Impact of the EIC on Collinear PDFs and the Strong Coupling

Physics Opportunities at an Electron-Ion Collider (POETIC) XI

Thomas Cridge

28th February 2025

In collaboration with MSHT colleagues - T.C., L.A. Harland-Lang, R.S. Thorne, and others - N. Armesto, F. Giuli, P. Newman, B. Schmookler, K. Wichmann

MSHT Global PDF Fitting

- Global fit of collinear unpolarised PDFs. More than 60 different datasets - Fixed Target, HERA DIS, neutrinos, Drell-Yan, Tevatron, LHC. 6 neutrinos, 2 fixed target DY, 8 HERA, 8 Tevatron, 27 LHC.
- Almost 5000 datapoints included over wide range of (x, Q^2): $10^{-4} \lesssim x \lesssim 0.8$ and $2\ \mathrm{GeV}^2 \lesssim Q^2 \lesssim 10^6\ \mathrm{GeV}^2$.
- Robust methodology with developments on all three fronts:
  1. Theoretical - Vast majority of processes included have full NNLO QCD theory, with NLO EW where relevant. Recent extension to approximate N3LO with theoretical uncertainties for first time. TC et al, 2207.04739
  2. Experimental - Many new datasets, more precise, more channels, more differential.
  3. Methodological - Extended parameterisation, 52 PDF parameters - allow fitting to accuracy $< 1\%$. Closure tests performed to examine central value and uncertainties. Harland-Lang, TC, Thorne, 2407.07944

MSHT Global PDF Fitting

- Global fit of collinear unpolarised PDFs. More than 60 different datasets - Fixed Target, HERA DIS, neutrinos, Drell-Yan, Tevatron, LHC. 6 neutrinos, 2 fixed target DY, 8 HERA, 8 Tevatron, 27 LHC.
- Almost 5000 datapoints included over wide range of (x, Q^2): $10^{-4} \lesssim x \lesssim 0.8$ and $2\ \mathrm{GeV}^2 \lesssim Q^2 \lesssim 10^6\ \mathrm{GeV}^2$.
- Robust methodology with developments on all three fronts:
  1. Theoretical - Vast majority of processes included have full NNLO QCD theory, with NLO EW where relevant. Recent extension to approximate N3LO with theoretical uncertainties for first time. TC et al, 2207.04739
  2. Experimental - Many new datasets, more precise, more channels, more differential.
  3. Methodological - Extended parameterisation, 52 PDF parameters - allow fitting to accuracy $< 1\%$. Closure tests performed to examine central value and uncertainties. Harland-Lang, TC, Thorne, 2407.07944

What can the EIC contribute to this? $\Rightarrow$ Precise, new DIS data.

EIC Kinematic Coverage:

- Consider NC and CC DIS at EIC.
  - Higher x coverage, still at moderate Q^2.
  - Complements HERA data, which are backbone of PDF fits still.
  - EIC less sensitive to higher twists than fixed target data in global fits.
  - Study here - generate pseudodata for e^-p data with updated beam energies, configurations, lumis and uncertainty projections.
  - Kinematic coverage: $Q^2 > 2\ \mathrm{GeV}^2$, $0.01 < y < 0.95$, $W^2 > 15\ \mathrm{GeV}^2$.
  - Only highest $\sqrt{s}$ has CC DIS.

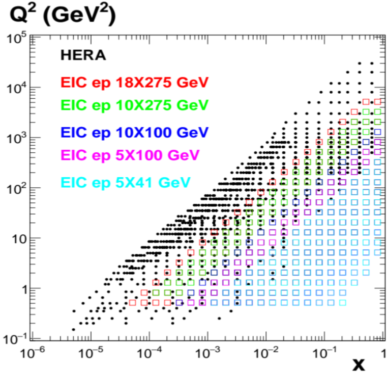


TABLE I. Beam energies, center-of-mass energies and annual integrated luminosities of the different configurations considered for the EIC.

