Nuclear PDFs and the Impact of an Electron-Ion Collider

Jacob Ethier

EINN 2019 Workshop "The Road to an Electron-Ion Collider"

October 30th, 2019

Overview

- Motivation
- Nuclear PDFs from NNPDF (R. Abdul Khalek, JJE, and J. Rojo)

  → nNNPDF1.0 – Monte Carlo approach to nPDFs using NNs

  → Towards nNNPDF2.0 – Impact of CC DIS + W/Z production at LHC

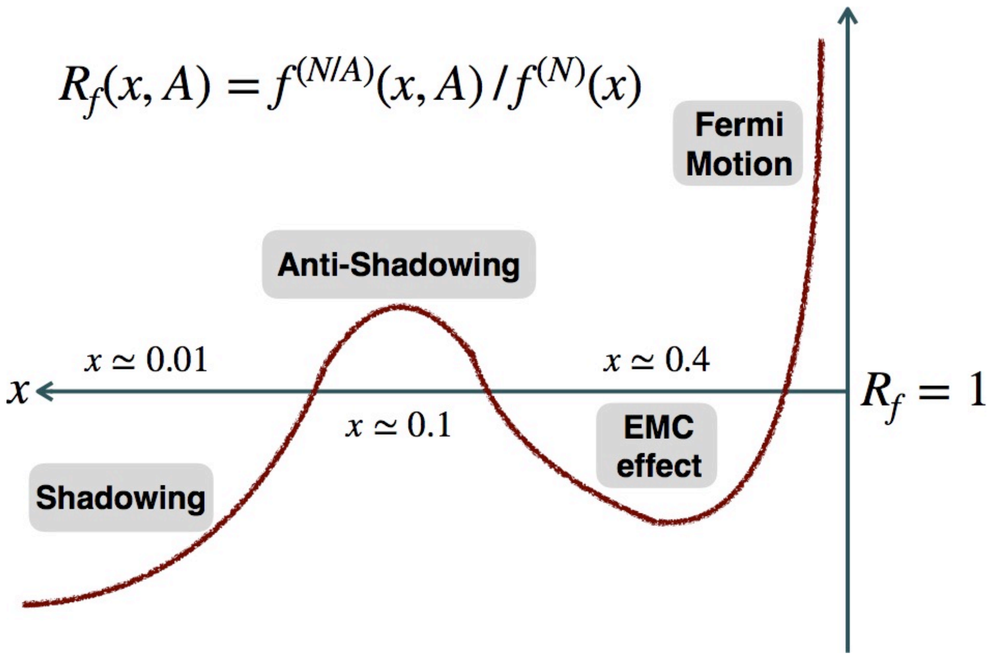
  → Impact of EIC pseudo-data
- Summary and outlook

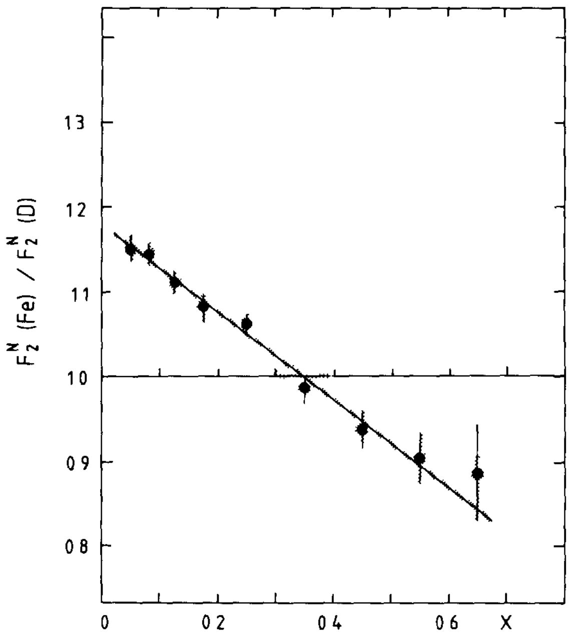
Motivation

- Deep-inelastic scattering of leptons from nuclei cannot be described by free nucleon formalism (EMC effect), e.g.

