UNIVERSITY OF BIRMINGHAM | SCHOOL OF PHYSICS AND ASTRONOMY

EIC Early Science PDFs

S. Maple

Overview

- Lots of progress for EIC Early Science in the past year
- In the inclusive group we've produced studies for unpolarised and polarised structure functions, PDFs, and α_s extraction, which are to be shown between the main Early Science Report (ESR), NIM-A special edition paper, and other papers.
- For the ESR, space is limited, so the focus has been on "impact plots" where we can directly see reductions in error bars
  - "Impact" is most easily shown through PDF error plots

I'll be discussing the work that's been done for collinear PDF/nPDFs, which hasn't been shown in our meetings recently

How we study/extract PDFs

- Inclusive DIS data from HERA are the backbone of global proton PDF fits
- So far, we've used a HERAPDF-like approach for studies of the proton PDFs
- In HERAPDF, proton PDFs, xf(x), are generically parametrised at starting scale μ_{f0}^2 as:

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

- **The central parametrisation is:**

$$\begin{aligned}
xg(x) &= A_g x^{B_g}(1-x)^{C_g} - A'_g x^{B'_g}(1-x)^{C'_g}, \\
xu_v(x) &= A_{u_v} x^{B_{u_v}}(1-x)^{C_{u_v}}\left(1+E_{u_v}x^2\right), \\
xd_v(x) &= A_{d_v} x^{B_{d_v}}(1-x)^{C_{d_v}}, \\
x\bar{U}(x) &= A_{\bar{U}} x^{B_{\bar{U}}}(1-x)^{C_{\bar{U}}}(1+D_{\bar{U}}x), \\
x\bar{D}(x) &= A_{\bar{D}} x^{B_{\bar{D}}}(1-x)^{C_{\bar{D}}}.
\end{aligned}$$

- Additional Constraints:
  - A_{uv}, A_{dv}, A_g constrained by quark number sum rules and momentum sum rule
  - $B_{\bar{U}} = B_{\bar{D}}$
  - $x\bar{s} = f_s x\bar{D}$ at starting scale, $f_s = 0.4$

- **We perform studies in xFitter, fitting EIC pseudo-data alongside HERA data, and comparing to HERA-only**

Early Science DIS Pseudo-data

- EIC Pseudo-data prepared for PDF fits → need to choose a binning in x and Q^2, and estimate statistical, and uncorrelated and correlated systematic uncertainties

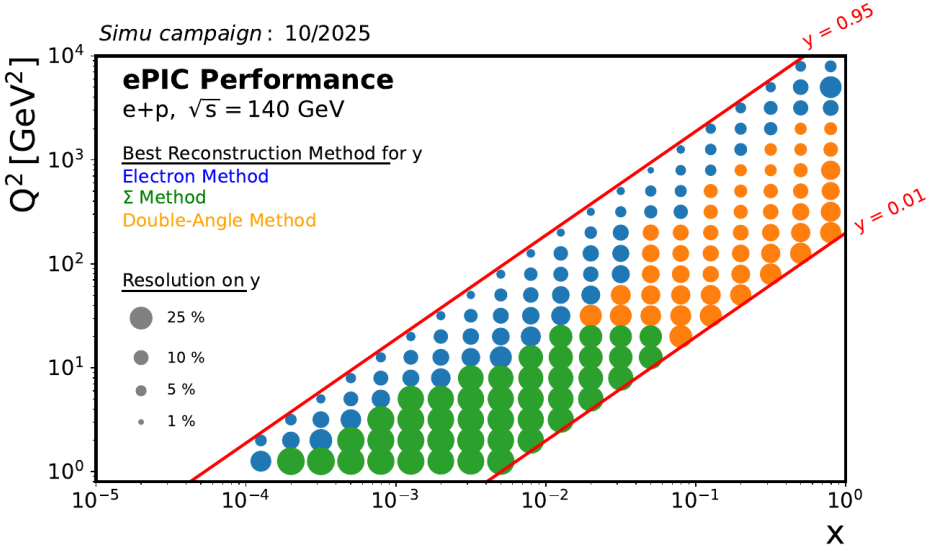
- Statistical uncertainties from full simulation of Pythia6 events (npsim) and subsequent reconstruction in EICrecon
  - Cuts typical of an inclusive DIS cross section analysis are applied
  - Fractional uncertainty on the reduced cross section goes as $1/(\sqrt{N_{rec}})$ where N_{rec} is the number of events reconstructed in a given x-Q^2 bin after losses due to inefficiencies, cuts, and bin migration