e-beam energy (GeV)	p-beam energy (GeV)	$\sqrt{s}$ (GeV)	Integrated lumi (fb^{-1})
18	275	141	15.4
10	275	105	100
10	100	63	79.0
5	100	45	61.0
5	41	29	4.4

EIC Kinematic Coverage:

- Consider NC and CC DIS at EIC.
  - Higher x coverage, still at moderate Q^2.
  - Complements HERA data, which are backbone of PDF fits still.
  - EIC less sensitive to higher twists than fixed target data in global fits.
  - Study here - generate pseudodata for $e^- p$ data with updated beam energies, configurations, lumis and uncertainty projections.
  - Kinematic coverage: $Q^2 > 2\ \mathrm{GeV}^2$, $0.01 < y < 0.95$, $W^2 > 15\ \mathrm{GeV}^2$.
  - Only highest $\sqrt{s}$ has CC DIS.

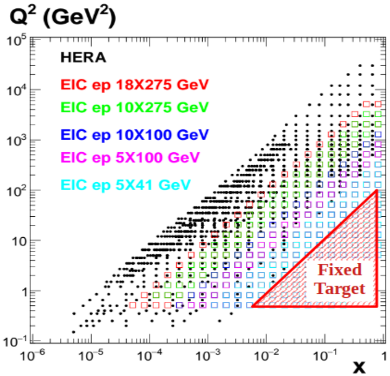
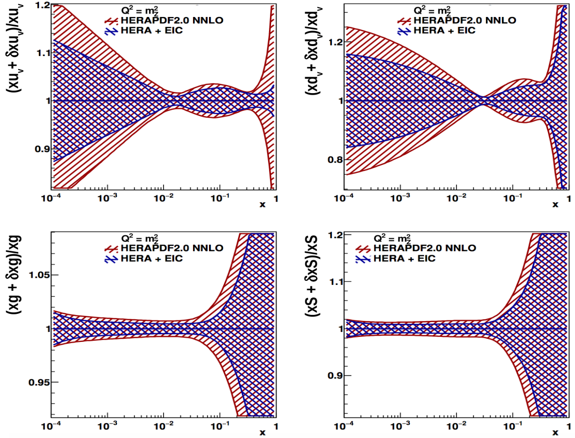


TABLE I. Beam energies, center-of-mass energies and annual integrated luminosities of the different configurations considered for the EIC.

e-beam energy (GeV)	p-beam energy (GeV)	$\sqrt{s}$ (GeV)	Integrated lumi (fb^{-1})
18	275	141	15.4
10	275	105	100
10	100	63	79.0
5	100	45	61.0
5	41	29	4.4

PDF Impact in HERAPDF:

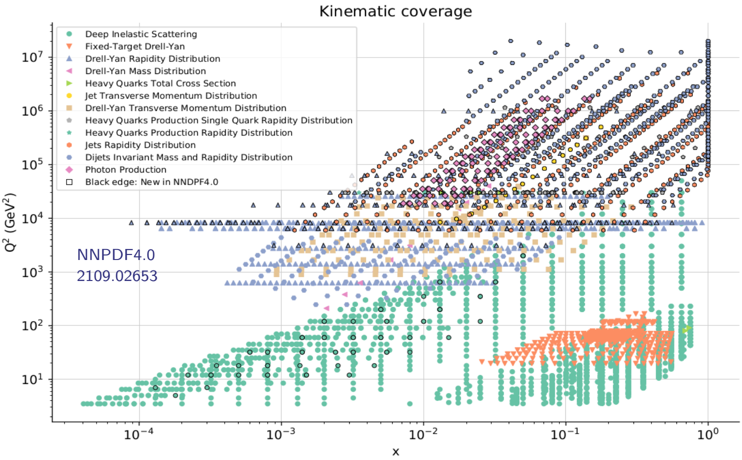
- Observed large reductions in PDF uncertainty when EIC data added on top of HERAPDF, no fixed target or LHC data.