DIS structure functions: $$F_2^A \neq ZF_2^p + (A-Z)F_2^n$$

- Parton structure is modified in bound systems





J.J. Aubert et. al. Phys. Lett. B 123B (1983)

- Mechanisms that originate nuclear effects still not well understood

nPDFs from Global QCD Analyses

- Empirical determination of in-medium modifications to parton structure
- Relevant for interpretation of heavy ion collisions at RHIC and LHC
- Relies on collinear factorization formalism analogous to free proton PDF fits, e.g.

$$d\sigma^A(x,Q^2) \simeq \sum_f \int_x^{M_A/M} \frac{d\xi}{\xi} f^A(\frac{x}{\xi},Q^2) d\hat{\sigma}_f(\xi,Q^2)$$

nPDFs from Global QCD Analyses

- Empirical determination of in-medium modifications to parton structure
- Relevant for interpretation of heavy ion collisions at RHIC and LHC
- Relies on collinear factorization formalism analogous to free proton PDF fits, e.g.

$$d\sigma^A(x,Q^2) \simeq \sum_f \int_x^{M_A/M} \frac{d\xi}{\xi}\, f^A(\frac{x}{\xi},Q^2)\, d\hat{\sigma}_f(\xi,Q^2)$$

Parton distribution function (PDF) of nucleus with atomic mass *A*

Hard scattering

nPDFs from Global QCD Analyses

- Empirical determination of in-medium modifications to parton structure
- Relevant for interpretation of heavy ion collisions at RHIC and LHC
- Relies on collinear factorization formalism analogous to free proton PDF fits, e.g.

$$d\sigma^A(x,Q^2) \simeq \sum_f \int_x^{M_A/M} \frac{d\xi}{\xi} f^A(\frac{x}{\xi},Q^2) d\hat{\sigma}_f(\xi,Q^2)$$

Parton distribution function (PDF) of nucleus with atomic mass A

Hard scattering

- Scattering from nuclei treated as incoherent from single bound nucleon

$$d\sigma^A(x,Q^2) \simeq Z d\sigma^{p/A} + (A-Z) d\sigma^{n/A}$$

$$d\sigma^{N/A}(x,Q^2) \simeq \sum_f \int_x^1 \frac{d\xi}{\xi} f^{N/A}(\frac{x}{\xi},Q^2) d\hat{\sigma}_f(\xi,Q^2)$$

- nPDFs are parameterized and fitted to global lepton-nucleus and hadron-nucleus scattering measurements

nPDFs from Global QCD Analyses

- Can reveal onset of non-linear evolution effects at low x and Q^2

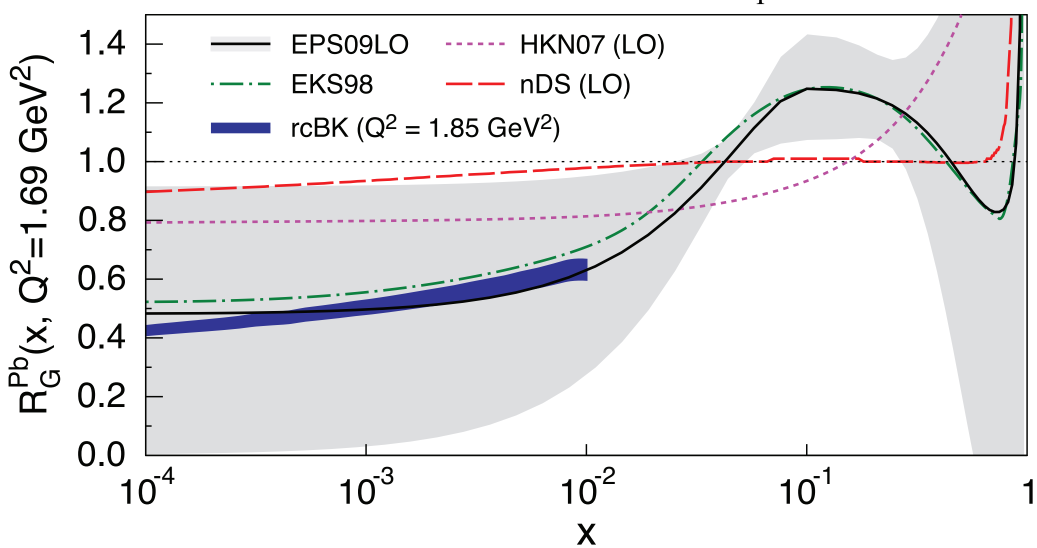
  → Enhancement for heavier nuclei – saturation region expected to begin at larger x