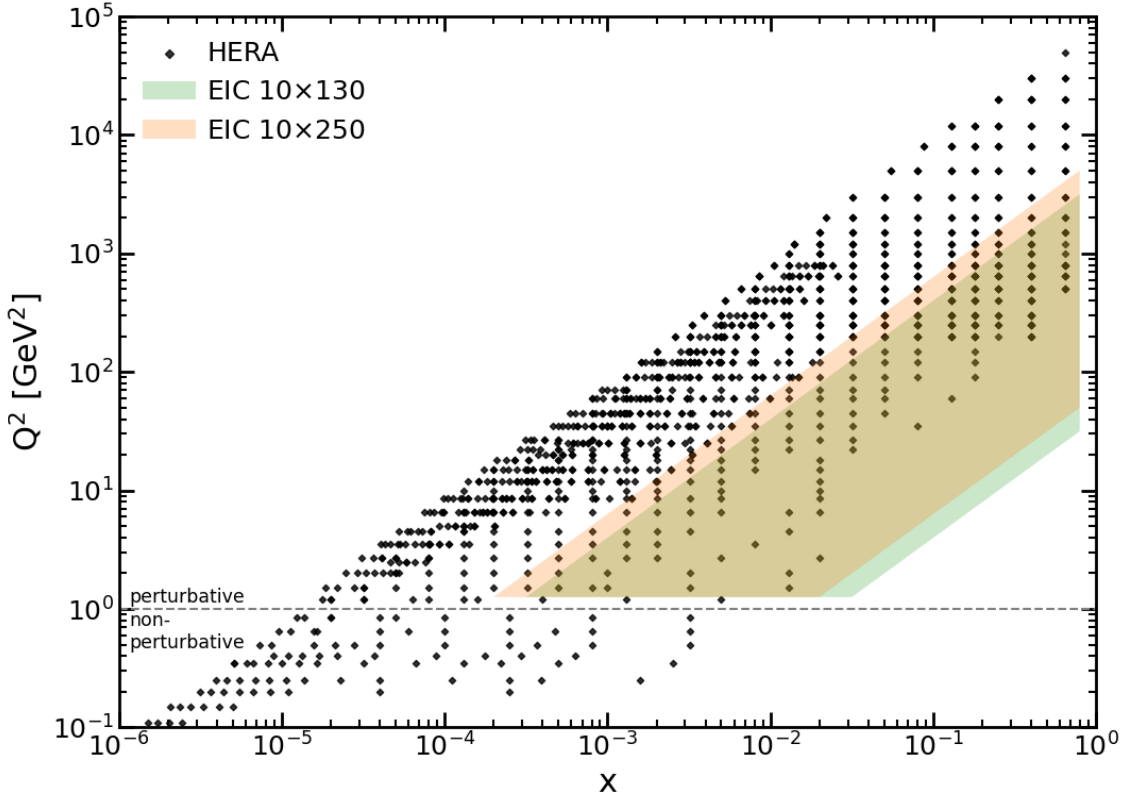
- Systematic uncertainties use (conservative) estimates based on Yellow Report studies:
  - Point-to-point uncorrelated uncertainty = 1.9%
  - Totally correlated normalisation uncertainty = 3.4% (1.5% from luminosity measurement, rest from electron purity and detector efficiency effects) → totally uncorrelated between different beam scenarios

- Cross section values generated according to central value of PDF prediction → smeared according to uncertainties

Early Science DIS Pseudo-data

- Choose an x-Q^2 binning to optimise overlap with HERA data points while maintaining resolution
  - ~5 bins per decade in x and ~5-8 in Q^2
- We gain valuable high precision points at high-x and moderate Q^2