Armesto, TC et al 2309.11269

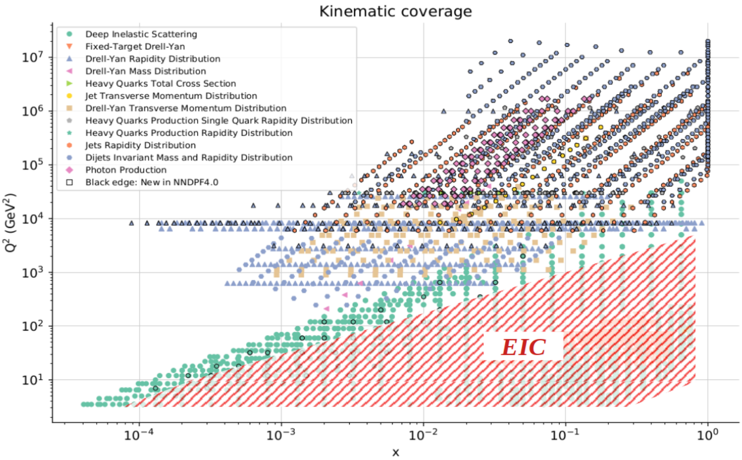
PDF Impact in MSHT:

- Does the same hold for global PDFs? Also fixed target and LHC data.



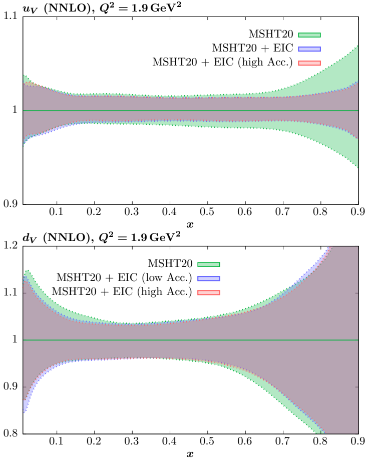
PDF Impact in MSHT:

- Does the same hold for global PDFs? Also fixed target and LHC data.



PDF Impact in MSHT:

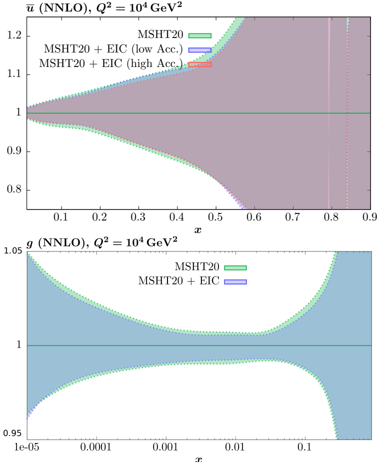
- Does the same hold for global PDFs? Also fixed target and LHC data.
- Add the pseudodata to global MSHT PDFs at NNLO and assess impact:
- Largest impact on u PDF at large x as $\sigma^{\mathrm{NC\ DIS}}_{e^-p} \propto \sum_i Q_i^2 f_i(x)$.
  $\Rightarrow$ Uncertainty reduced by up to 50%.
- Smaller impact on d PDF.
- Impact of larger y acceptance negligible as different beam energy configurations provide constraints.
- Positron or deuteron data would increase constraints on d PDF. As would tagged DIS studies and PVDIS with polarised electrons. (see e.g. CJ/JAM in 2103.05419 and S. Liu's talk yesterday)



Armesto, TC et al 2309.11269

PDF Impact in MSHT:

