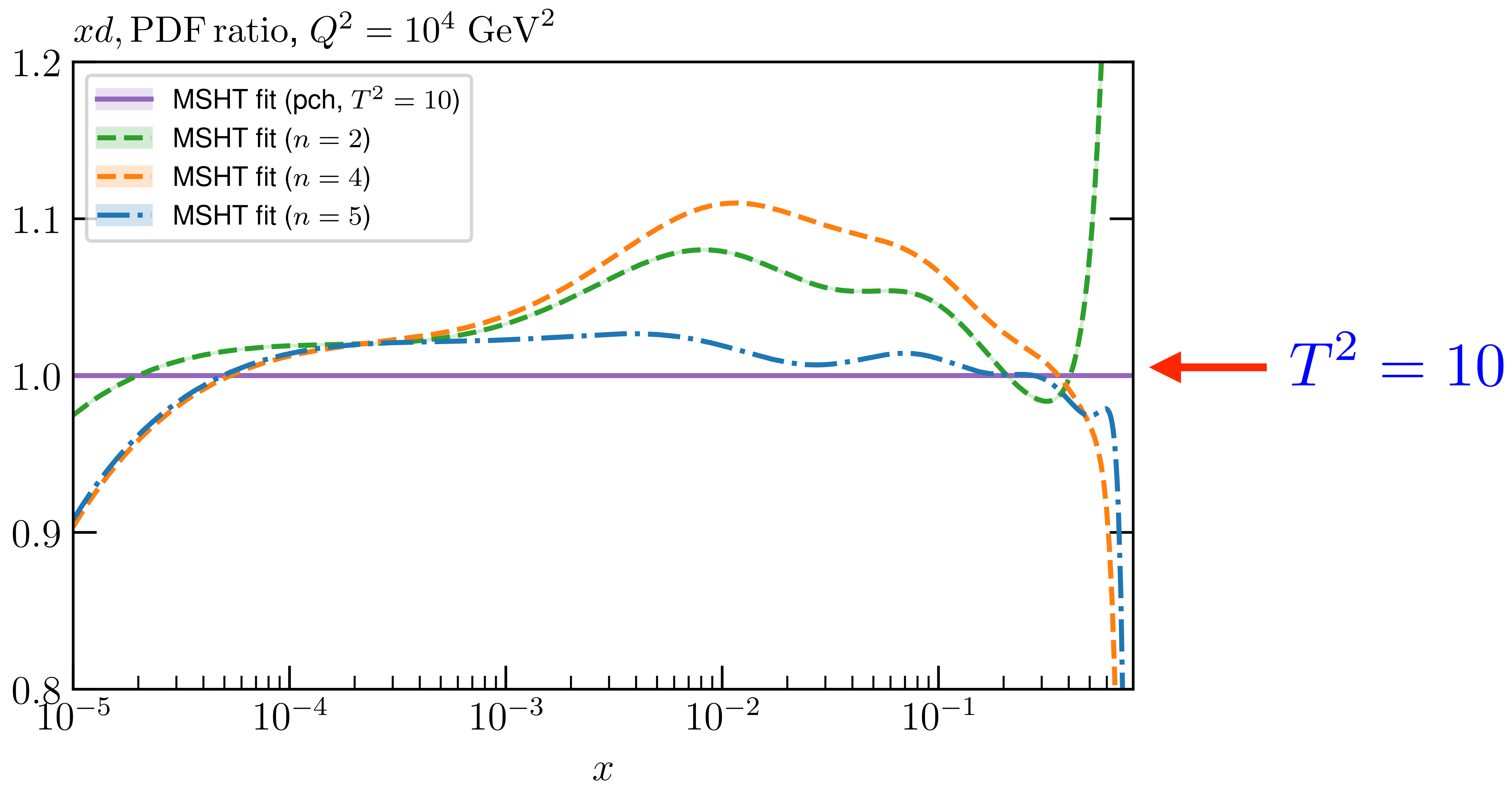
- Impact on fit quality slightly larger, e.g. ~ 200 point deterioration for $n = 2$ (~ 100 in closure)

$$xf(x, Q_0) = Ax^{\delta}(1-x)^{\eta}\left(1 + \sum_{i=1}^{n} a_i T_i(y(x))\right),$$

$n = 2, 4 \quad \Rightarrow \quad n_{\rm par} = 28, 40$

$n = 6 \quad \Rightarrow \quad n_{\rm par} = 52$

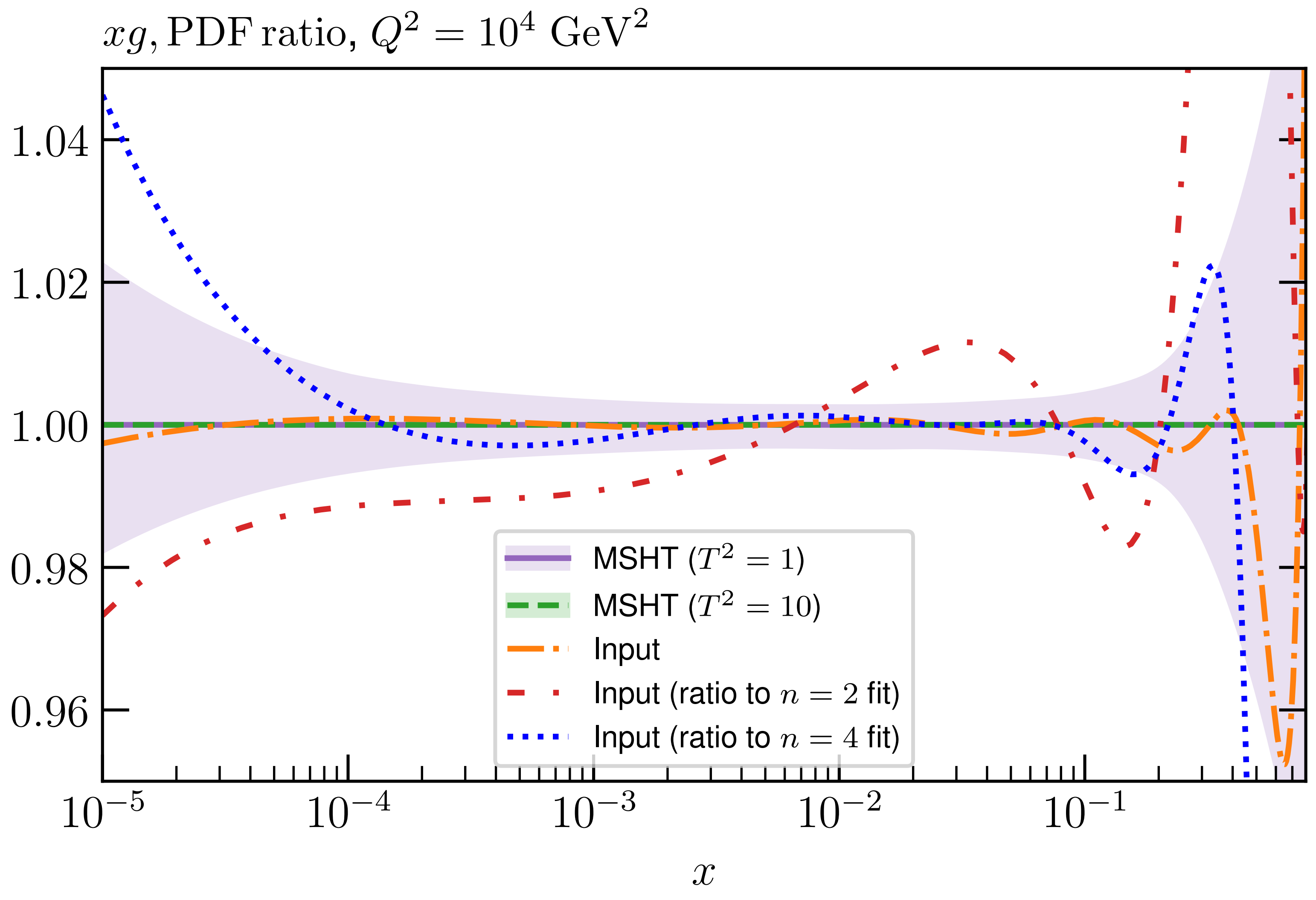




- Impact on PDFs also somewhat larger than in closure test - for $n = 2$ often at edge of even outside $T^2 = 10$ uncertainties.

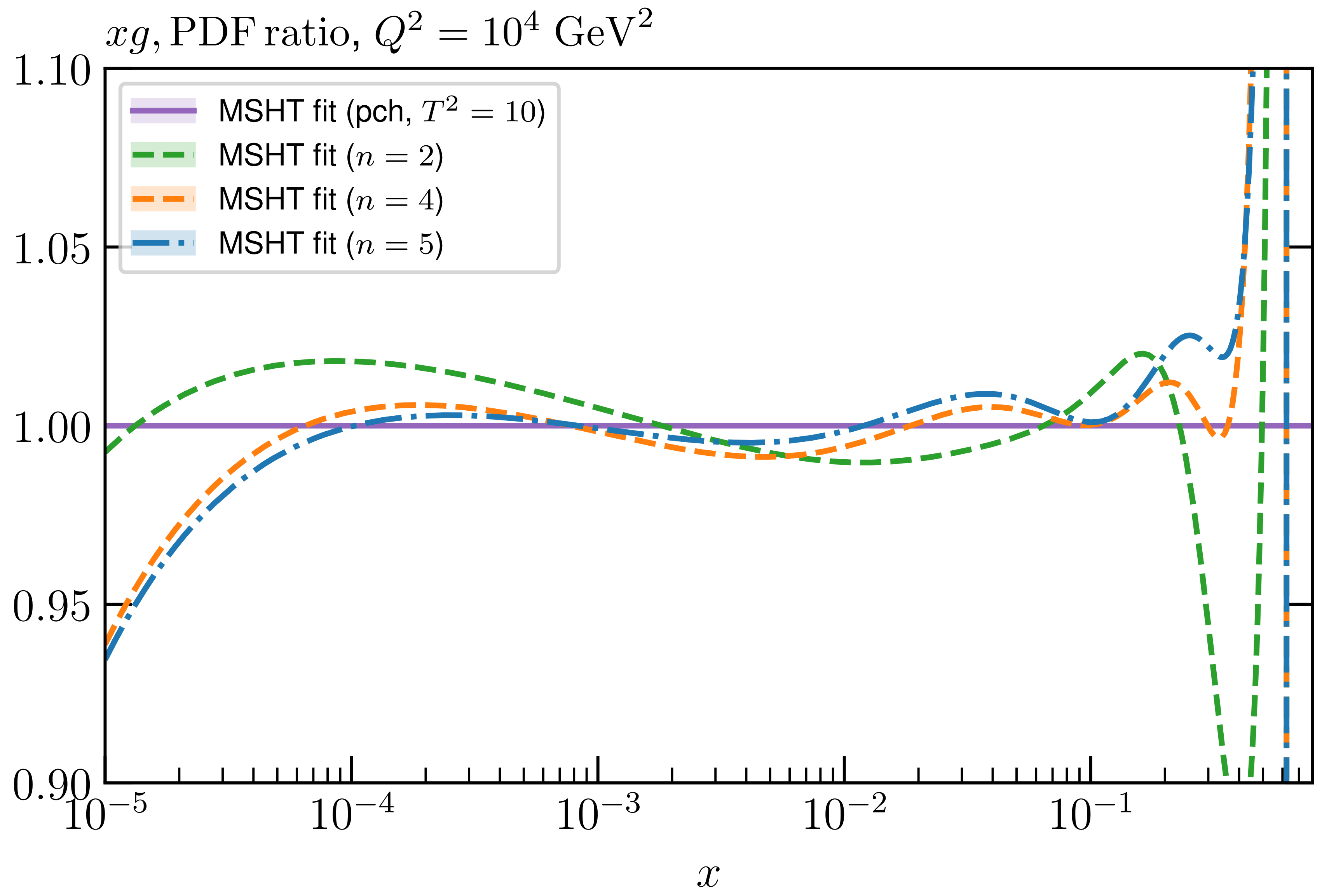
Overfitting

- Various tests performed to examine whether the default MSHT parameterisation ($n_{\rm par} = 52$) or more restricted ones results in overfitting, i.e. fitting noise in data.

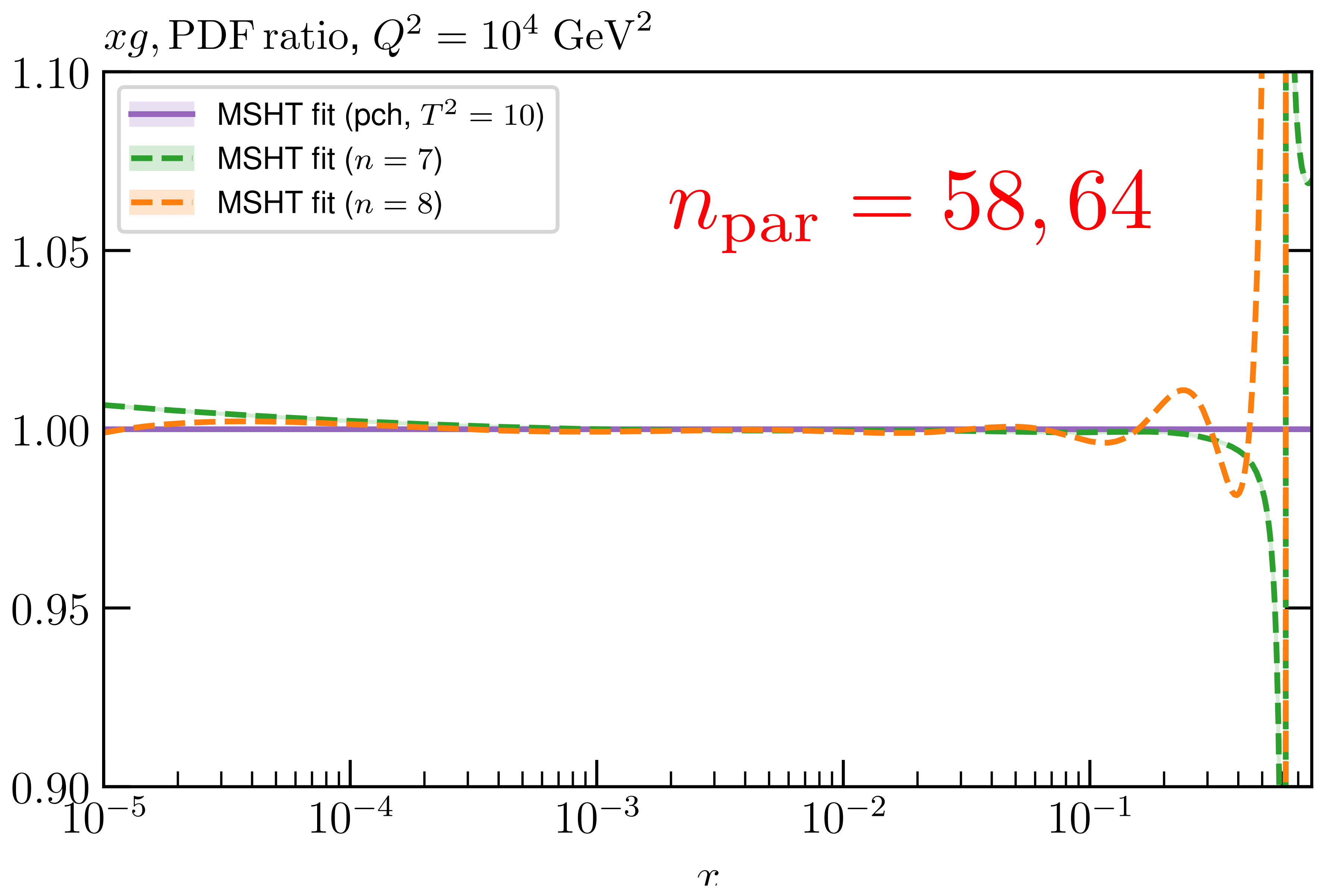


★ Closure test - restricted ($n_{\rm par} = 28, 40$) parameterisation cannot match pseuodata well. Unlikely to result in overfitting in real fit!

★ Indeed precisely same effect seen in real fit…



- Indeed can try **extending** parameterisation, and find very mild impact on fit quality + PDFs - i.e. no clear trend for overfitting setting in with this ~ of parameters.



- In NNPDF fit extensive training/validation used to prevent overfitting (in context of NN fit). Fitting same data/theory with MSHT parameterisation we find similar fit qualities.