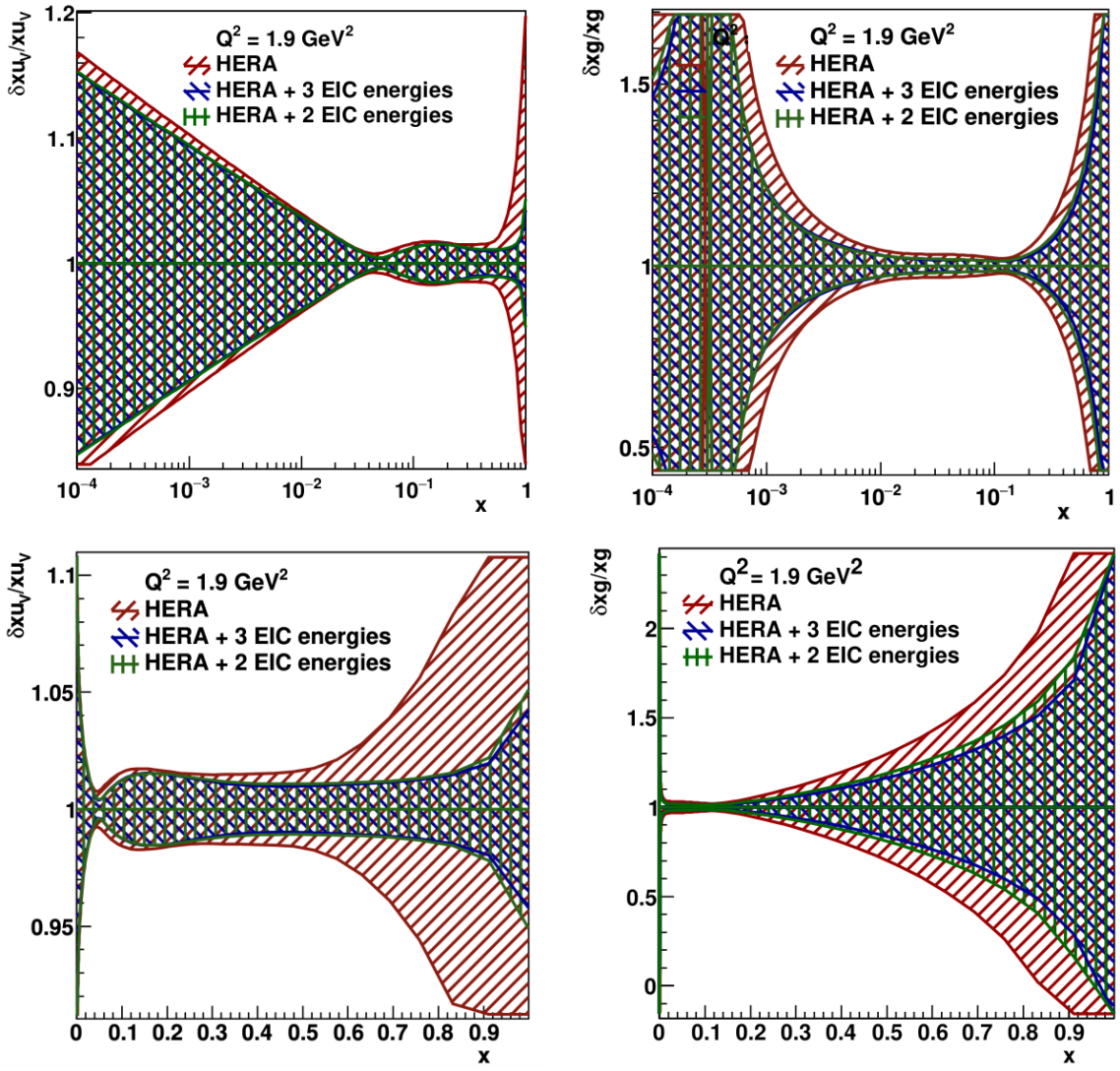




Proton PDF fits

- Evaluate impact on HERA-only fit
  - Early science data not as powerful as full EIC dataset → main impact is on the U_v and g PDFs
  - Stronger impact at high-x where HERA data lose sensitivity due to limited statistical precision