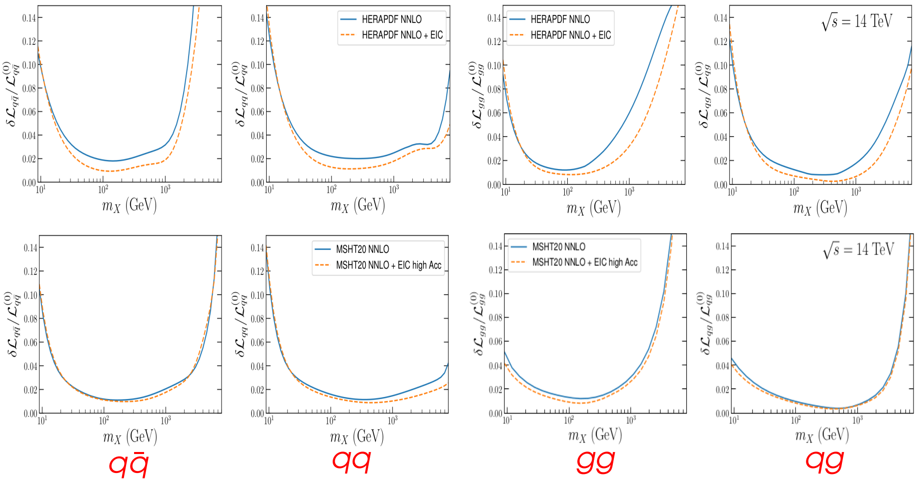
- Add pseudodata to global MSHT PDFs at NNLO and assess impact:
- Inclusive DIS has smaller impact on sea quark PDFs, where uncertainties are larger.
- Mild reduction in gluon uncertainty across all x.
- Comes from scaling violations, $dF_2/dQ^2 \sim \alpha_S g$.
- Similar EIC constraints seen in HERAPDF but greater in magnitude there as it's not a global PDF fit.
- Also investigated sensitivity to small-x $\ln(1/x)$ resummation
  - no difference in fit quality observed.



Armesto, TC et al 2309.11269

PDF Luminosity Impact in MSHT:

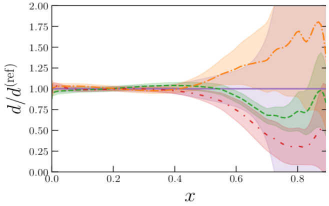
- Knock-on impact on PDF luminosity uncertainties in HERA/MSHT:

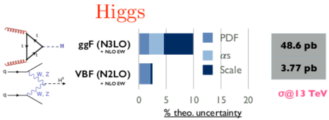


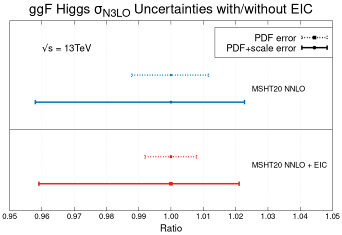
Armesto, TC et al 2309.11269

Consequences for Phenomenology:

- Why is this important?
- High x PDF (quark or gluon) uncertainties currently grow rapidly.
- Limits sensitivity to BSM physics at large invariant masses.
- Reason is lack of data and tensions observed between fixed target/LHC data $\Rightarrow$ EIC can help resolve these!
- Gluon uncertainty key for Higgs production cross-section uncertainty.
- Observe reduction in gg luminosity PDF uncertainty from 1.2% to 0.8% $\Rightarrow$ impact on $gg \to H$ cross-section.



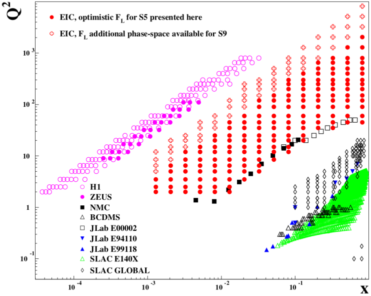


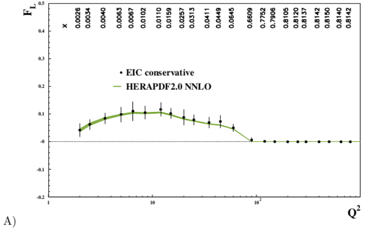


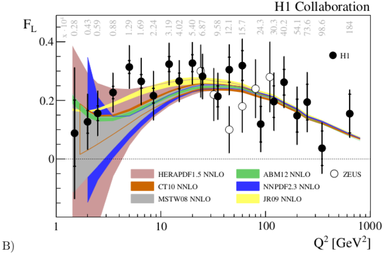
Further PDF Constraints - F_L in HERAPDF:

- Additional direct sensitivity to the gluon from F_L measurements. $F_L \sim \alpha_S g$
- Possible over larger range than HERA.
- Separate by Rosenbluth method using:

$$\sigma_{red}^{NC} \sim F_2 - \frac{y^2}{Y_+} F_L$$





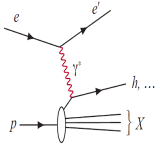


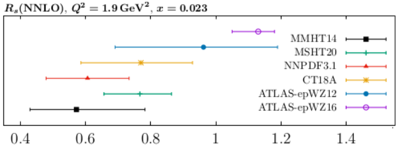
Jiménez-Lopez, Newman, Wichmann 2412.16123

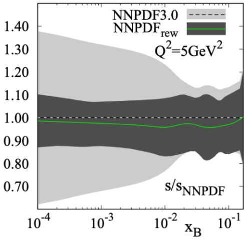
Further PDF Constraints - Strangeness:

- Limited strangeness sensitivity from inclusive DIS EIC measurements.
- Use SIDIS - parton content of outgoing hadron is connected to fragmenting parton and via CC/NC vertex to the parton in the proton.
- Pickup uncertainties from fragmentation functions.
- Similar to νDIS already used from NuTeV, which provides main constraint on $s - \bar{s}$ asymmetry, and from future FPF at CERN.
- Proton strangeness observed to be larger at LHC.
- Further s constraints come from charm jets.

(see e.g. Arratia, Hobbs et al in 2006.12520)



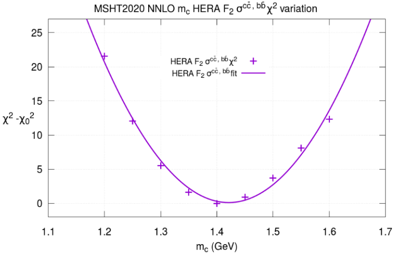




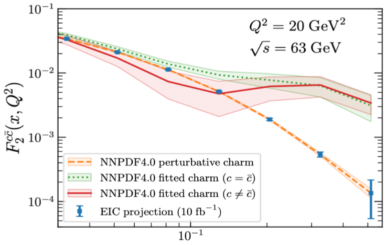
NNPDF in 2103.05419

Further PDF Constraints - Heavy Quarks:

- Measurements of charm and bottom structure functions will be extended to higher x.
- Gives sensitivity to high x heavy quark PDFs, and to heavy quark masses. E.g. used HERA data in MSHT (lower left).
- Recent suggestions of a fitted charm component of proton at high x by NNPDF, using EMC F_2^c and LHCb (Z+c) data.
- Several questions in community about this ⇒ can be resolved by EIC.



MSHT, TC et al 2106.10289



NNPDF, R.D. Ball et al 2311.00743

Determination of the Strong Coupling Constant:

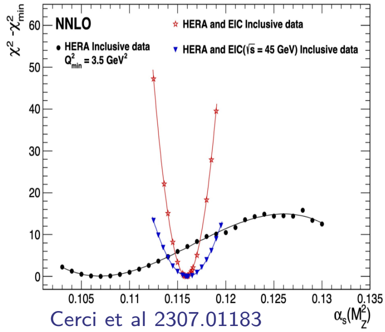
- $\alpha_S(M_Z^2)$ sensitivity in global PDF fit come from:
  - Direct $\alpha_S(M_Z^2)$ dependence in coefficient functions.

$$\mathbf{C}(\alpha_S) = \alpha_S^i[\mathbf{C}_0 + \alpha_S \mathbf{C}_1 + \alpha_S^2 \mathbf{C}_2 + \alpha_S^3 \mathbf{C}_3 + \ldots]$$

  - Indirect $\alpha_S(M_Z^2)$ dependence through PDF evolution.

$$\frac{df}{d\log\mu_F^2} = \begin{bmatrix} P_{qq} & n_f P_{qg} \\ P_{gq} & P_{gg} \end{bmatrix} \begin{bmatrix} \Sigma \\ g \end{bmatrix}$$