	NNPDF4.0 pch	MSHT fit (w positivity)
$\chi^2_{t_0}$	5928.3 (1.282)	5837.8 (1.262)

$\Delta\chi^2_{t_0}$: $-90.5\,(-0.02)$

NNPDF4.0	MSHT fit (w positivity)
5692.1 (1.233)	5651.0 (1.224)

$\Delta\chi^2_{t_0}$: $-41.1\,(0.009)$

(After controlling for positivity)

- True that we find somewhat better fit quality (reason for this open question), but improvement is spread across range of e.g. LHC DY datasets that independently constrain similar PDFs.

Summary

- In this talk I have presented:

  ★ **First** global **closure test** of fixed parameterisation (MSHT) approach: is parameterisation flexible enough to give faithful description of global pseudodata?

  ✦ **Yes**: no issue in passing (unfluctuated) global closure test.

  ★ **First** completely direct **comparison** between fixed parameterisation (MSHT) and NN approaches. How do these compare in full global fit?

  ✦ At level of errors $\sigma(\text{NNPDF}) \sim \sigma(\text{MSHT}, T^2 = 1)$ in general with some exceptions - gluon larger though less than $T^2 = 10$ (MSHT20 default).

  ✦ At level of PDFs, surprisingly find fit quality is **lower** in **MSHT** fixed parameterisation case, and outside of NNPDF uncertainties. Reason for this is currently unclear. Positivity clearly important in p. charm case, but not only source (PDF basis? Flexibility?).

Summary

- Returning to the original possibilities (focus on MSHT as only considered here). We need to - and can - work out which is true:

    1. NNPDF4.0 uncertainty not conservative enough (too small).
    2. MSHT uncertainty too conservative (too large).
    3. MSHT fit less accurate, due to parameterisation inflexibility, and hence enlarged errors needed (less precise).

Summary

- Returning to the original possibilities (focus on MSHT as only considered here). We need to - and can - work out which is true:

    1. NNPDF4.0 uncertainty not conservative enough (too small).
    2. MSHT uncertainty too conservative (too large).
    3. ~~MSHT fit less accurate, due to parameterisation inflexibility, and hence enlarged errors needed (less precise).~~

- Successful closure test + comparison to NNPDF4.0 global fit suggests 3 is not dominant issue (at least in data region) for MSHT. Can be issue for less flexible ones and for MSHT in extrapolation regions.

Summary

- Returning to the original possibilities (focus on MSHT as only considered here). We need to - and can - work out which is true:

  1. NNPDF4.0 uncertainty not conservative enough (too small).
  2. MSHT uncertainty too conservative (too large).
  3. ~~MSHT fit less accurate, due to parameterisation inflexibility, and hence enlarged errors needed (less precise).~~

- Successful closure test + comparison to NNPDF4.0 global fit suggests 3 is not dominant issue (at least in data region) for MSHT. Can be issue for less flexible ones and for MSHT in extrapolation regions.
- First direct comparison to NNPDF4.0 global fit finds that this gives inherently different (smaller) uncertainties than MSHT fixed parameterisation, keeping everything else equal.

Summary

- Returning to the original possibilities (focus on MSHT as only considered here). We need to - and can - work out which is true:

    1. NNPDF4.0 uncertainty not conservative enough (too small).
    2. MSHT uncertainty too conservative (too large).
    3. ~~MSHT fit less accurate, due to parameterisation inflexibility, and hence enlarged errors needed (less precise).~~

- Successful closure test + comparison to NNPDF4.0 global fit suggests 3 is not dominant issue (at least in data region) for MSHT. Can be issue for less flexible ones and for MSHT in extrapolation regions.
- First direct comparison to NNPDF4.0 global fit finds that this gives inherently different (smaller) uncertainties than MSHT fixed parameterisation, keeping everything else equal.
- Put together, this implies that **either 1 or 2 is true** (or both). This study has not addressed which, though question of tolerance discussed elsewhere, but either way suggests more work needed.

Summary

- Have also examine impact of restricting PDF parameterisation to something with fewer free parameters than default MSHT case ($n_{\rm par} = 52 \to 28$).
- Closure test input **not reproduced** within $T^2 = 1$ (and even $T^2 = 10$) uncertainties.
- Ability to pass such a closure test is arguably necessary (though not sufficient) condition within any approach. Should be used to guide this.
- No obvious sign of overfitting within MSHT parameterisation, though question of why this gives better fit quality than NNPDF4.0 is open question (for further analysis - see paper to come!).

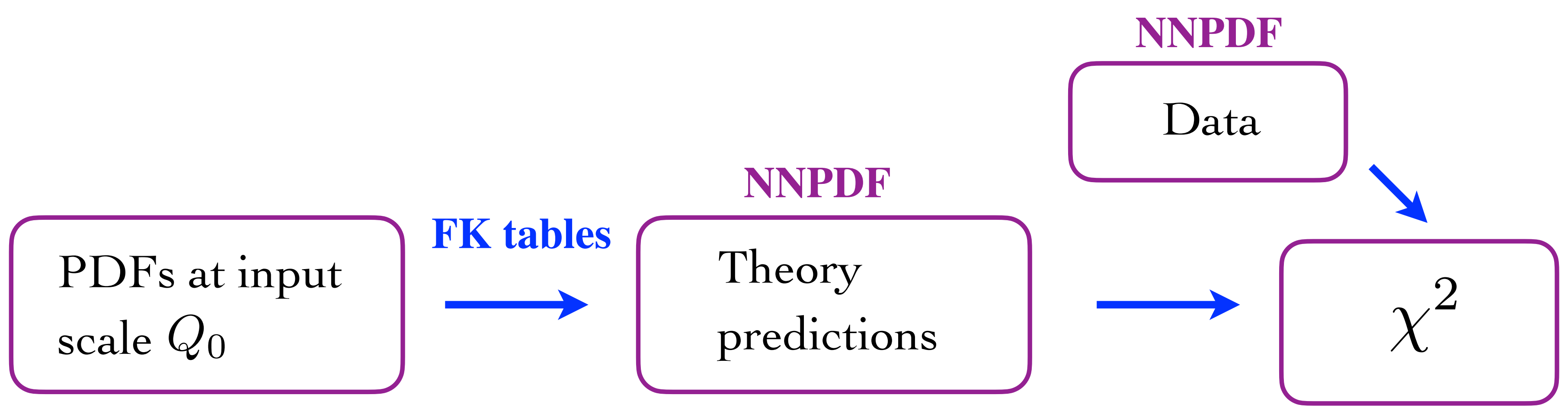
Thank you for listening!

Backup

Full fit with fitted charm

- Readily extend previous study to include **fitted charm**. NNPDF theory inputs change accordingly, while PDFs parametrised at $Q_0 = 1.65\,\mathrm{GeV}\ (> m_c)$ rather than 1 GeV, and parameterise charm:

$$xc_+(x,Q_0) = A_{c_+} x^{\delta_{c+}} (1-x)^{\eta_{c+}} \left(1 + \sum_{i=1}^{6} a_{c,i} T_i(y(x))\right)$$
$$xc_-(x,Q_0) = 0$$



- Find:

	NNPDF4.0	MSHT fit	MSHT fit (w positivity)
$\chi^2_{t_0}$	5692.1 (1.233)	5645.2 (1.222)	5651.0 (1.224)

$\Delta\chi^2_{t_0}$: $-46.9\ (0.011)$ $-41.1\ (0.009)$

⟶ Fit quality with **MSHT** parameterisation again **better** than result of central NNPDF set, albeit by less than in perturbative charm case.

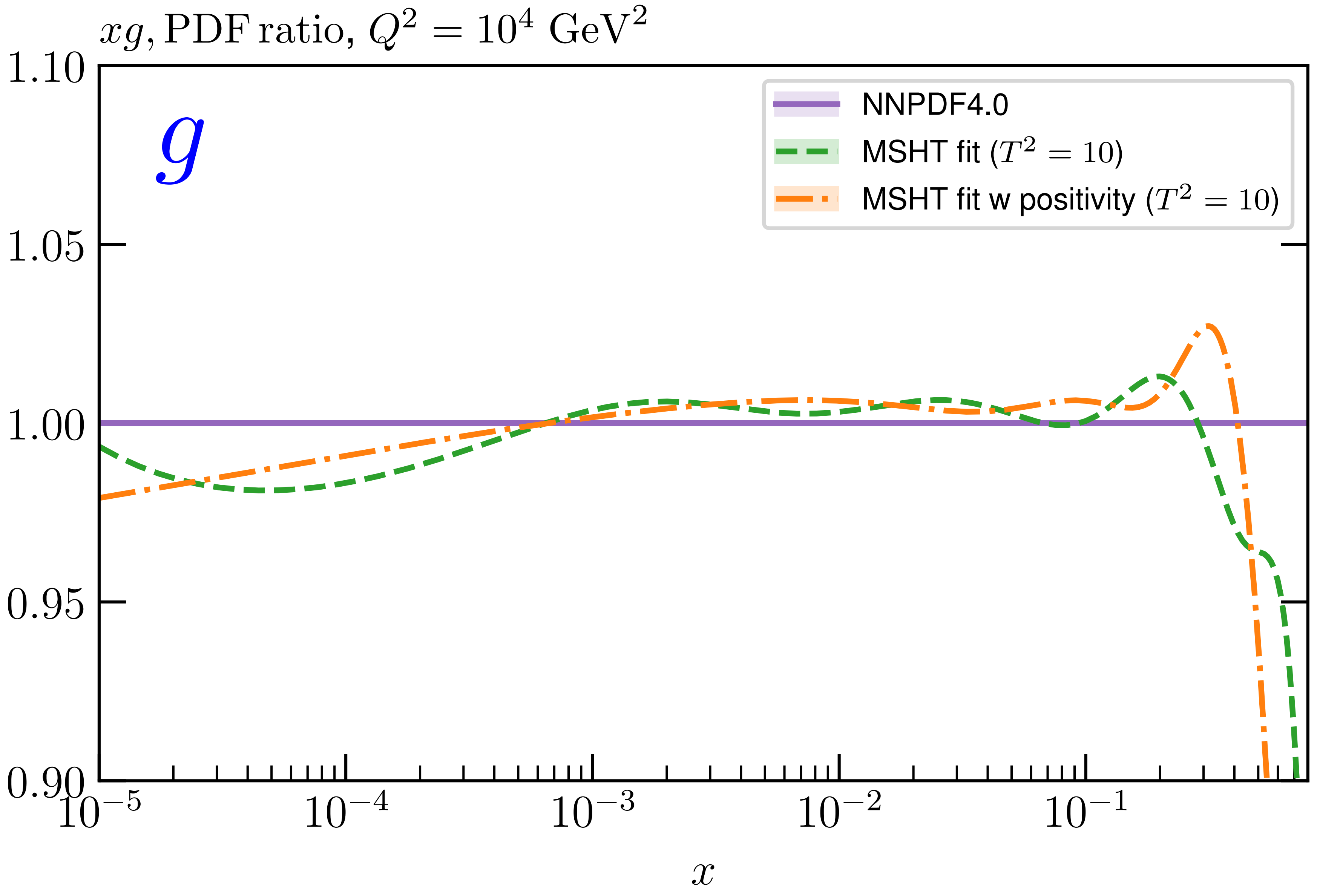
- Improvement spread across fixed target and LHC DY data.
- Role of **positivity** now **marginal** (confirmed with direct NNPDF fit).

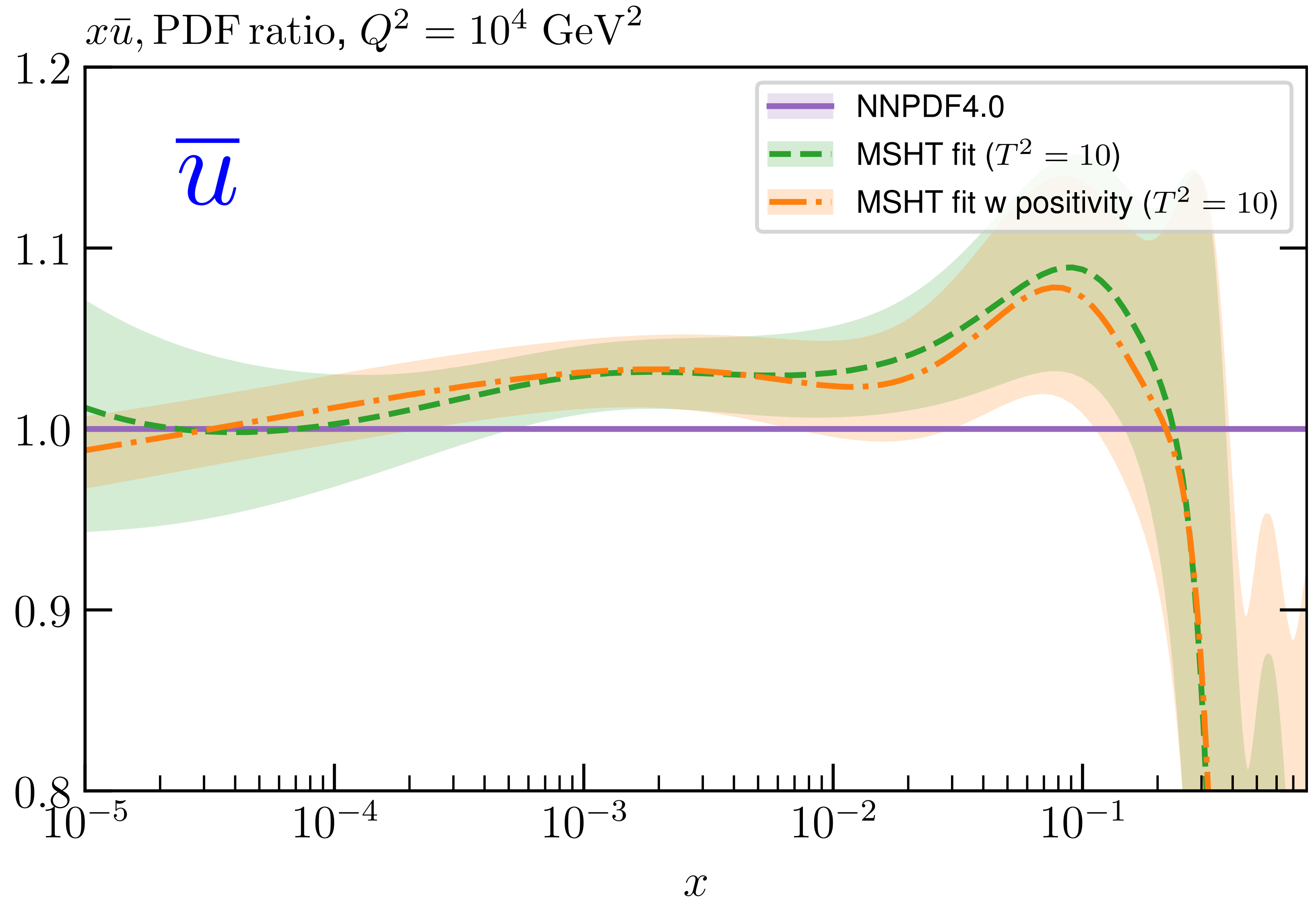
Integrability checked and not issue: Backup