- Effect given in ratio of nuclear to proton PDFs:

$$R_G^A = \frac{g^{(p/A)}(x, Q^2, A)}{g^{(p)}(x, Q^2)}$$

- Existing nuclear PDF information contains large uncertainties in the saturation region

A. Accardi et. al. EIC White Paper *arXiv:1212.1701*



- Precise determination of nuclear PDFs is highly relevant for Electron-Ion Collider

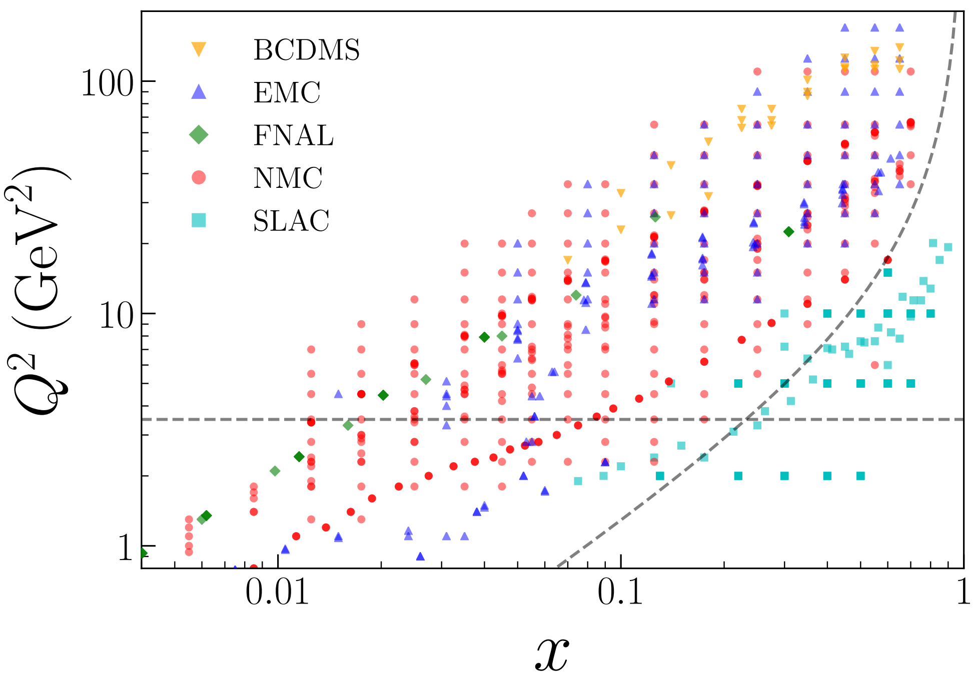
nNNPDF1.0 Analysis

- Includes all available neutral current DIS data from CERN, SLAC, and FNAL experiments

→ Kinematic cuts: $W^2 > 12.5\ \mathrm{GeV}^2$

$Q^2 > 3.5\ \mathrm{GeV}^2$

- Significant range in atomic mass values (A from 2 to 208)
- 451 total data points



Experiment	A_1/A_2	N_{dat}
SLAC E-139	^{4}He/^{2}D	3
NMC 95, re.	^{4}He/^{2}D	13
NMC 95	^{6}Li/^{2}D	12
SLAC E-139	^{9}Be/^{2}D	3
NMC 96	^{9}Be/^{12}C	14
EMC 88, EMC 90	^{12}C/^{2}D	12
SLAC E-139	^{12}C/^{2}D	2
NMC 95, NMC 95, re.	^{12}C/^{2}D	26
FNAL E665	^{12}C/^{2}D	3
NMC 95, re.	^{12}C/^{6}Li	9
BCDMS 85	^{14}N/^{2}D	9
SLAC E-139	^{27}Al/^{2}D	3
NMC 96	^{27}Al/^{12}C	14
SLAC E-139	^{40}Ca/^{2}D	2
NMC 95, re.	^{40}Ca/^{2}D	12
EMC 90	^{40}Ca/^{2}D	3
FNAL E665	^{40}Ca/^{2}D	3
NMC 95, re.	^{40}Ca/^{6}Li	9
NMC 96	^{40}Ca/^{12}C	23
EMC 87	^{56}Fe/^{2}D	58
SLAC E-139	^{56}Fe/^{2}D	8
NMC 96	^{56}Fe/^{12}C	14
BCDMS 85, BCDMS 87	^{56}Fe/^{2}D	16
EMC 88, EMC 93	^{64}Cu/^{2}D	27
SLAC E-139	^{108}Ag/^{2}D	2
EMC 88	^{119}Sn/^{2}D	8
NMC 96, Q^2 dependence	^{119}Sn/^{12}C	119
FNAL E665	^{131}Xe/^{2}D	4
SLAC E-139	^{197}Au/^{2}D	3
FNAL E665	^{208}Pb/^{2}D	3
NMC 96	^{208}Pb/^{12}C	14
Total		**451**