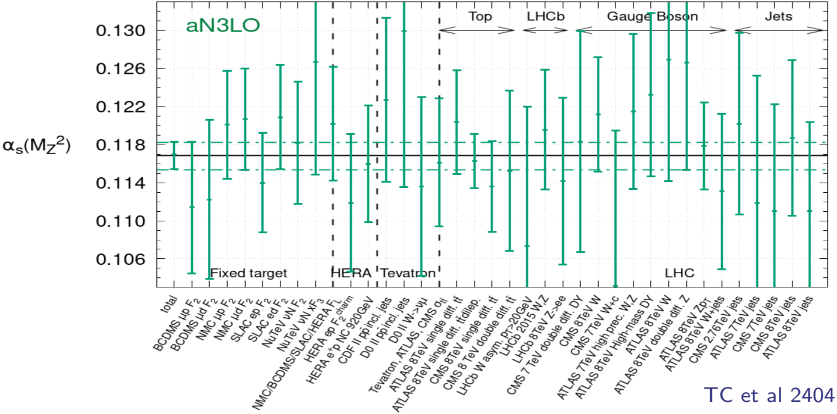
- DIS has limited sensitivity indirectly via scaling violations.
- HERA at low/intermediate x driven by gluon splitting, hard to disentangle α_S.
- EIC at higher x driven by non-singlet splitting, so α_S less correlated to g.
- Improved precision + more datapoints on structure function evolution.



Cerci et al 2307.01183

Determination of the Strong Coupling Constant:

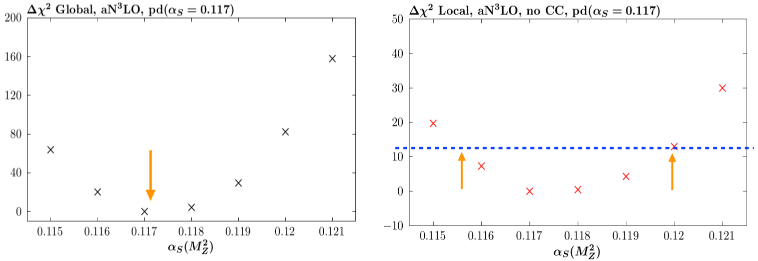
- Can we improve our global bounds? Again have fixed target and LHC data which also bound α_S.
- MSHT recently performed first determination in PDF fit at approximate N3LO: $\alpha_{S,\mathrm{aN3LO}}(M_Z^2) = 0.1170 \pm 0.0016$.



TC et al 2404.02964

Determination of the Strong Coupling Constant:

- Utilise same pseudodata now generated at aN3LO and with $\alpha_S(M_Z^2) = 0.117$ - consistent with PDFs. Fit simultaneously PDF+α_S.
- Examine χ^2 profile of EIC pseudodata to determine its bounds on α_S.

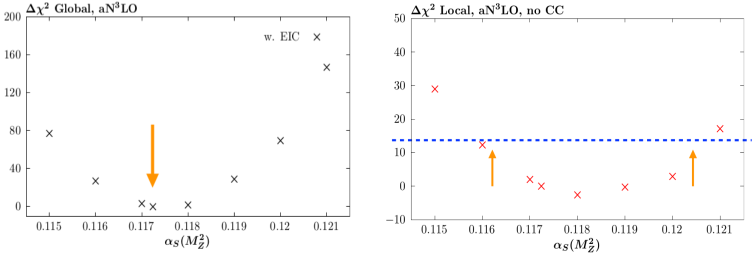


- Bounds set via dynamical tolerance $\Delta\chi^2 < (1 - \frac{\xi_{68}}{\xi_{50}}\chi_0^2) \sim 13$ for EIC NC data. Upper bound on α_S found not competitive.
- Lower bound ~ -0.0015 competitive $\Rightarrow$ would be best lower bound, better than SLAC/NMC d $\sim -0.0016/17$ which currently set limits.