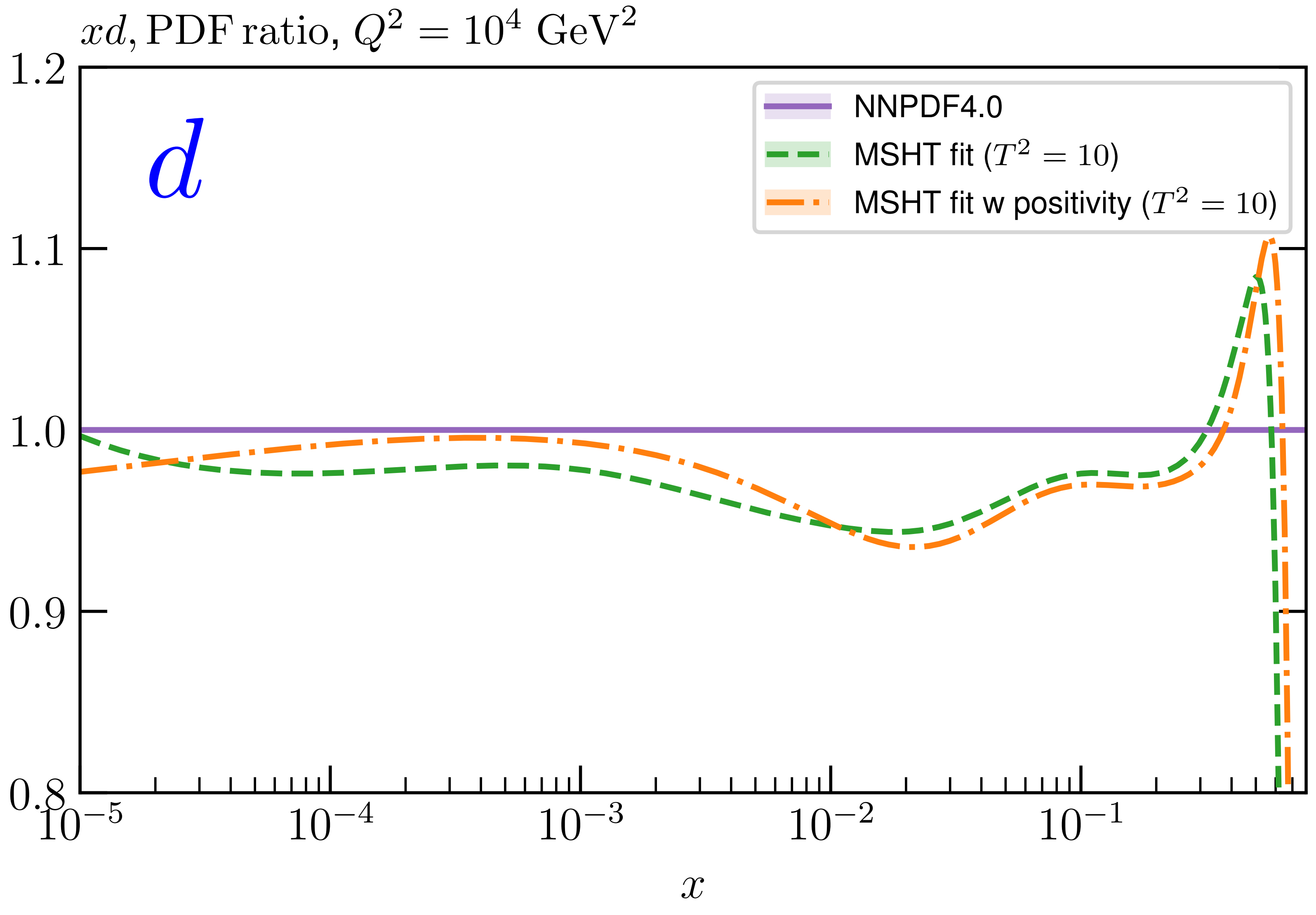
PDFs

- Confirm that role of **positivity** now **marginal**. Also seen if positivity removed from pure NNPDF fit.
- Difference less than in p. charm case, but **clear difference** in **flavour decomposition** remains.
- Again qualitatively this follows trend of using flavour rather than evolution basis (though difference larger and not identical). **See Backup**
- Show $T^2 = 10$ for concreteness and fact that there is reasonable (not perfect) agreement within these not relevant factor, given fit quality is better for MSHT sets.

NNPDF4.0pch MSHT (pos) $T^2 = 10$ MSHT, $T^2 = 10$

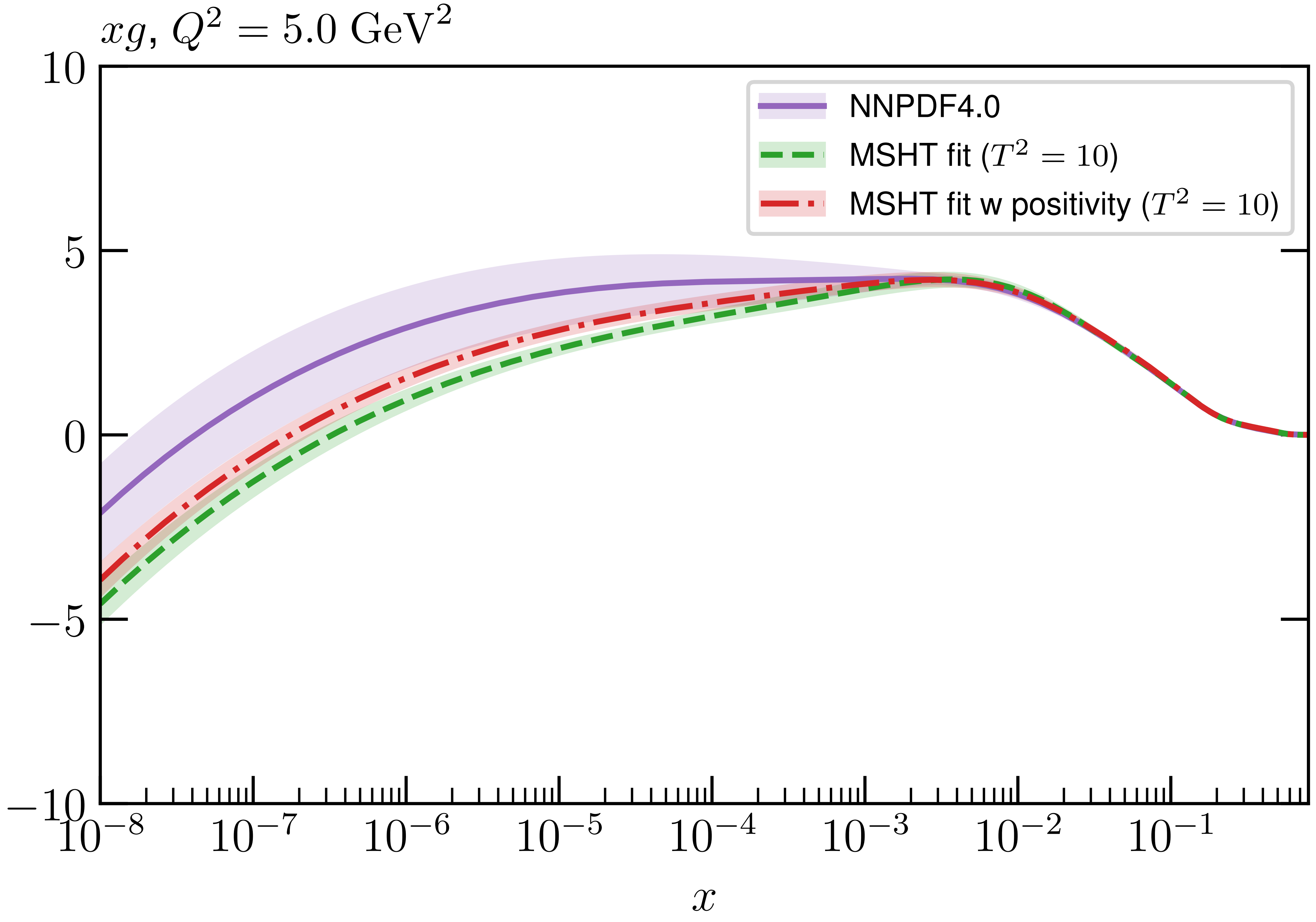






Positivity?

- With fitted charm tendency for preferred gluon to be negative reduced and pushed to lower x : why positivity requirement plays minor role - also seen in direct NNPDF fit.
- Not clear why this is (under investigation) but given intrinsic charm is expected to be high x phenomena might be concern?



Comparison to NNPDF uncertainties

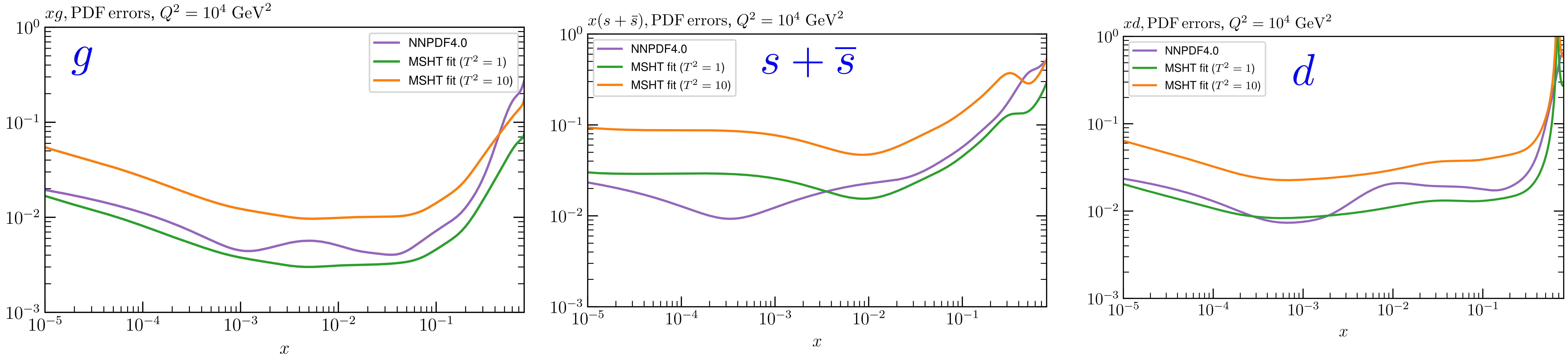
- Can do same comparison of MSHT vs. NNPDF PDF uncertainties but now with fitted charm. Again completely like-for-like. Results very similar to closure test comparison:

  ⋆ Quark flavour decomposition: $\sigma(\text{NNPDF}) \sim \sigma(\text{MSHT}, T^2 = 1)$

  ⋆ Gluon (singlet at intermediate x): $\sigma(\text{MSHT}, T^2 = 1) \lesssim \sigma(\text{NNPDF}) \lesssim \sigma(\text{MSHT}, T^2 = 10)$

MSHT, $T^2 = 1$
MSHT, $T^2 = 10$
NNPDF4.0

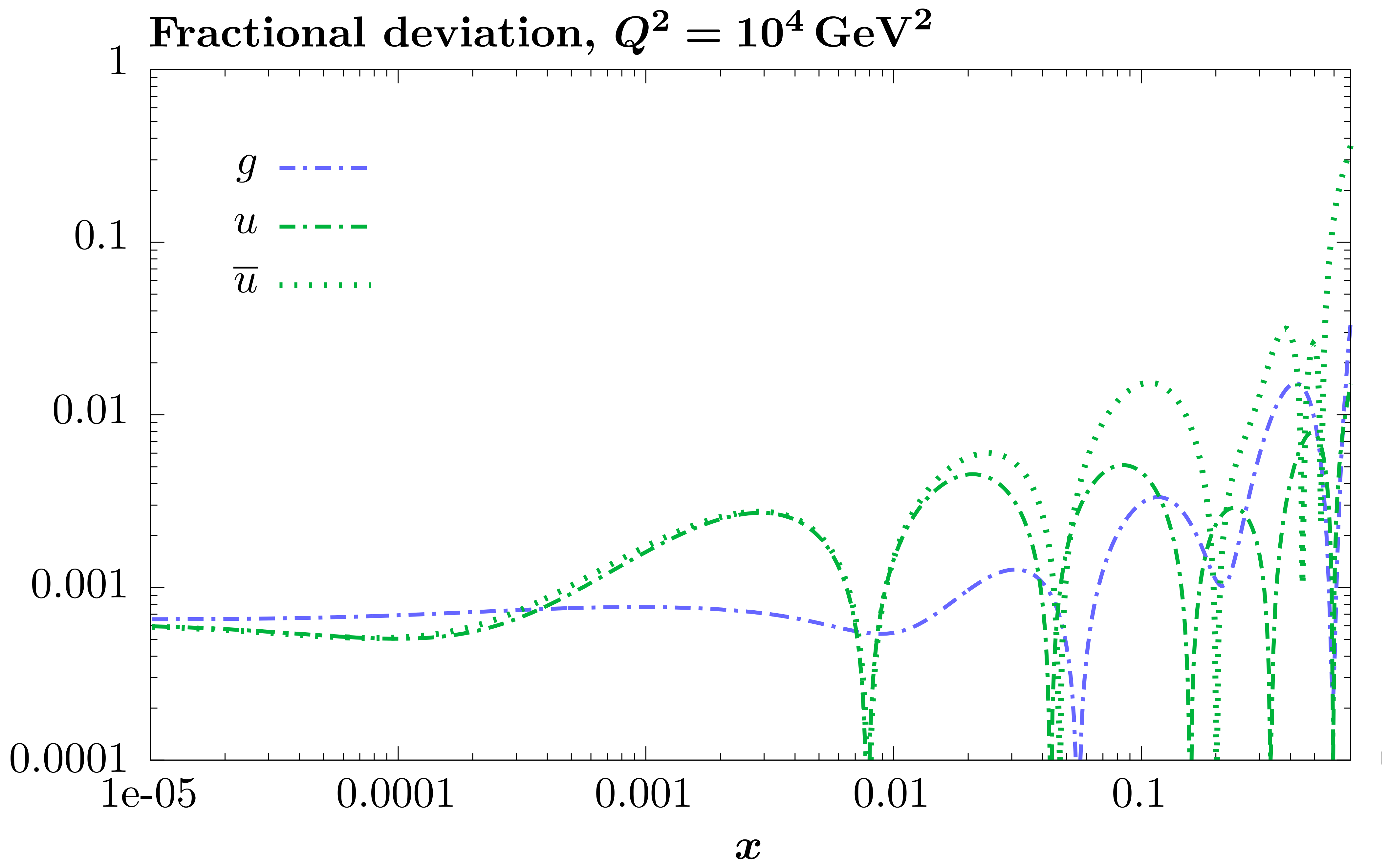
- With rather similar overall trends with x .
- Exception at high x where NNPDF uncertainy can become larger.
- Some trend for gluon to be a little closer to $T^2 = 1$ case.

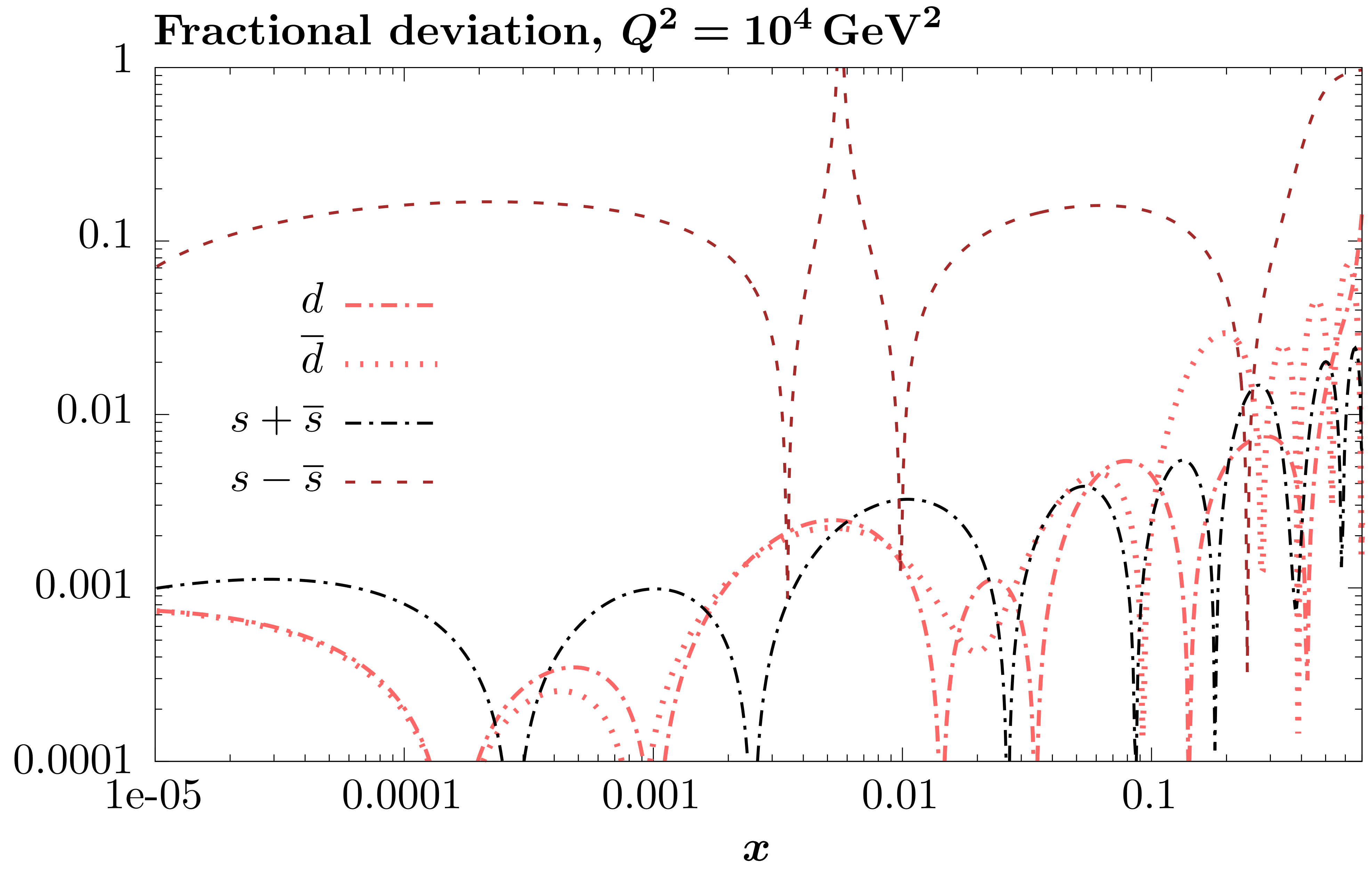


Closure test - warm up

Similar study: A.D. Martin et al., arXiv:1211.1215

- Before considering global closure test, consider related question. Given PDF-level pseudodata, how closely can MSHT parameterisation match it Basic point: for LHC precision aim for sub-1% agreement.
- 500 PDF points logarithmically in $x \in \{10^{-5}, 0.99\}$ scattered by 1% uncertainty, for $u_V, d_V, S, s_+, s_-, g, \bar{d}/\bar{u}$
- Take NNPDF4.0 (p. charm) as input and plot fractional deviation. Find this is $\ll 0.01$ for most of the x region. Biggest deviations at high x and for s_- (as expected - MSHT parameterisation limited at moment).