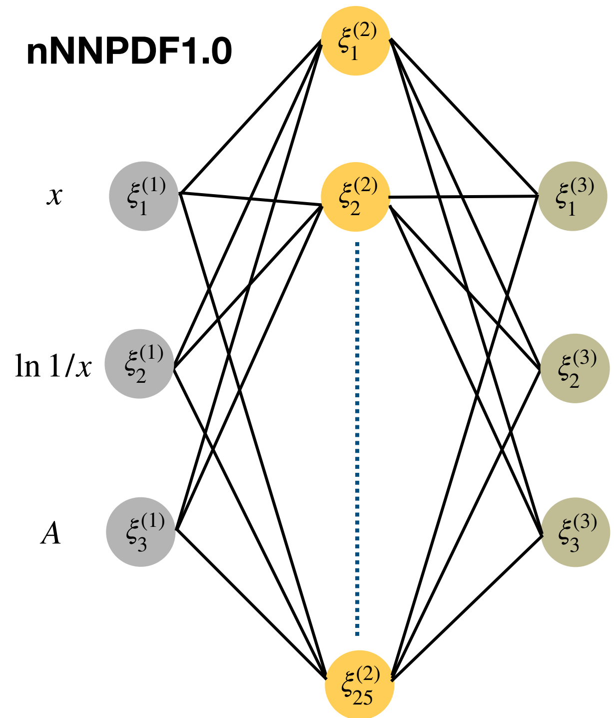
nPDF Parameterization

- Single NN with architecture 3-25-3
- Input scale: $Q_0 = 1$ GeV
- PDFs parameterized with NN output multiplied by preprocessing function

$$x\Sigma(x, Q_0, A) = x^{-\alpha_\Sigma}(1-x)^{\beta_\sigma}\, \xi_1^{(3)}(x, A)$$

$$xT_8(x, Q_0, A) = x^{-\alpha_{T_8}}(1-x)^{\beta_{T_8}}\, \xi_2^{(3)}(x, A)$$

$$xg(x, Q_0, A) = B_g x^{-\alpha_g}(1-x)^{\beta_g}\, \xi_3^{(3)}(x, A)$$



$$g(x, Q_0, A) = B_g x^{-\alpha_g}(1-x)^{\beta_g}\, \xi_1^{(3)}$$

$$\Sigma(x, Q_0, A) = x^{-\alpha_\Sigma}(1-x)^{\beta_\Sigma}\, \xi_2^{(3)}$$

$$T_8(x, Q_0, A) = x^{-\alpha_{T_8}}(1-x)^{\beta_{T_8}}\, \xi_3^{(3)}$$

$$\Sigma \equiv \sum_q q^+ = \sum_q q + \bar{q} \qquad T_8 \equiv u^+ + d^+ - 2s^+$$

- Exponents treated as free parameters
- Momentum Sum Rule: $\int_0^1 dx x(\Sigma(x, A) + g(x, A)) = 1 \;\rightarrow\; B_g = \dfrac{1 - \int_0^1 dx x \Sigma(x, A)}{\int_0^1 dx x g(x, A)}$
- Parameters optimized by stochastic gradient descent in TensorFlow (open-source ML software library)

nPDF Optimization

- Minimizing the cost function:

$$\chi^2 \equiv \sum_{i,j=1}^{N_{\rm dat}} \left(R_i^{(\rm exp)} - R_i^{(\rm th)}(\{f_m\})\right) (\mathrm{cov}_{\mathrm{t}_0})^{-1}_{ij} \left(R_j^{(\rm exp)} - R_j^{(\rm th)}(\{f_m\})\right)$$