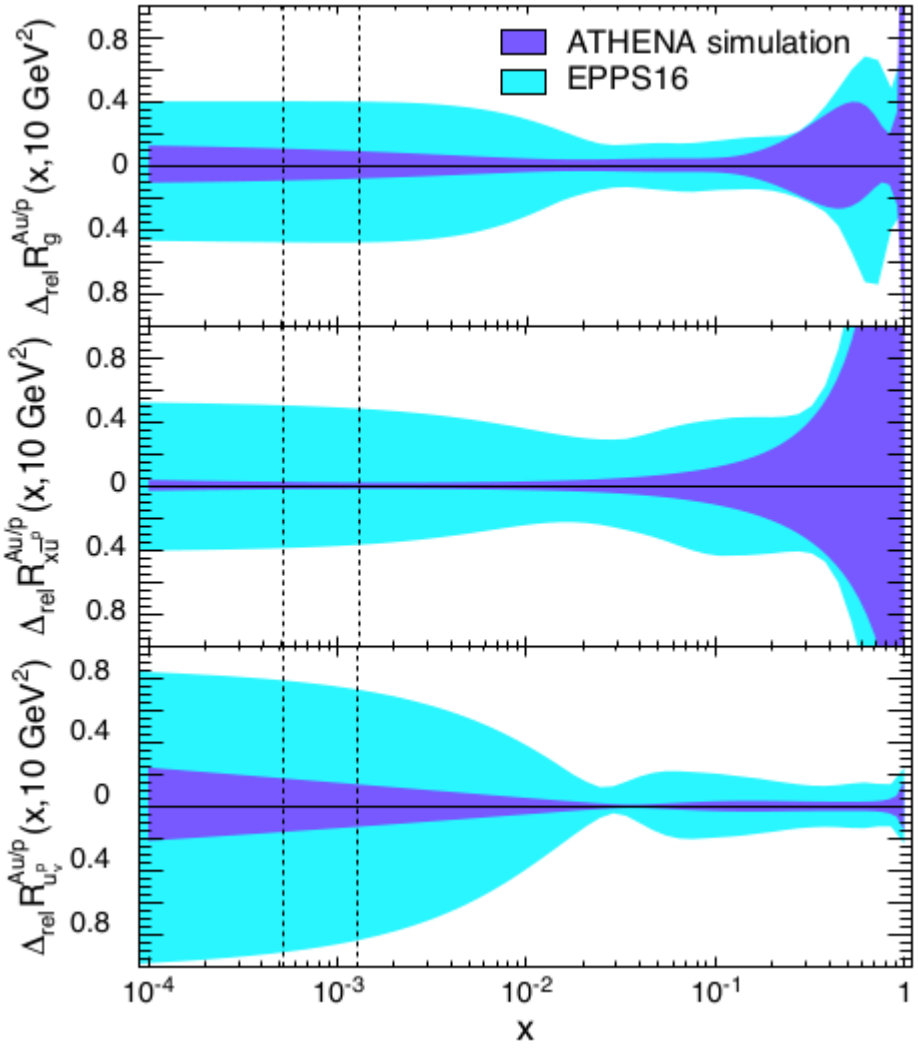
- Consider also an extra scenario where we have a lower energy (5GeV electron) run
  - Beneficial for structure functions, but only make marginal difference to PDFs



Fits + plots by K. Wichmann

Nuclear PDFs

- Ideally we would do something similar for nPDFs → We generate EIC eAu pseudo-data, and use xFitter to estimate the precision on EIC-only PDFs similarly to what was done for ATHENA
- **Problem:** we only have one eAu energy in the provisional run plan, with a maximum of 2
  - We will likely need 3 beam configurations which span a larger range in c.o.m. energy to get a fit to converge (Neither Katarzyna nor I could get the fits to converge with the energies we had – maybe with further effort)
- We can revisit this for preTDR, where things should work better.



The workaround: PDF reweighting

- Thanks to P. Zurita we managed to get an nPDF impact plot for the ESR by using a Bayesian reweighting procedure

- Bayesian reweighting essentially uses PDF probability distributions as input, applies weights to these distributions based on how well the new data is described and outputs an updated PDF probability distribution
  - For a PDF set with errors in MC replica format → N_{rep} replica PDFs

$$\langle O(\mathrm{PDF})\rangle = \frac{1}{N_{\mathrm{rep}}}\sum_{k=1}^{N_{\mathrm{rep}}} O(\mathrm{PDF}_k)$$

Observable e.g. DIS cross section

$$\sigma_{O(\mathrm{PDF})} = \sqrt{\frac{1}{N_{\mathrm{rep}}-1}\sum_{k=1}^{N_{\mathrm{rep}}}(O(\mathrm{PDF}_k)-\langle O(\mathrm{PDF})\rangle)^2}$$