- Encouraging, but rather artificial - really want to see how deviation compares in data region of global fit.

MSHT parameterisation

$$u_V(x,Q_0^2) = A_u(1-x)^{\eta_u} x^{\delta_u}\left(1+\sum_{i=1}^{6} a_{u,i}T_i(y(x))\right)$$

$$d_V(x,Q_0^2) = A_d(1-x)^{\eta_d} x^{\delta_d}\left(1+\sum_{i=1}^{6} a_{d,i}T_i(y(x))\right)$$

$$S(x,Q_0^2) = A_S(1-x)^{\eta_S} x^{\delta_S}\left(1+\sum_{i=1}^{6} a_{S,i}T_i(y(x))\right)$$

$$s_+(x,Q_0^2) = A_{s_+}(1-x)^{\eta_{s+}} x^{\delta_S}\left(1+\sum_{i=1}^{6} a_{s_+,i}T_i(y(x))\right)$$

$$g(x,Q_0^2) = A_g(1-x)^{\eta_g} x^{\delta_g}\left(1+\sum_{i=1}^{4} a_{g,i}T_i(y(x))\right) + A_{g-}(1-x)^{\eta_{g-}} x^{\delta_{g-}}$$

$$s_-(x,Q_0^2) = A_{s_-}(1-x)^{\eta_{s-}}(1-x/x_0)x^{\delta_{s-}}$$

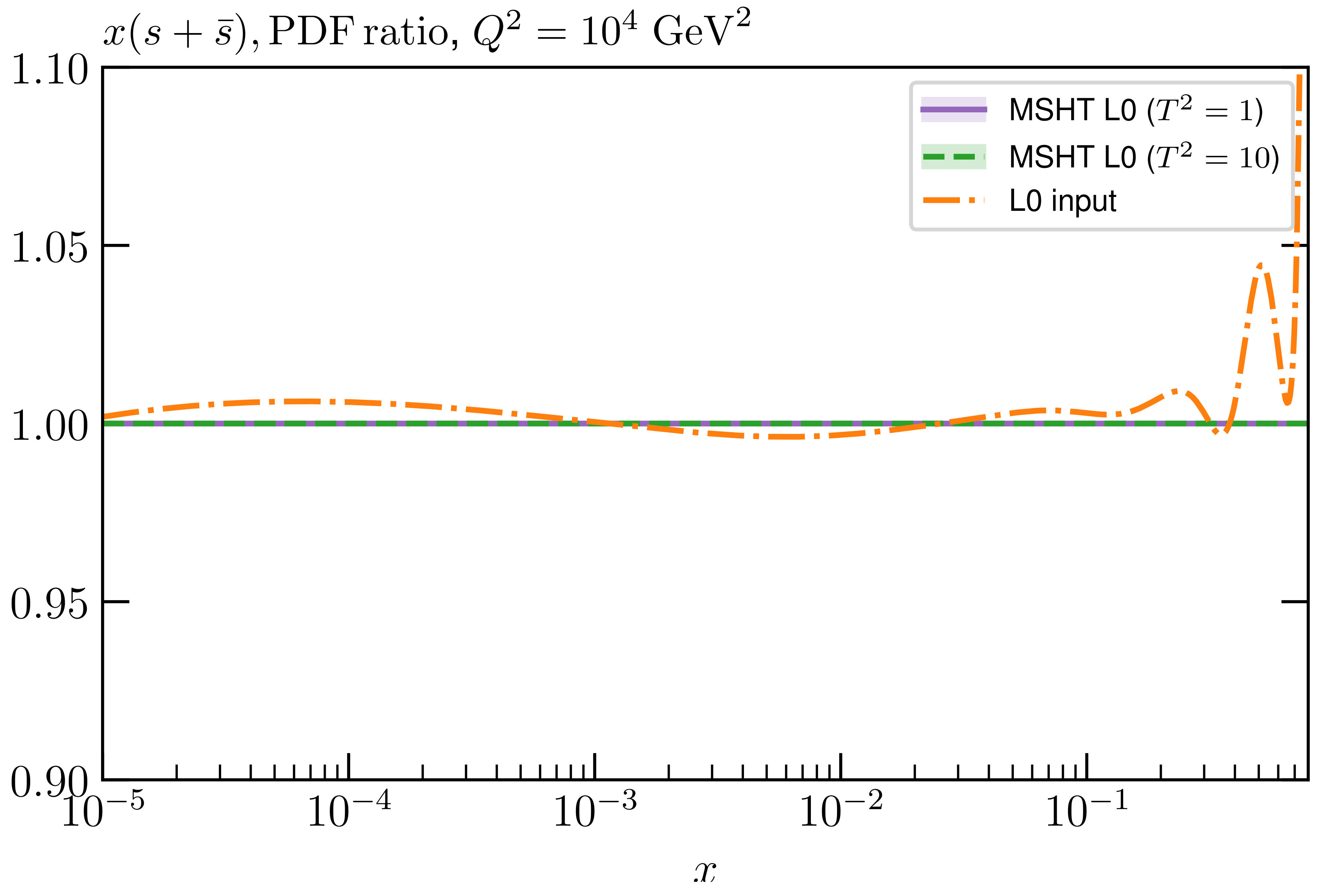
$$(\bar{d}/\bar{u})(x,Q_0^2) = A_\rho(1-x)^{\eta_\rho}\left(1+\sum_{i=1}^{6} a_{\rho,i}T_i(y(x))\right)$$

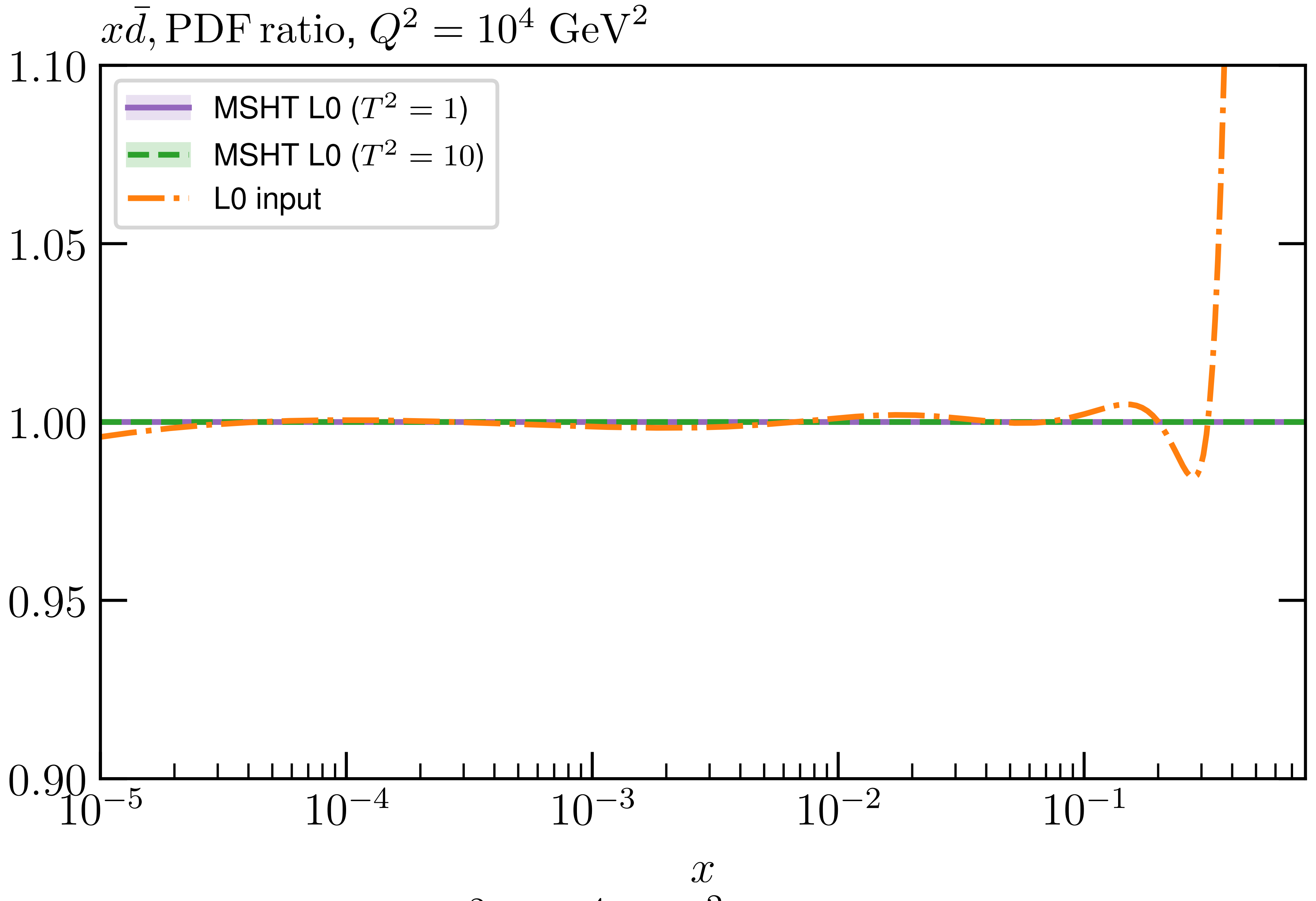
$$y(x) = 1-2\sqrt{x}$$

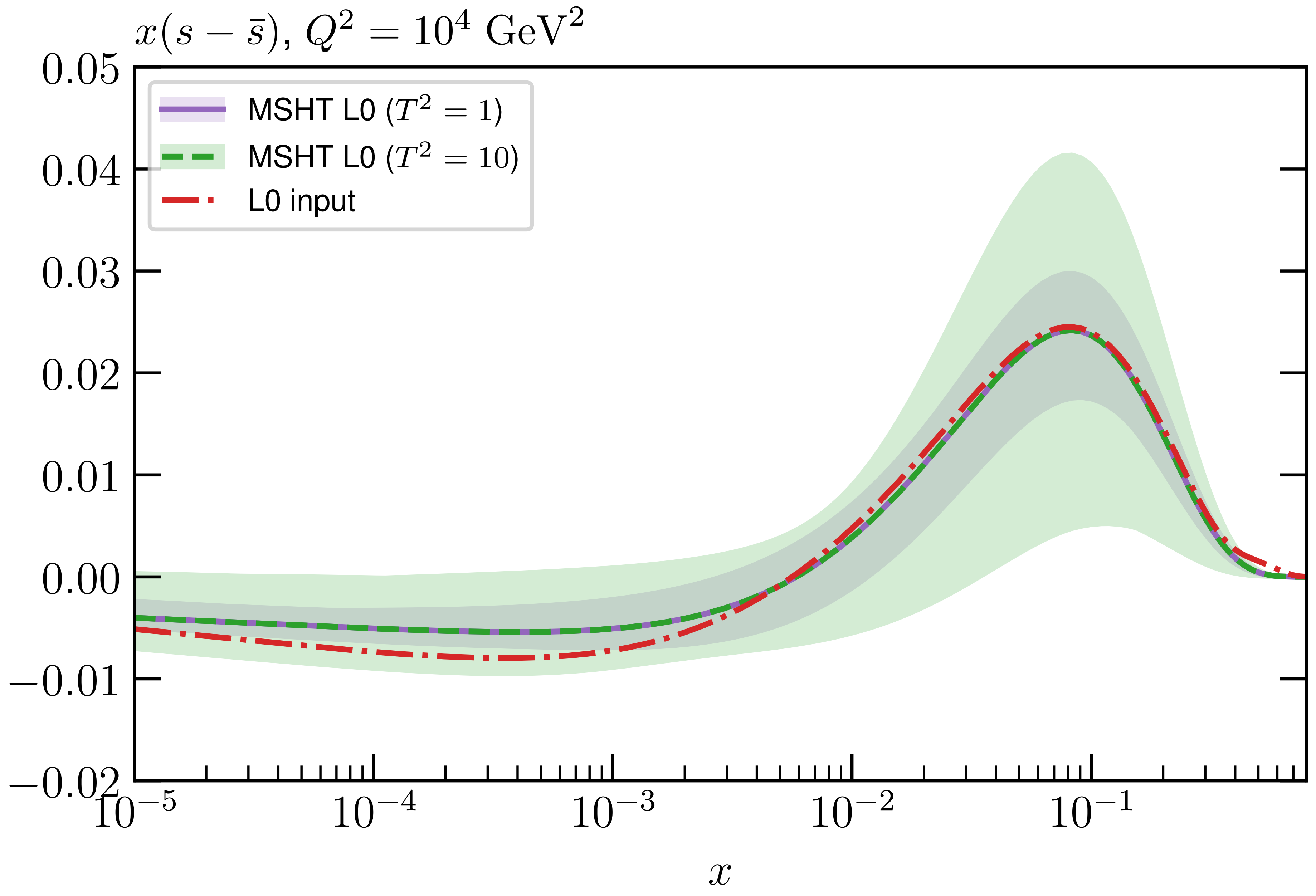
$$S(x) = 2(\bar{u}(x)+\bar{d}(x))+s(x)+\bar{s}(x)$$