$$+ \lambda \sum_{m=g,\Sigma,T_8} \sum_{l=1}^{N_x} \left(f_m(x_l, Q_0, A) - f_m^{(p+n)/2}(x_l, Q_0)\right)^2$$

nPDF Optimization

- Minimizing the cost function:

$$\chi^2 \equiv \sum_{i,j=1}^{N_{\mathrm{dat}}} \left(R_i^{(\mathrm{exp})} - R_i^{(\mathrm{th})}(\{f_m\}) \right) (\mathrm{cov}_{t_0})_{ij}^{-1} \left(R_j^{(\mathrm{exp})} - R_j^{(\mathrm{th})}(\{f_m\}) \right)$$

$$+ \lambda \sum_{m=g,\Sigma,T_8} \sum_{l=1}^{N_x} \left(f_m(x_l, Q_0, A) - f_m^{(p+n)/2}(x_l, Q_0) \right)^2$$

Artificial data replicas – experimental measurements smeared by Gaussian

$$\mathcal{O}_i^{(\mathrm{art})(k)} = S_{i,N}^{(k)} \mathcal{O}_i^{(\mathrm{exp})} \left(1 + r_i^{(k)} \sigma_i^{(\mathrm{stat})} + \sum_{\alpha=1}^{N_{\mathrm{sys}}} r_{i,\alpha}^{(k)} \sigma_{i,\alpha}^{(\mathrm{sys})} \right), \quad k = 1, \ldots, N_{\mathrm{rep}}$$

nPDF Optimization

- Minimizing the cost function:

$$\chi^2 \equiv \sum_{i,j=1}^{N_{\rm dat}} \left(R_i^{(\rm exp)} - R_i^{(\rm th)}(\{f_m\}) \right) (\mathrm{cov}_{\mathrm{t}_0})_{ij}^{-1} \left(R_j^{(\rm exp)} - R_j^{(\rm th)}(\{f_m\}) \right)$$
$$+ \lambda \sum_{m=g,\Sigma,T_8} \sum_{l=1}^{N_x} \left(f_m(x_l, Q_0, A) - f_m^{(p+n)/2}(x_l, Q_0) \right)^2$$

Artificial data replicas – experimental measurements smeared by Gaussian

Theoretical predictions (functions of the parameterized PDFs)

nPDF Optimization

- Minimizing the cost function:

$$\chi^2 \equiv \sum_{i,j=1}^{N_{\rm dat}} \left(R_i^{\rm (exp)} - R_i^{\rm (th)}(\{f_m\}) \right) \left({\rm cov}_{t_0} \right)^{-1}_{ij} \left(R_j^{\rm (exp)} - R_j^{\rm (th)}(\{f_m\}) \right)$$

$$+ \lambda \sum_{m=g,\Sigma,T_8} \sum_{l=1}^{N_x} \left(f_m(x_l, Q_0, A) - f_m^{(p+n)/2}(x_l, Q_0) \right)^2$$

Artificial data replicas – experimental measurements smeared by Gaussian

Theoretical predictions (functions of the parameterized PDFs)

Covariance matrix – encodes all uncorrelated and correlated experimental uncertainties

$$({\rm cov}_{t_0})_{ij}^{\rm (exp)} \equiv \left(\sigma_i^{\rm (stat)} R_i^{\rm (exp)} \right)^2 \delta_{ij} + \left(\sum_{\alpha=1}^{N_{\rm add}} \sigma_{i,\alpha}^{\rm (sys,a)} \sigma_{j,\alpha}^{\rm (sys,a)} R_i^{\rm (exp)} R_j^{\rm (exp)} + \sum_{\beta=1}^{N_{\rm mult}} \sigma_{i,\beta}^{\rm (sys,m)} \sigma_{j,\beta}^{\rm (sys,m)} R_i^{\rm (th,0)} R_j^{\rm (th,0)} \right)$$

nPDF Optimization

- Minimizing the cost function:

$$\chi^2 \equiv \sum_{i,j=1}^{N_{\text{dat}}} \left(R_i^{(\text{exp})} - R_i^{(\text{th})}(\{f_m\}) \right) (\text{cov}_{t_0})_{ij}^{-1} \left(R_j^{(\text{exp})} - R_j^{(\text{th})}(\{f_m\}) \right)$$

$$+ \lambda \sum_{m=g,\Sigma,T_8} \sum_{l=1}^{N_x} \left(f_m(x_l, Q_0, A) - f_m^{(p+n)/2}(x_l, Q_0) \right)^2$$