Harland-Lang, TC, Thorne et al (work in progress) - Preliminary!

Determination of the Strong Coupling Constant:

- What happens if preferred α_S value different to global PDF fit?
- Instead generate pseudodata with $\alpha_S(M_Z^2) = 0.118$ and repeat.
- Now global $\alpha_S(M_Z^2)$ best fit is shifted up, from 0.1170 to 0.1172.

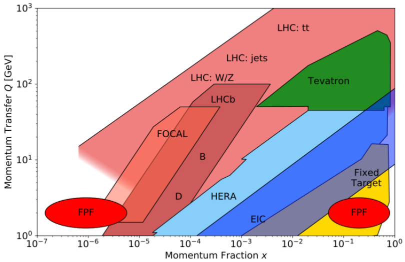


- Now sets tighter lower bound than before ~ -0.0012 (as it prefers larger $\alpha_S(M_Z^2)$) and weaker upper bound.
- Interplay of preferred α_S and uncertainty on bounds is often neglected.
- Improved precision on α_S directly reduces uncertainties on many key SM processes.

Harland-Lang, TC, Thorne et al (work in progress) - Preliminary!

Conclusions:

- EIC provides important constraints on collinear PDFs in its own right.
- Constrains proton in high x low/moderate Q^2 region, complementary to HERA at lower x and current/future LHC data at higher Q^2.
- Combination with collider programs elsewhere enhances this further.
- Further constraints from Future Physics Facility at LHC (ν CC DIS), future LHeC, etc. (see C. Gwenlan's talk yesterday)
- DIS at many x, Q^2 points ties down proton structure in clean environment
- Additional constraints on Standard Model parameters α_S, $\sin^2\theta_W$, $m_{c,b}$.
- Outputs feeds into important SM and BSM analyses at the LHC and beyond ⇒ Higgs, top, BSM.



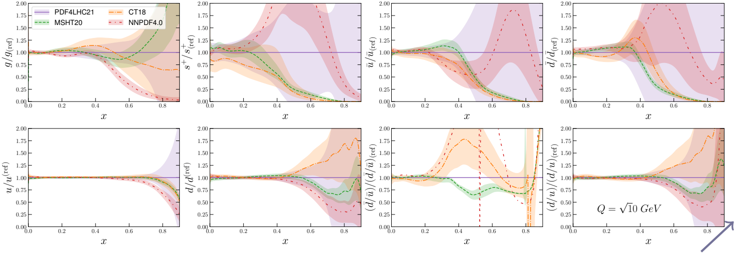
FPF White Paper 2203.05090

Acknowledgements

T. Cridge acknowledges funding funding by Research Foundation-Flanders (FWO) project number: 12E1323N.

High x PDF Comparison

- High x PDFs important for BSM searches, yet quite unconstrained.
- High x PDFs constrained by fixed target, asymmetries, LHC (e.g. jets, top, Zp_T). Use of high x low Q^2 data limited by Q^2, W^2 cuts.
- PDFs at very large x and low Q are connected to collider measurements at lower x and high Q by evolution.



- Quite large spread of the PDFs at high x + uncertainties grow rapidly!
- Both related to fact we have *limited data in this region*:
  - Data differences/tensions can have a larger effect.
  - More sensitive to methodological differences + theoretical assumptions.

Extrapolation

High x PDF Comparison

Data effects

- Strangeness raised by inclusion of ATLAS high precision 7, 8 TeV W, Z data - not in CT18.
- Overall strangeness is balance of this LHC precision DY data with older NuTeV dimuon data.
- $\bar{d}/\bar{u}$ raised at $x \sim 0.4$ by Seaquest data. Included only in NNPDF4.0. Seaquest tension with NuSea?
- Recent STAR data on W^+/W^- may also be relevan
- High x gluon affected by balance of LHC jet, top an Zp_T data + treatment of correlated systematics' issues.
- High x at low Q^2 connected to lower x at higher Q^2 by evolution ⇒ data at lower x may have indirect effects. Sum rules connect different x regions.