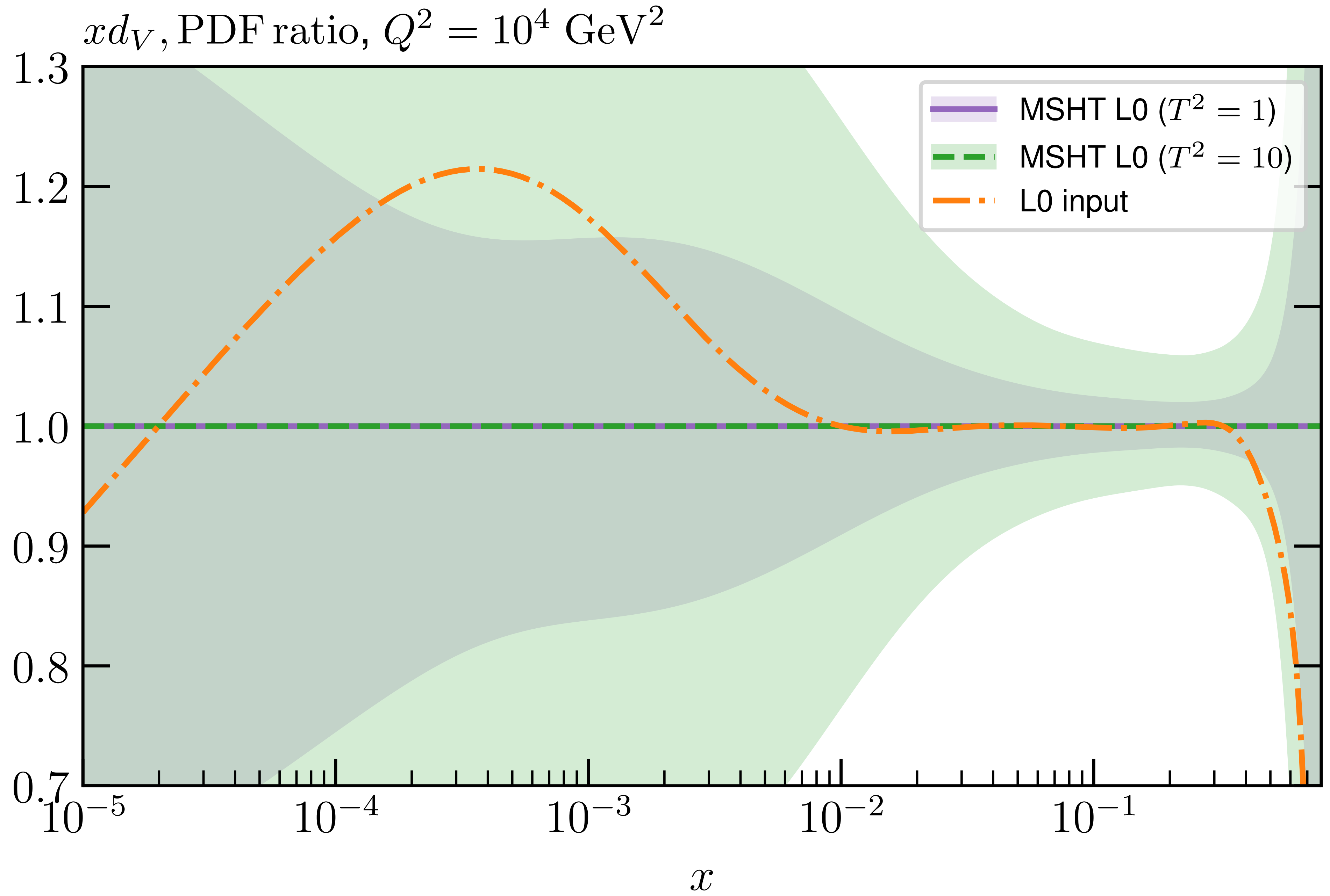
T_i: Chebyshev Polynomials

Global Closure PDFs



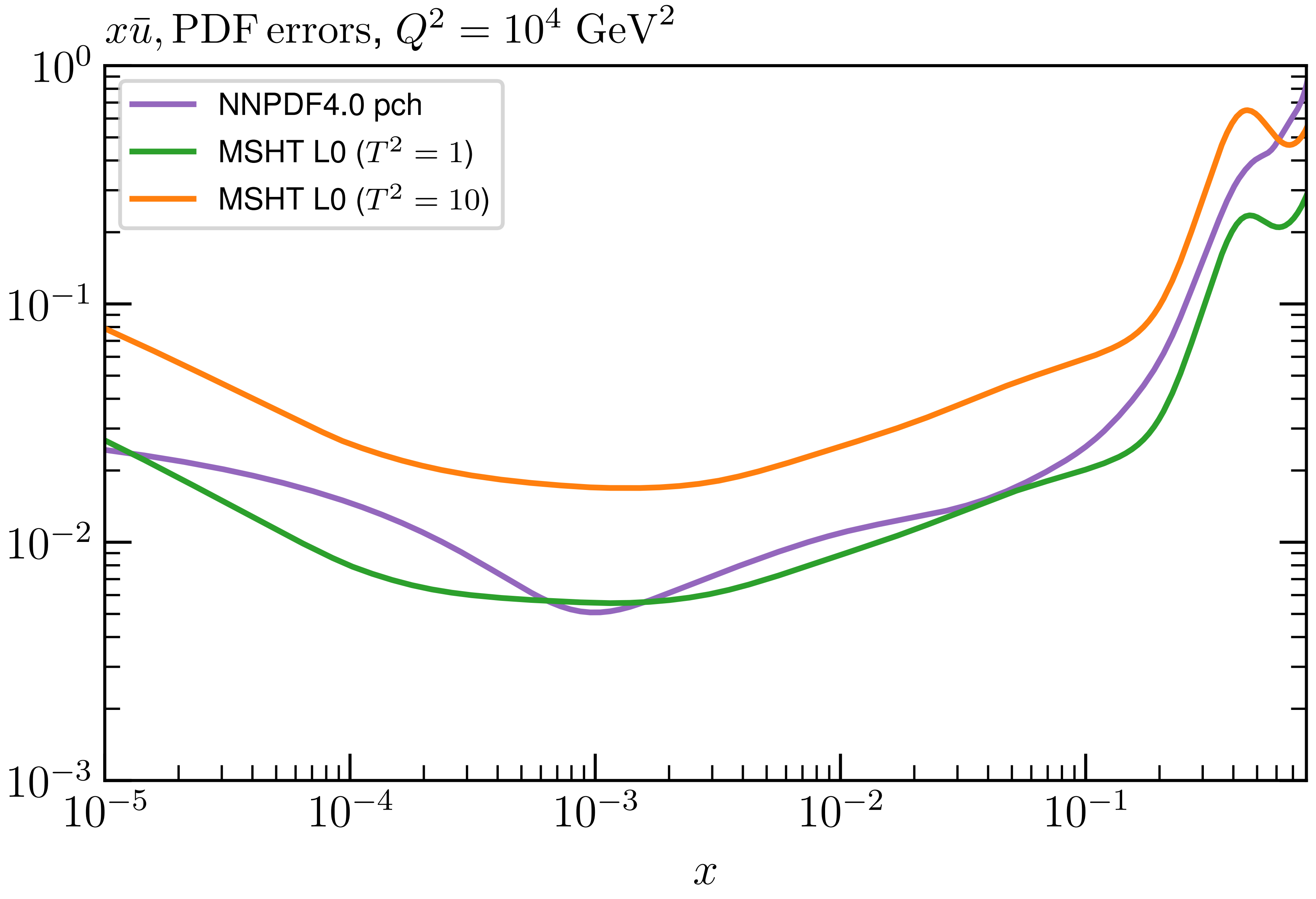


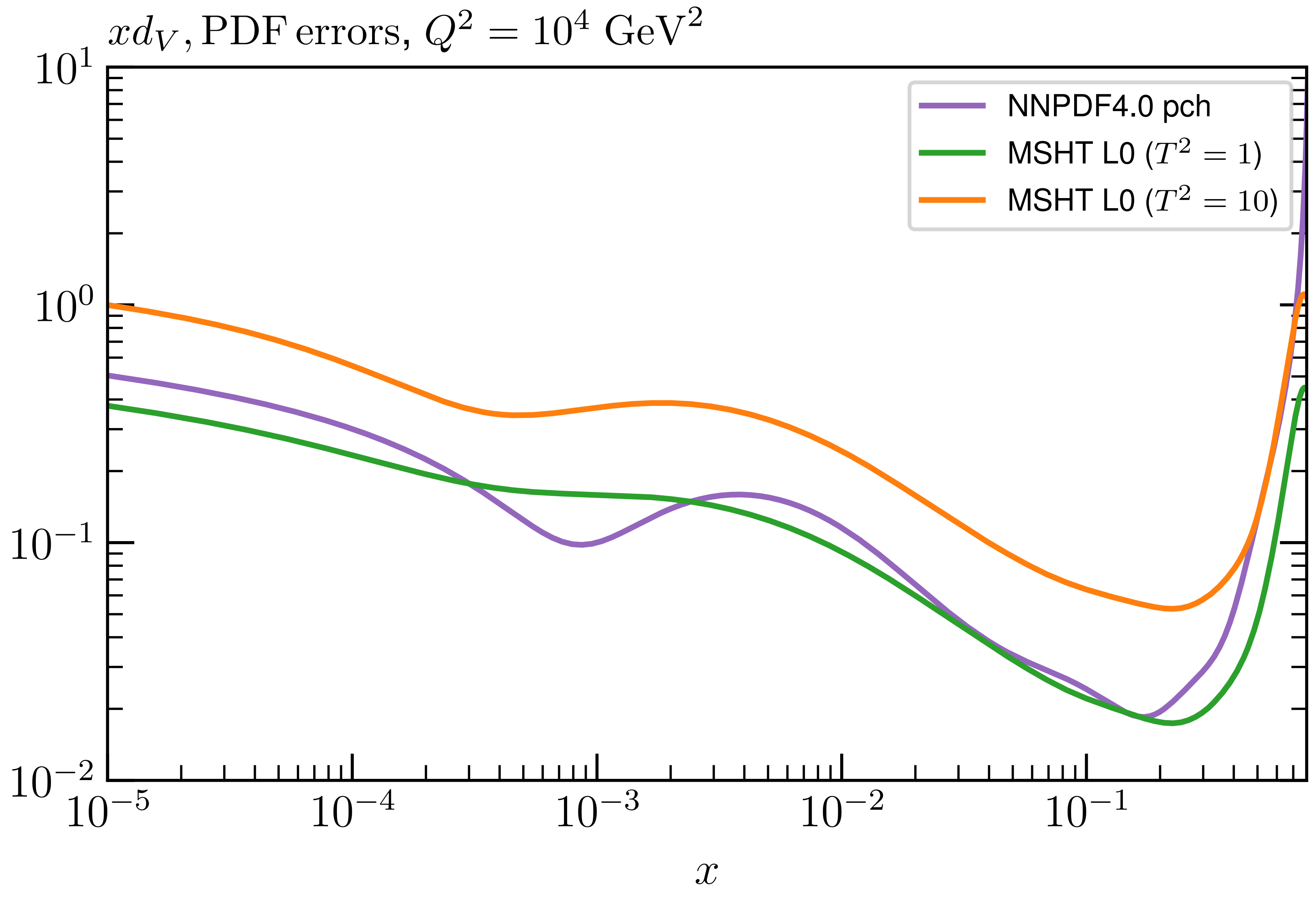


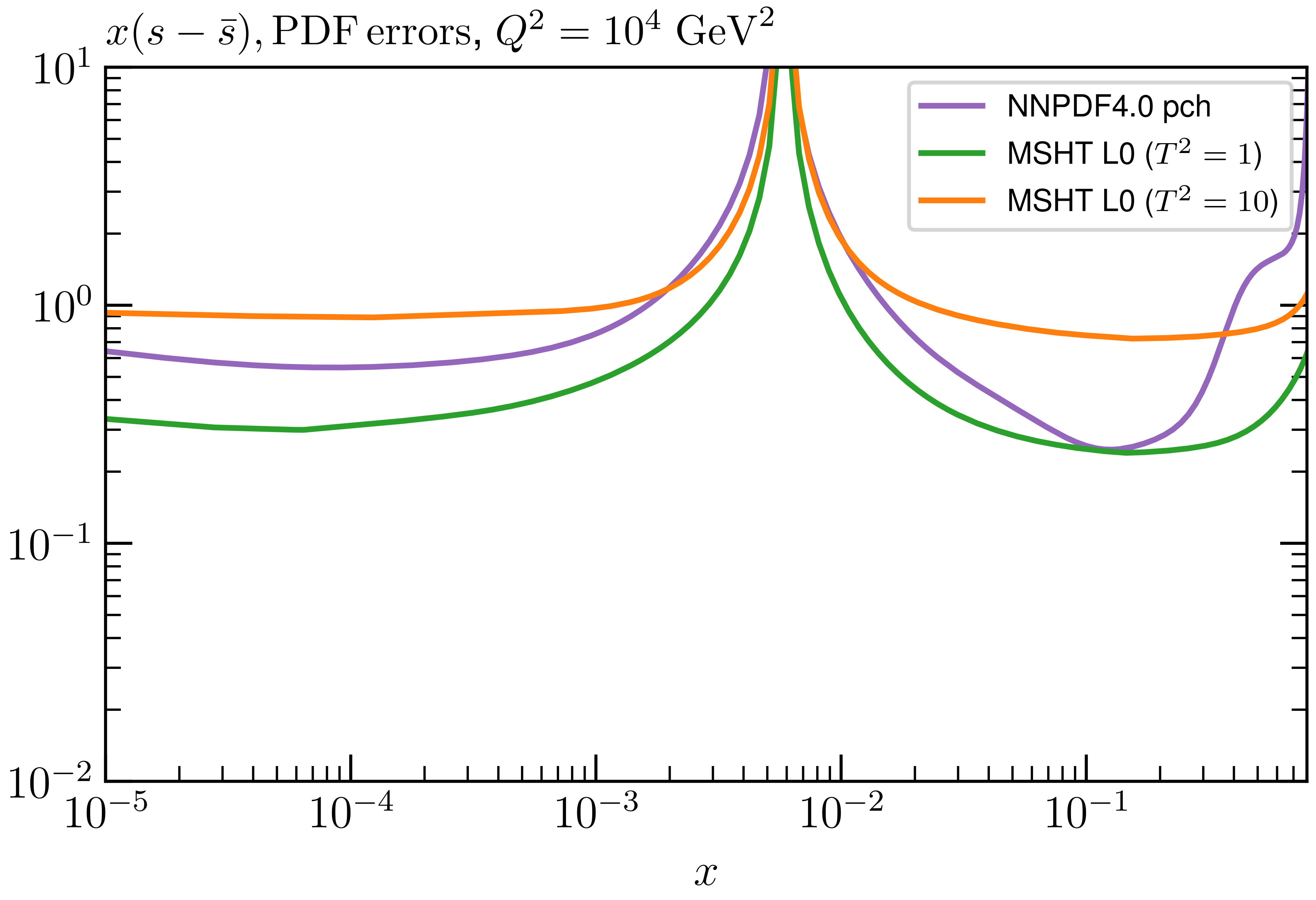


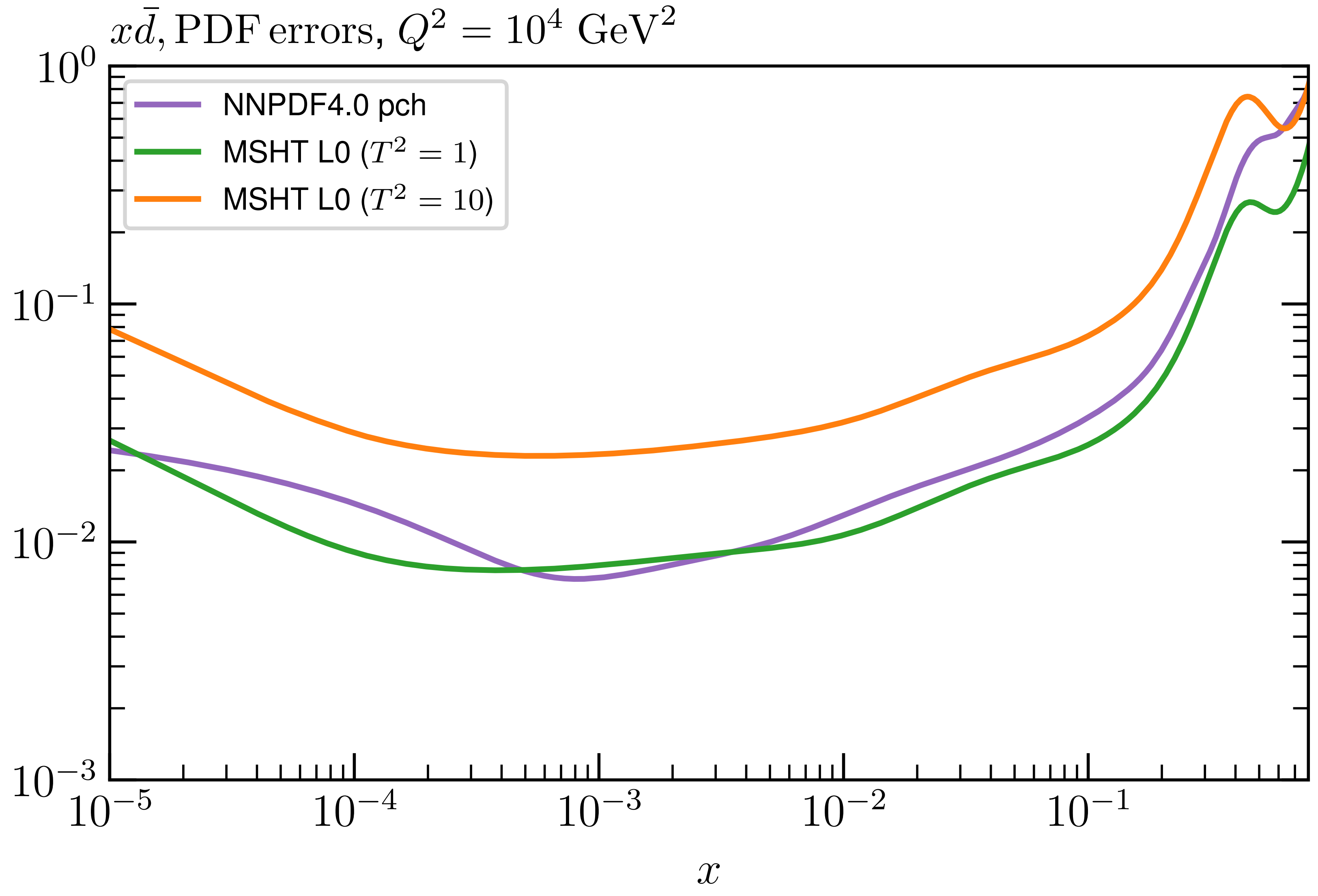

xd, $Q^2 = 10^4$ GeV^2
xu, $Q^2 = 10^4$ GeV^2
$x\bar{d}$, $Q^2 = 10^4$ GeV^2
$x(s+\bar{s})$, $Q^2 = 10^4$ GeV^2
xg, $Q^2 = 10^4$ GeV^2
MSHT L0 ($T^2 = 1$)
MSHT L0 ($T^2 = 10$)
L0 input
x