Artificial data replicas – experimental measurements smeared by Gaussian

Theoretical predictions (functions of the parameterized PDFs)

Covariance matrix – encodes all uncorrelated and correlated experimental uncertainties

$$(\text{cov}_{t_0})_{ij}^{(\text{exp})} \equiv \left(\sigma_i^{(\text{stat})} R_i^{(\text{exp})} \right)^2 \delta_{ij} + \left(\sum_{\alpha=1}^{N_{\text{add}}} \sigma_{i,\alpha}^{(\text{sys,a})} \sigma_{j,\alpha}^{(\text{sys,a})} R_i^{(\text{exp})} R_j^{(\text{exp})} + \sum_{\beta=1}^{N_{\text{mult}}} \sigma_{i,\beta}^{(\text{sys,m})} \sigma_{j,\beta}^{(\text{sys,m})} R_i^{(\text{th},0)} R_j^{(\text{th},0)} \right)$$

t_0 prescription: multiply correlated multiplicative uncertainties by central theory values from previous fit iterations (iterated until convergence)

nPDF Optimization

- Minimizing the cost function:

$$\chi^2 \equiv \sum_{i,j=1}^{N_{\rm dat}} \left(R_i^{(\rm exp)} - R_i^{(\rm th)}(\{f_m\})\right) \left(\mathrm{cov}_{t_0}\right)_{ij}^{-1} \left(R_j^{(\rm exp)} - R_j^{(\rm th)}(\{f_m\})\right)$$
$$+ \lambda \sum_{m=g,\Sigma,T_8} \sum_{l=1}^{N_x} \left(f_m(x_l, Q_0, A) - f_m^{(p+n)/2}(x_l, Q_0)\right)^2$$

Artificial data replicas – experimental measurements smeared by Gaussian

Theoretical predictions (functions of the parameterized PDFs)

Covariance matrix – encodes all uncorrelated and correlated experimental uncertainties

- Many fits performed to obtain representative MC sample distribution

$$E[\mathcal{O}(\vec{a})] = \sum_k w_k \mathcal{O}(\vec{a}_k) \qquad V[\mathcal{O}(\vec{a})] = \sum_k w_k (\mathcal{O}(\vec{a}_k) - E[\mathcal{O}])^2$$

- Fits are cross-validated to prevent over-fitting

nPDF Optimization

- Minimizing the cost function:

$$\chi^2 \equiv \sum_{i,j=1}^{N_{\rm dat}} \left(R_i^{(\rm exp)} - R_i^{(\rm th)}(\{f_m\}) \right) \left(\mathrm{cov}_{\mathrm{t}_0}\right)_{ij}^{-1} \left(R_j^{(\rm exp)} - R_j^{(\rm th)}(\{f_m\}) \right)$$
$$+ \lambda \sum_{m=g,\Sigma,T_8} \sum_{l=1}^{N_x} \left(f_m(x_l, Q_0, A) - f_m^{(p+n)/2}(x_l, Q_0) \right)^2$$

Boundary condition (imposed for x from 10^{-3} to 0.7)

nPDF Optimization

- Minimizing the cost function:

$$\chi^2 \equiv \sum_{i,j=1}^{N_{\rm dat}} \left(R_i^{(\rm exp)} - R_i^{(\rm th)}(\{f_m\}) \right) ({\rm cov}_{t_0})_{ij}^{-1} \left(R_j^{(\rm exp)} - R_j^{(\rm th)}(\{f_m\}) \right)$$
$$+ \lambda \sum_{m=g,\Sigma,T_8} \sum_{l=1}^{N_x} \left(f_m(x_l, Q_0, A) - f_m^{(p+n)/2}(x_l, Q_0) \right)^2$$

Boundary condition (imposed for *x* from 10^{-3} to 0.7)