Uncertainty on observable

$$\langle O^{\mathrm{new}}(\mathrm{PDF})\rangle = \frac{1}{N_{\mathrm{rep}}}\sum_{k=1}^{N_{\mathrm{rep}}} w_k O(\mathrm{PDF}_k)$$

$$w_k = \frac{(\chi^2_k)^{\frac{1}{2}(N_{\mathrm{data}}-1)}\exp^{-\frac{1}{2}\chi^2_k}}{\frac{1}{N_{\mathrm{rep}}}\sum_{k=1}^{N_{\mathrm{rep}}}(\chi^2_k)^{\frac{1}{2}(N_{\mathrm{data}}-1)}\exp^{-\frac{1}{2}\chi^2_k}},$$

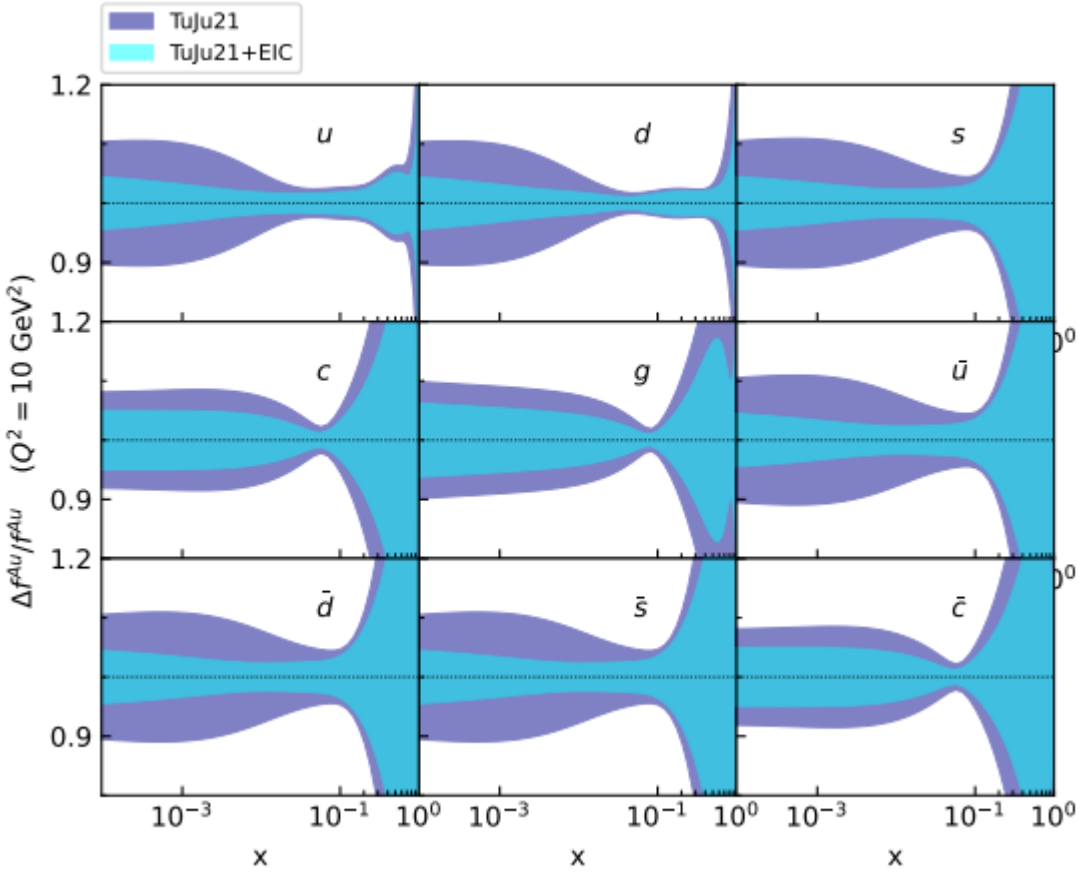
Update with N_{data} new data points

- The weighted PDF probability distribution can be turned into a new ensemble of PDF replicas
- → updated uncertainties (and central values)

Reweighted nPDFs

- Using reweighting we can obtain an estimate for the impact of a single 10x100 eAu configuration on the Au nPDF from a representative global nPDF fit (TUJU21)