Comparison to NNPDF errors - Closure









Positivity

- We take:

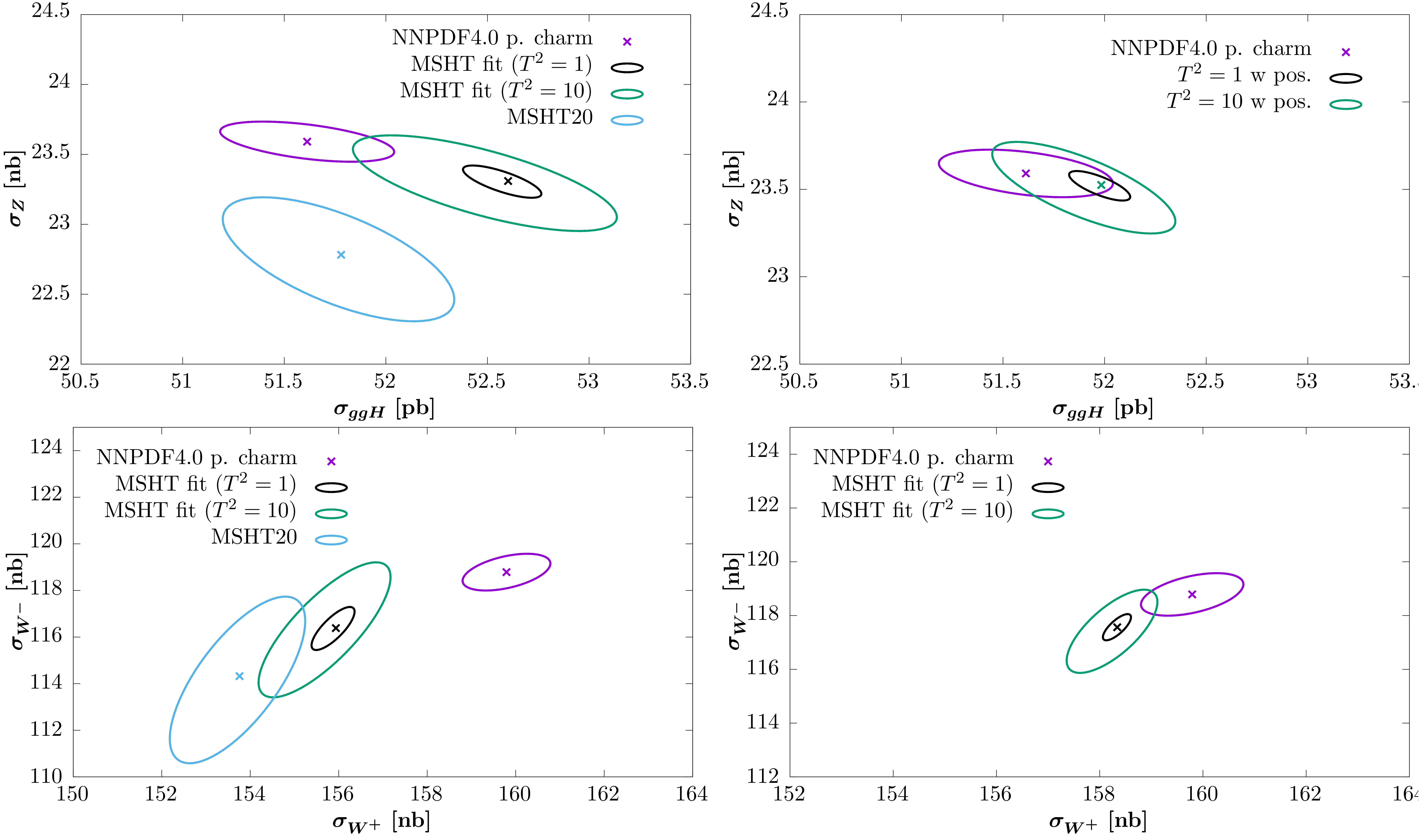
$$\chi^2_{\rm tot} \to \chi^2_{\rm tot} + \sum_{k=1}^{8} \Lambda_k \sum_{i=1}^{n_i} {\rm Elu}_\alpha \left(-\tilde{f}_k \left(x_i, Q^2 \right) \right) \,, \qquad \text{with} \qquad \begin{array}{c} \Lambda_k = 10^3 \\ x_i \in \{5 \cdot 10^{-7}, 0.9\} \end{array}$$

See NNPDF, arXiv:2109.02653

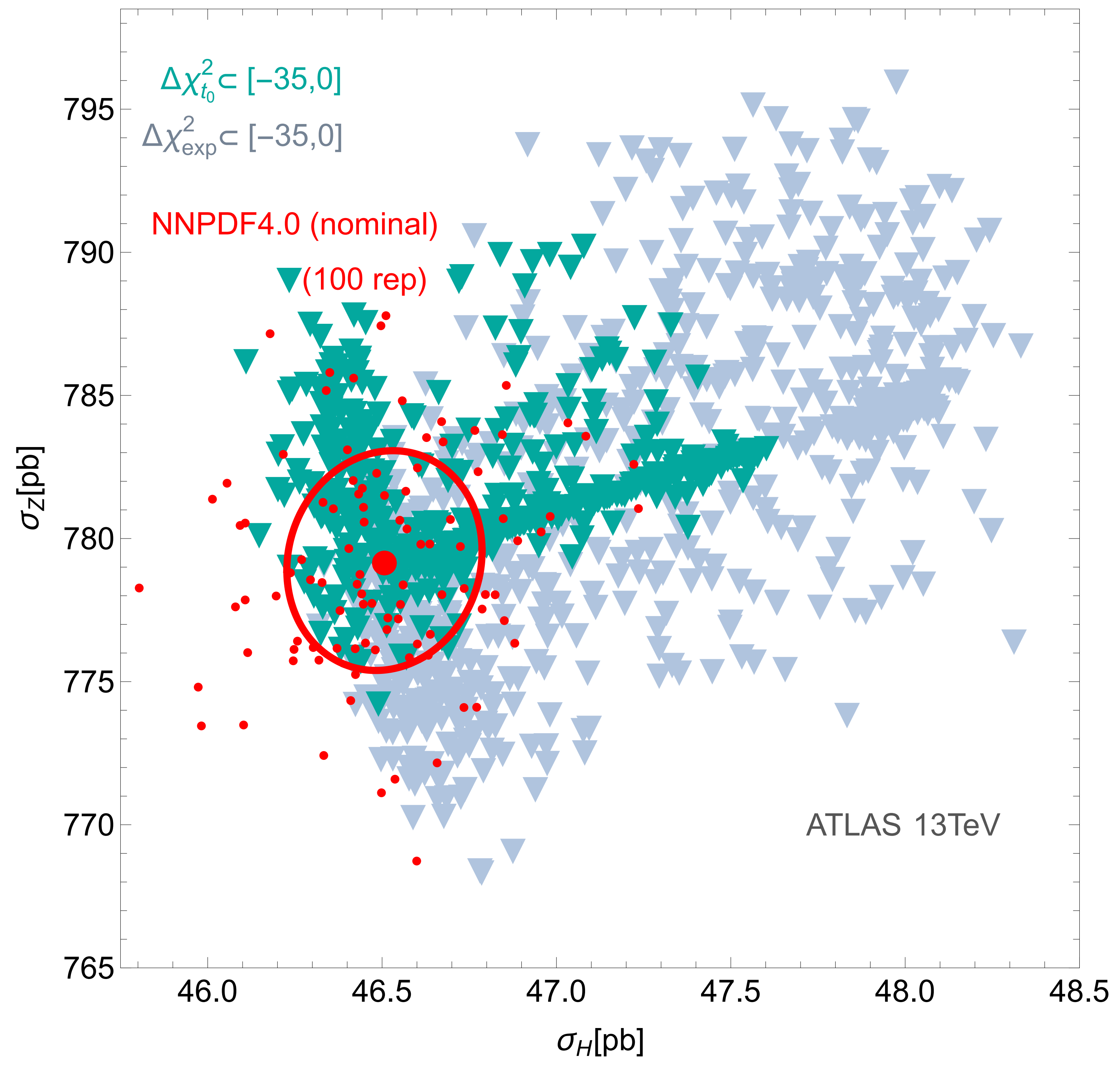
$${\rm Elu}_\alpha (t) = \begin{cases} t & \text{if} \quad t > 0 \\ \alpha \left(e^t - 1 \right) & \text{if} \quad t < 0 \end{cases} \,,$$

- And similarly for cross section constraints.

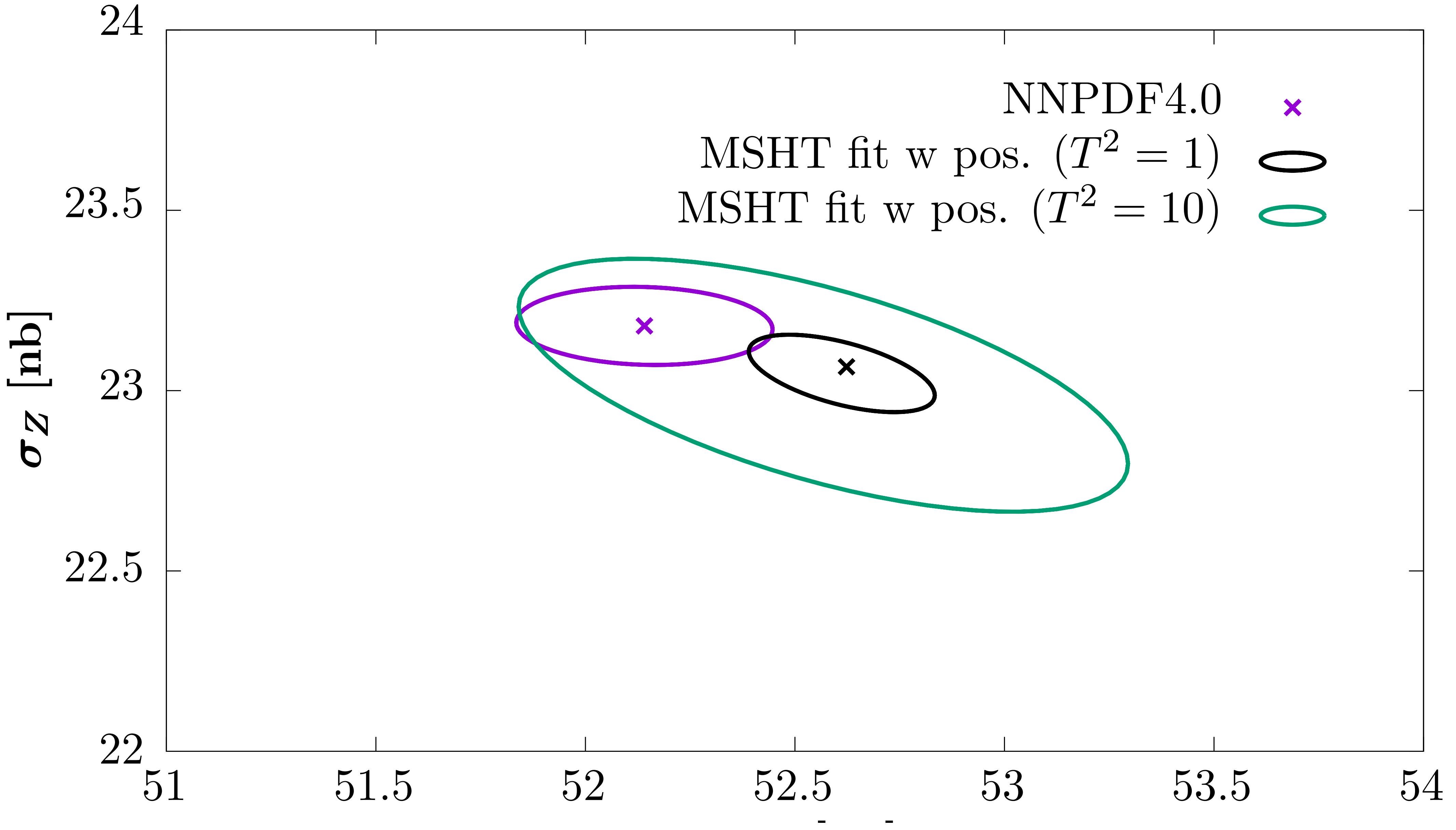
Cross Sections - p. charm



Comparison to hopscotch



A. Courtoy et al., arXiv: 2205.10444

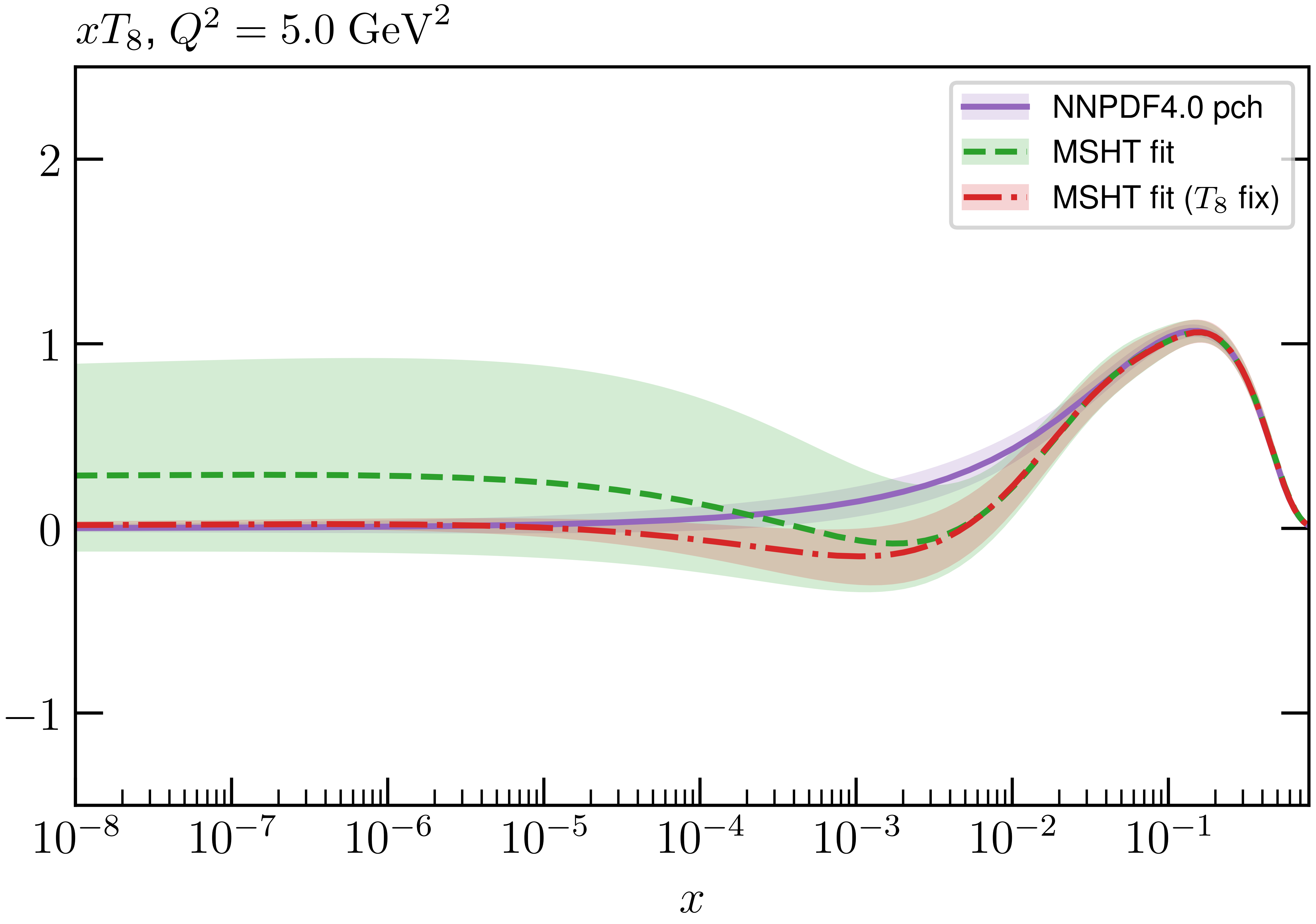


Integrability

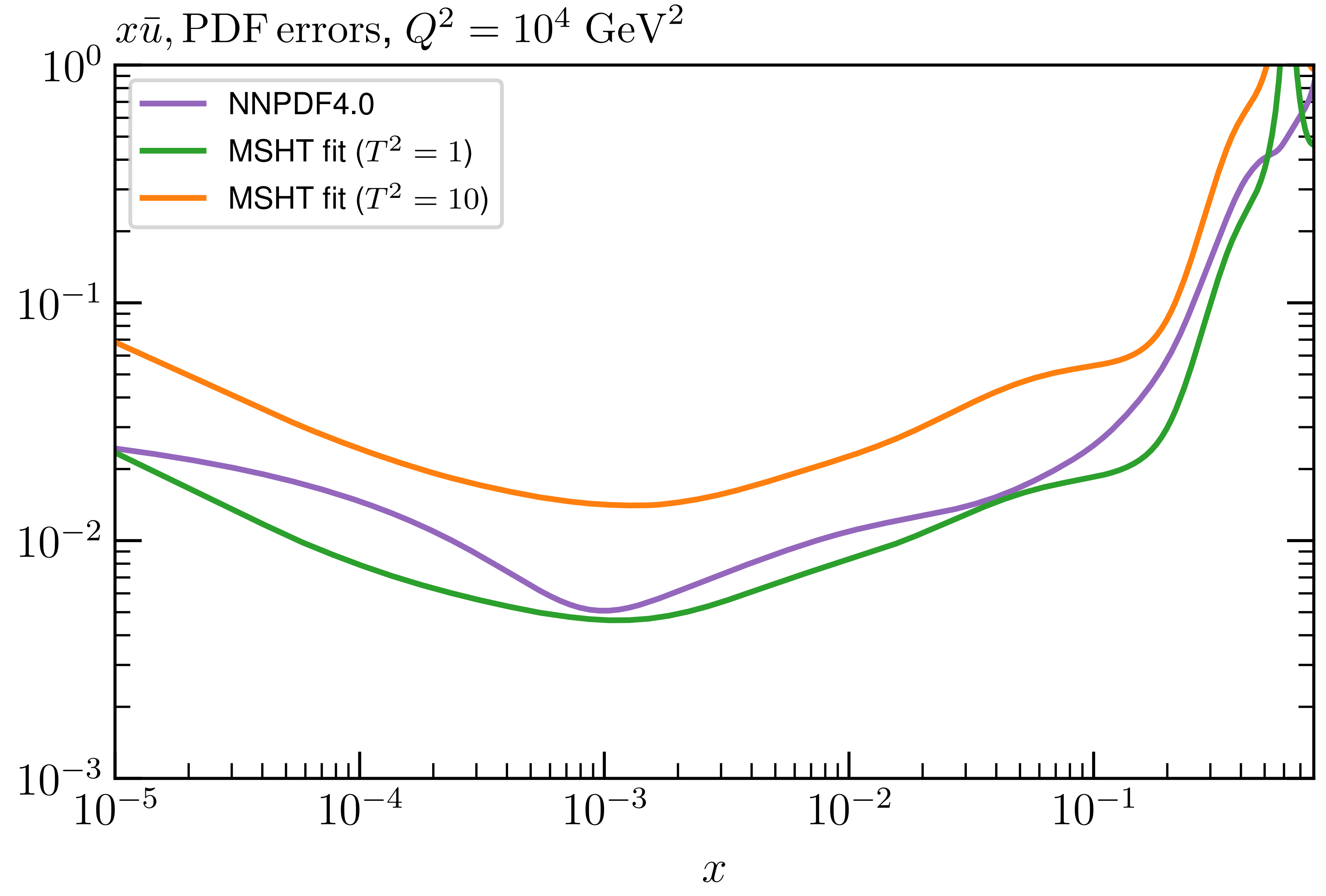
$$\lim_{x\to 0} x f_k(x,Q) = 0\,, \quad \forall\, Q\,, \qquad f_k = T_3,\, T_8\,.$$

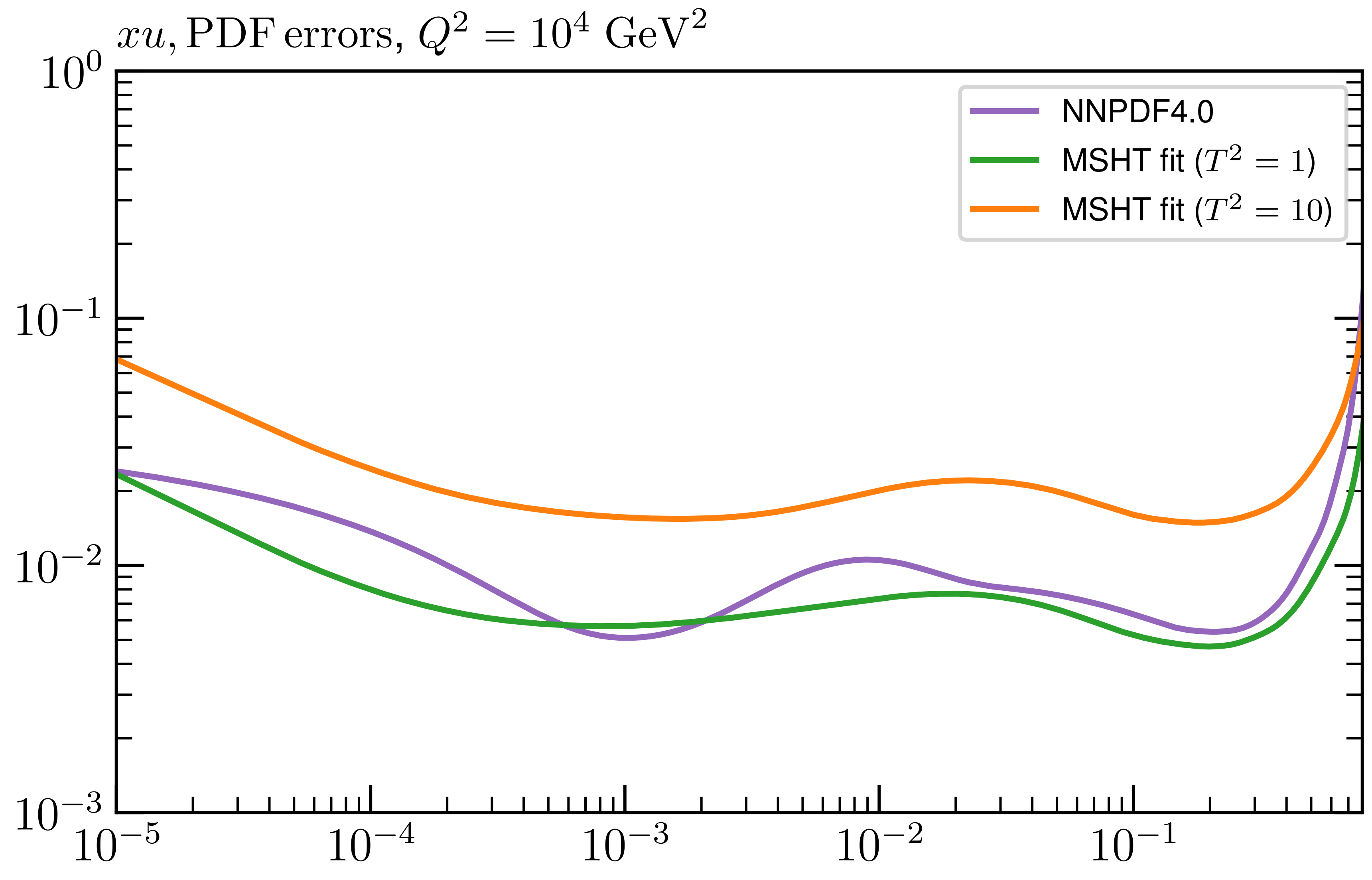
$$\chi^2_{\rm tot} \to \chi^2_{\rm tot} + \sum_k \Lambda_k^{(\rm int)} \sum_{i=1}^{n_i} \left[x f_k \left(x^{(i)}_{\rm int}, Q_i^2 \right) \right]^2\,, \qquad x^{(i)}_{\rm int} = 10^{-9},\, 10^{-8},\, 10^{-7}. \qquad \Lambda_k^{(\rm int)} = 100$$

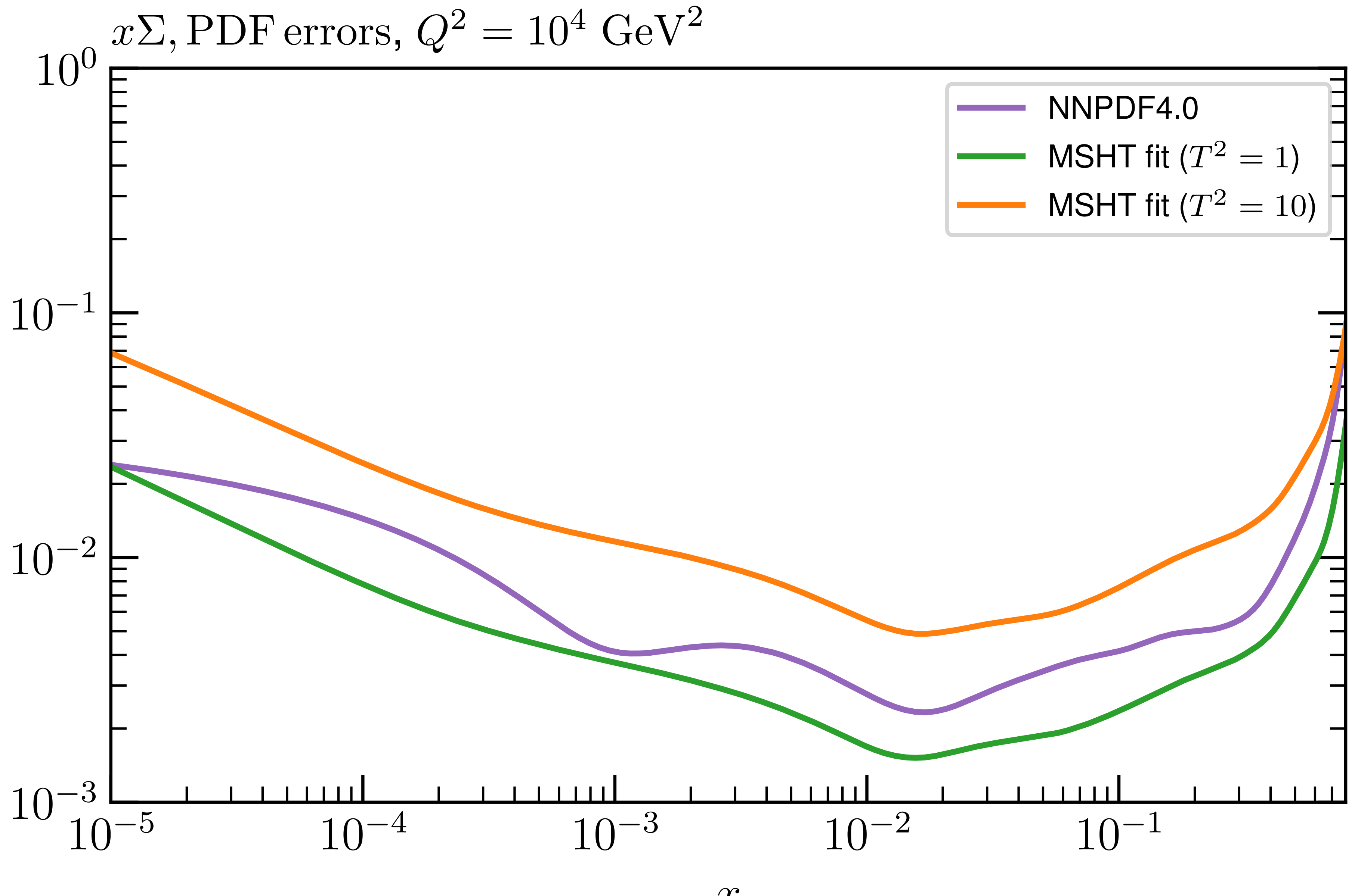
- Biggest deviation from this for MSHT is for the T_8 combination.
- In MSHT fixed parameterisation can impose that this vanishes at low x by simply fixing strangeness normalization.
- Gives ~ same fit quality for p. charm, and ~ 3 points worse for fitted.