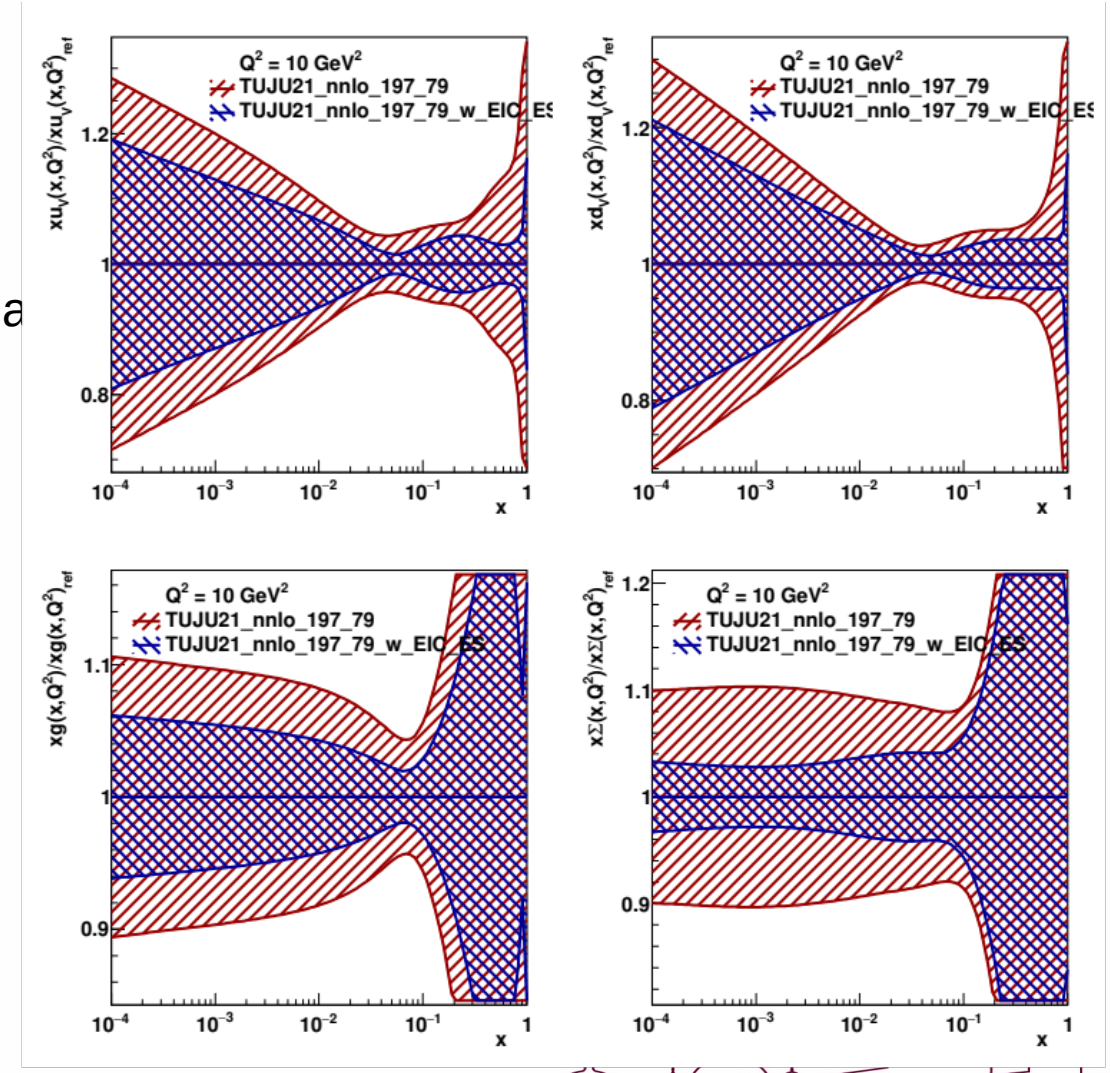
- Extend nuclear DIS data to x~10^{-3}
  → significant improvements at low x for all parton species



Plots by P. Zurita

Some sanity checks…

- TUJU21 LHAPDF set provides Hessian uncertainties as standard → Compare to a Hessian reweighting (profiling) in xFitter
  - Hessian profiling requires different mathematics but the goal is the same → estimate impact of new data on PDFs by comparing new data to theory predictions