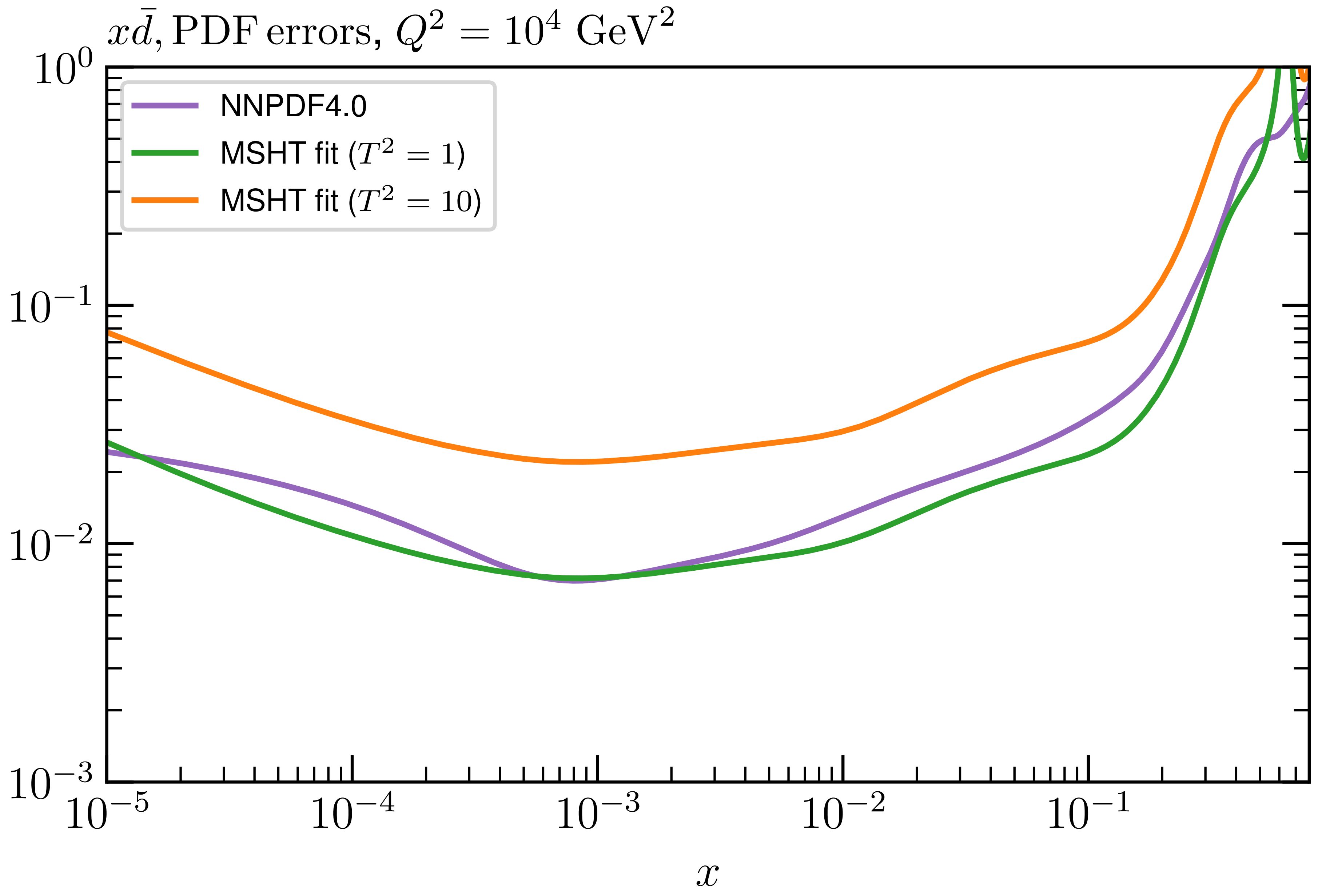


Uncertainties - p. charm

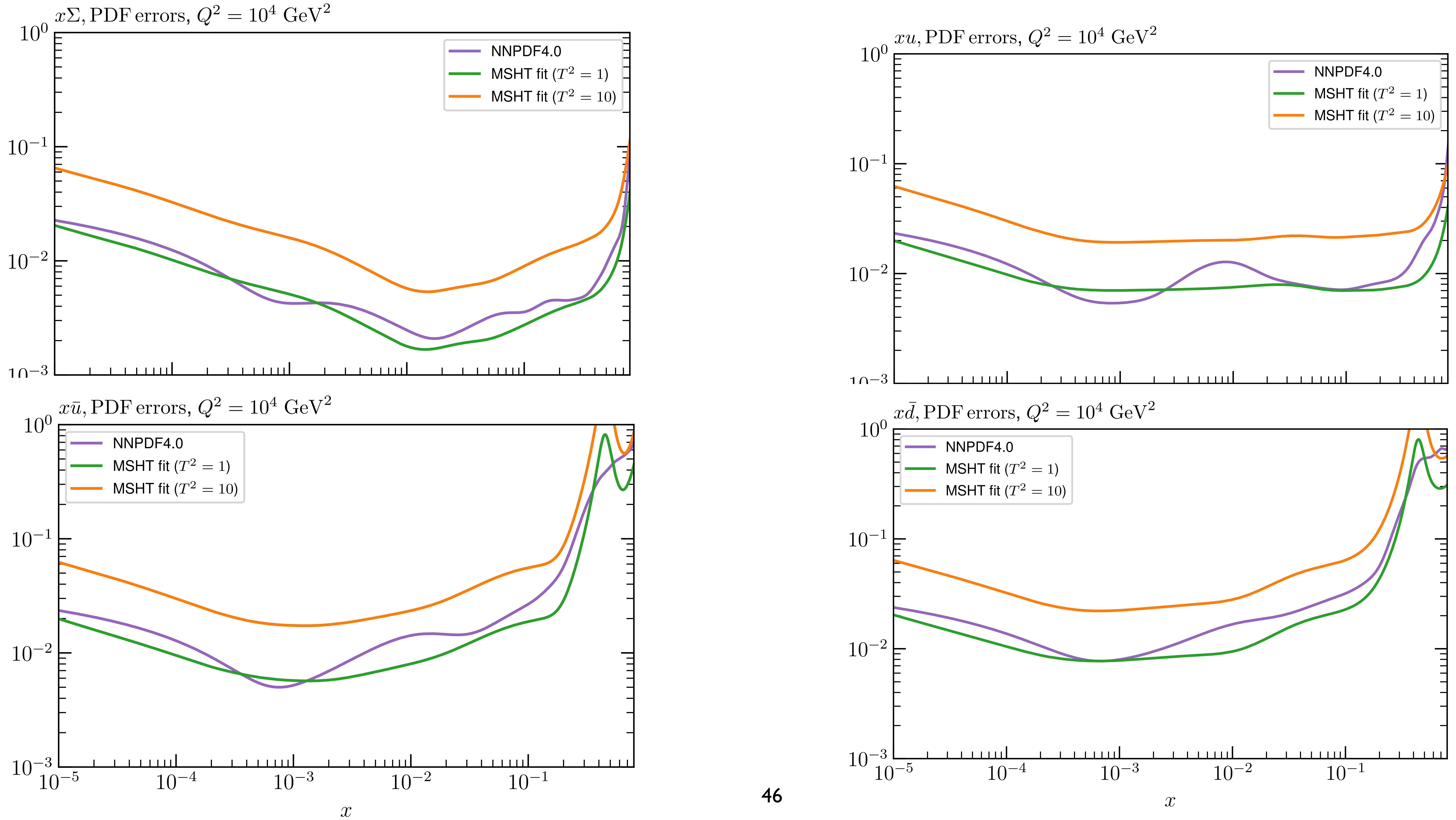








Uncertainties - fitted charm



Comparison to NNPDF uncertainties

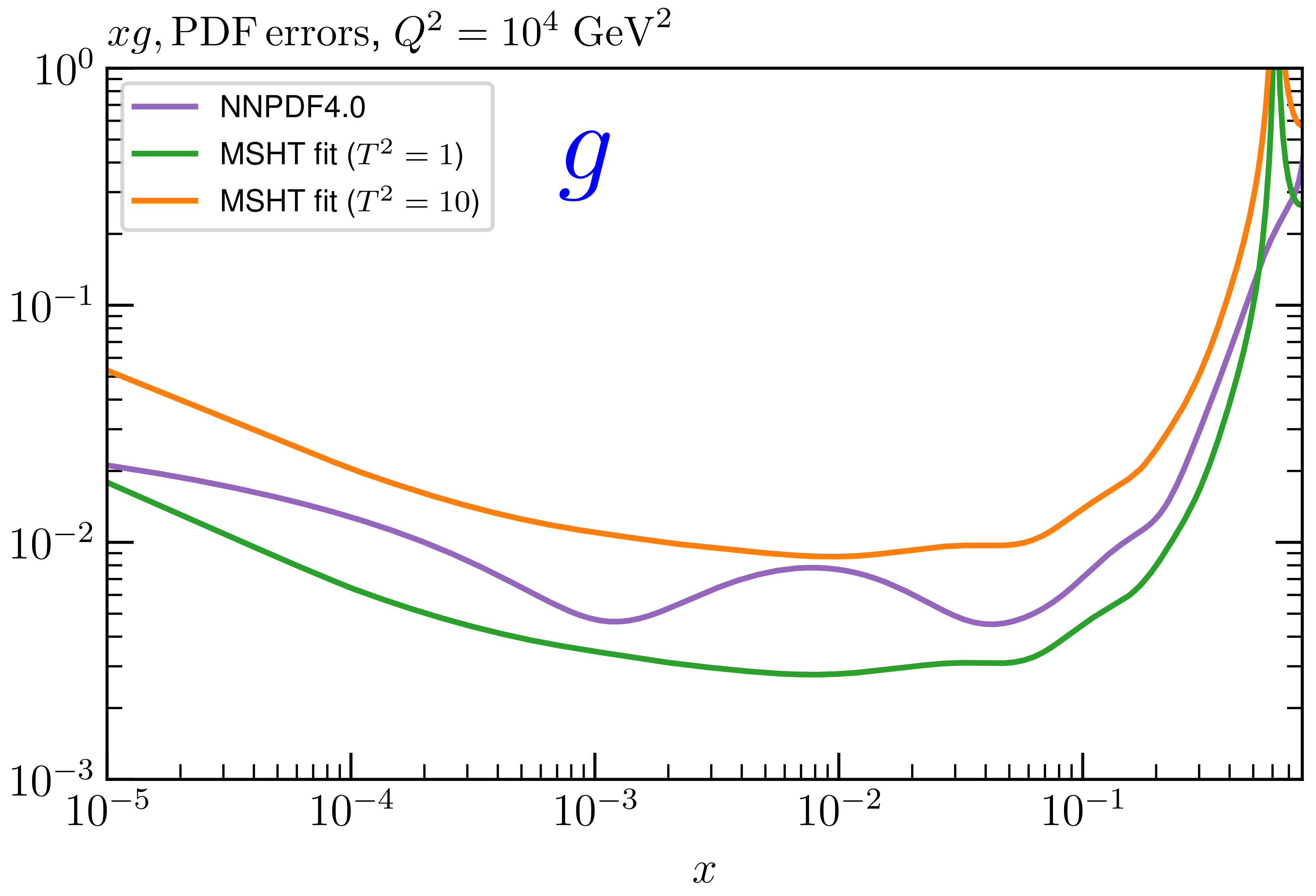
- Can do same comparison of MSHT vs. NNPDF PDF uncertainties but now in global fit. Completely like-for-like. Results very similar to closure test comparison:

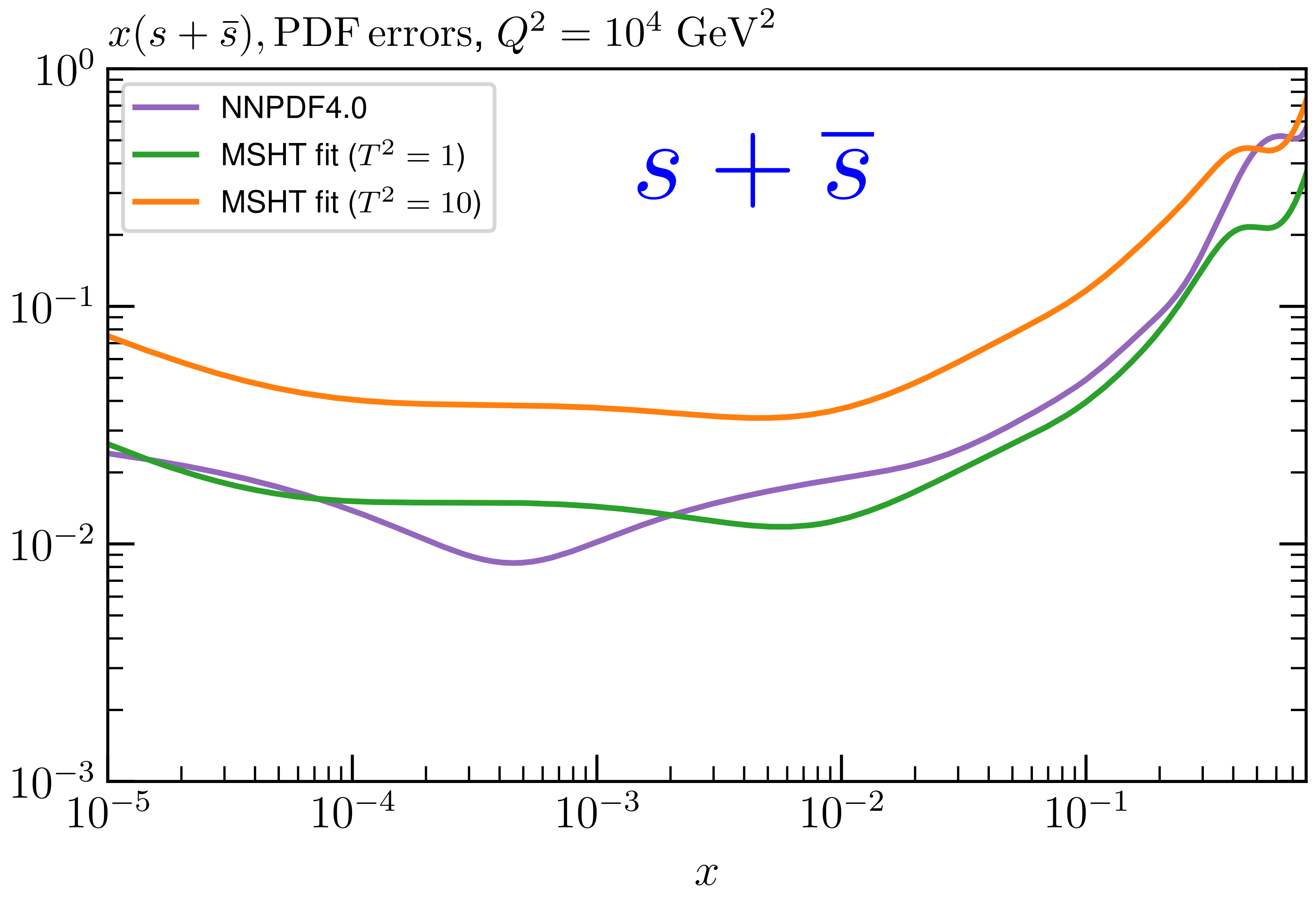
★ Quark flavour decomposition: $\sigma(\text{NNPDF}) \sim \sigma(\text{MSHT}, T^2 = 1)$

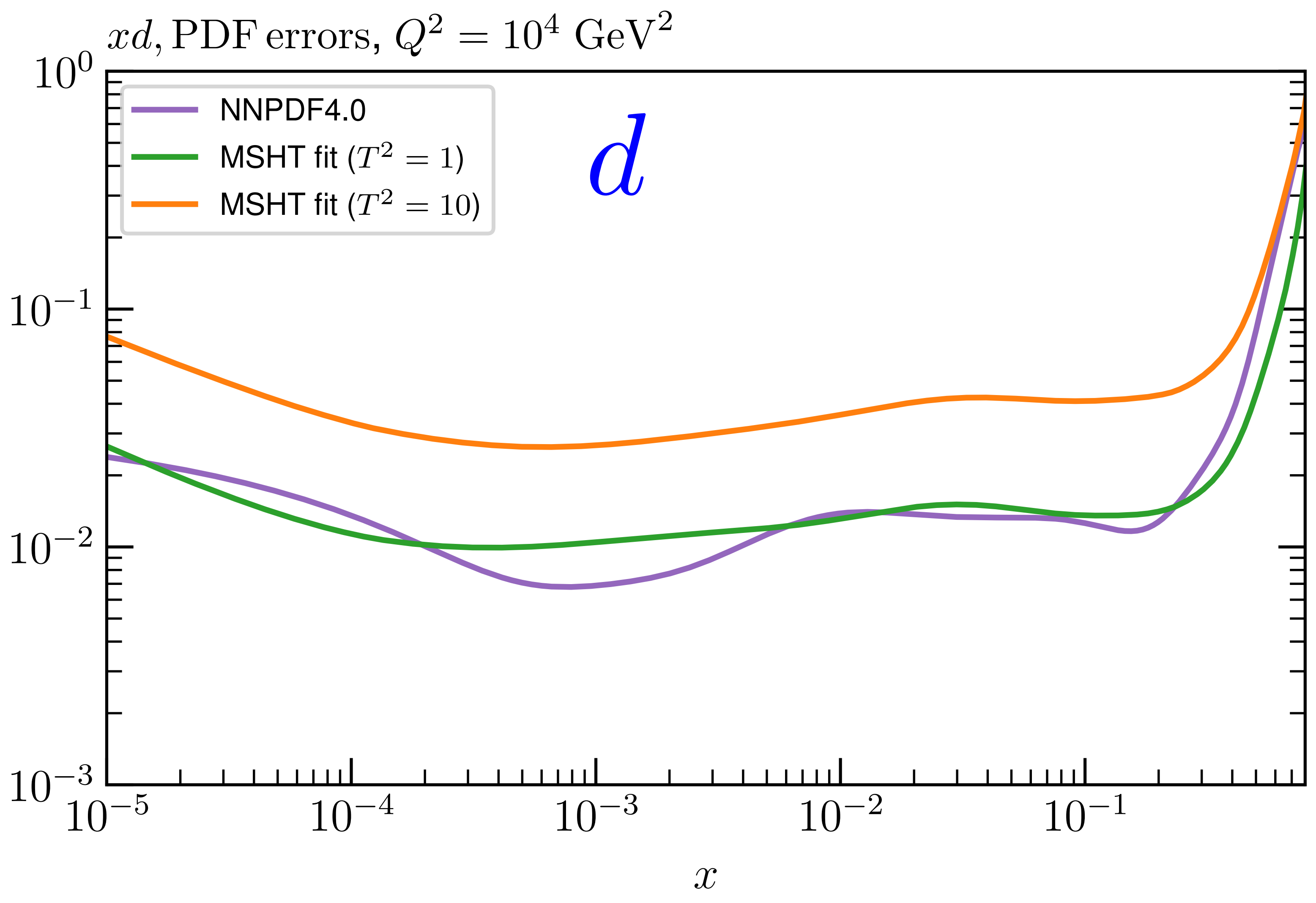
★ Gluon (singlet at intermediate x): $\sigma(\text{MSHT}, T^2 = 1) \lesssim \sigma(\text{NNPDF}) \lesssim \sigma(\text{MSHT}, T^2 = 10)$

MSHT, $T^2 = 1$
MSHT, $T^2 = 10$
NNPDF4.0pch

- With rather similar overall trends with x .
- Exception at high x where NNPDF uncertainty can become larger.







Positivity?

- General arguments for imposing strict positivity on PDFs outside of current data region rely on perturbative stability. Not clear for (very) low gluon - sensitivity to resummation etc.
- More importantly - **all cases** are actually **negative** at low x! Notable that the NNPDF gluon still prefers to be as negative as possible, i.e. just below the minimum x_i value where positivity imposed. Driving fit in undesirable way?