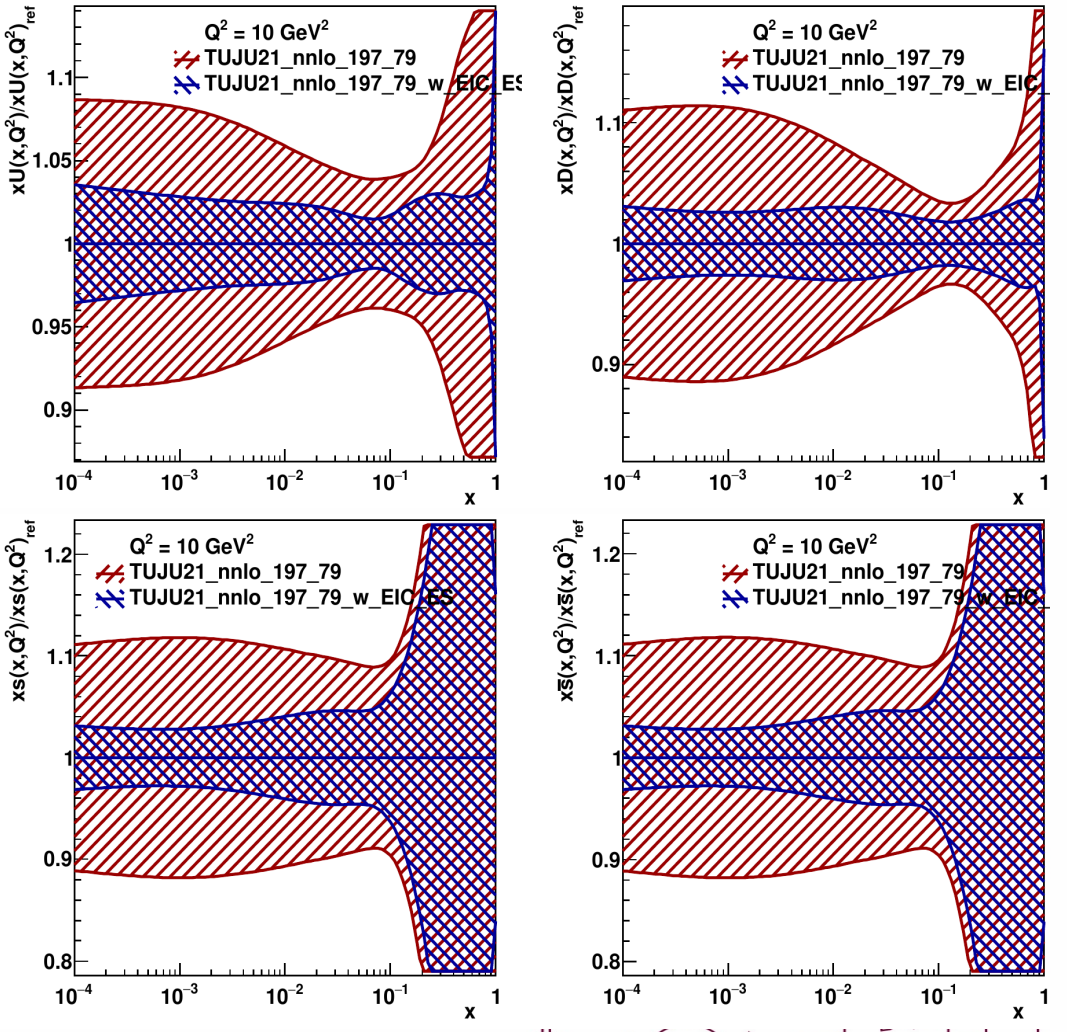


Improvements of similar magnitude and shape → slightly odd shape for the sea quarks at low x

Some sanity checks…

- TUJU21 LHAPDF set provides Hessian uncertainties as standard → Compare to a Hessian reweighting (profiling) in xFitter
  - Hessian profiling requires different mathematics but the goal is the same → estimate impact of new data on PDFs by comparing new data to theory predictions

Improvements of similar magnitude and shape → slightly odd shape for the sea quarks at low x



Summary and future work

- We have a good amount already in place
- Impact of early science proton data on MSHT global fit is in the works → hopefully to be included in NIM-A special edition
- Nuclear PDF plots based on reweighting are available
  - Goal is to use full fit of EIC only data for pTDR, similar to ATHENA
  - Also looking into inclusion of EIC pseudo-data in TUJU-like full fit in xFitter

Backup