→ Free nucleon PDFs must be reproduced at A=1

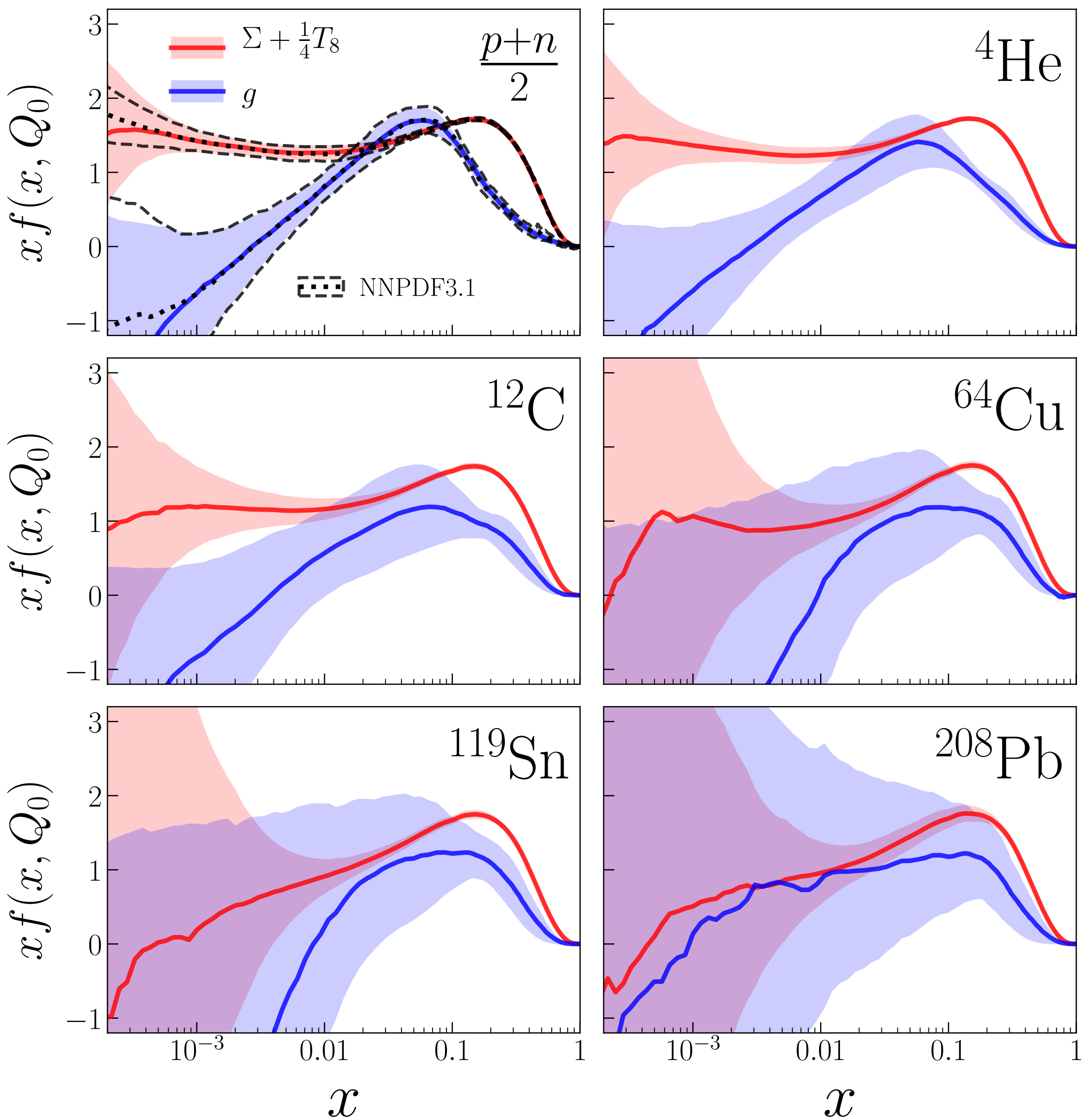
$$f(x, Q, A=1) = \frac{1}{2}\left[f_p(x, Q^2) + f_n(x, Q^2) \right]$$

→ NNPDF3.1 proton PDF fits are used as baseline (consistent methodology and theoretical assumptions)

→ Central values and *uncertainties* reproduced at minimization level – "simultaneous" fit of proton and nuclear PDFs

→ Correlations from proton PDFs are accounted for!

nNNPDF1.0 Results



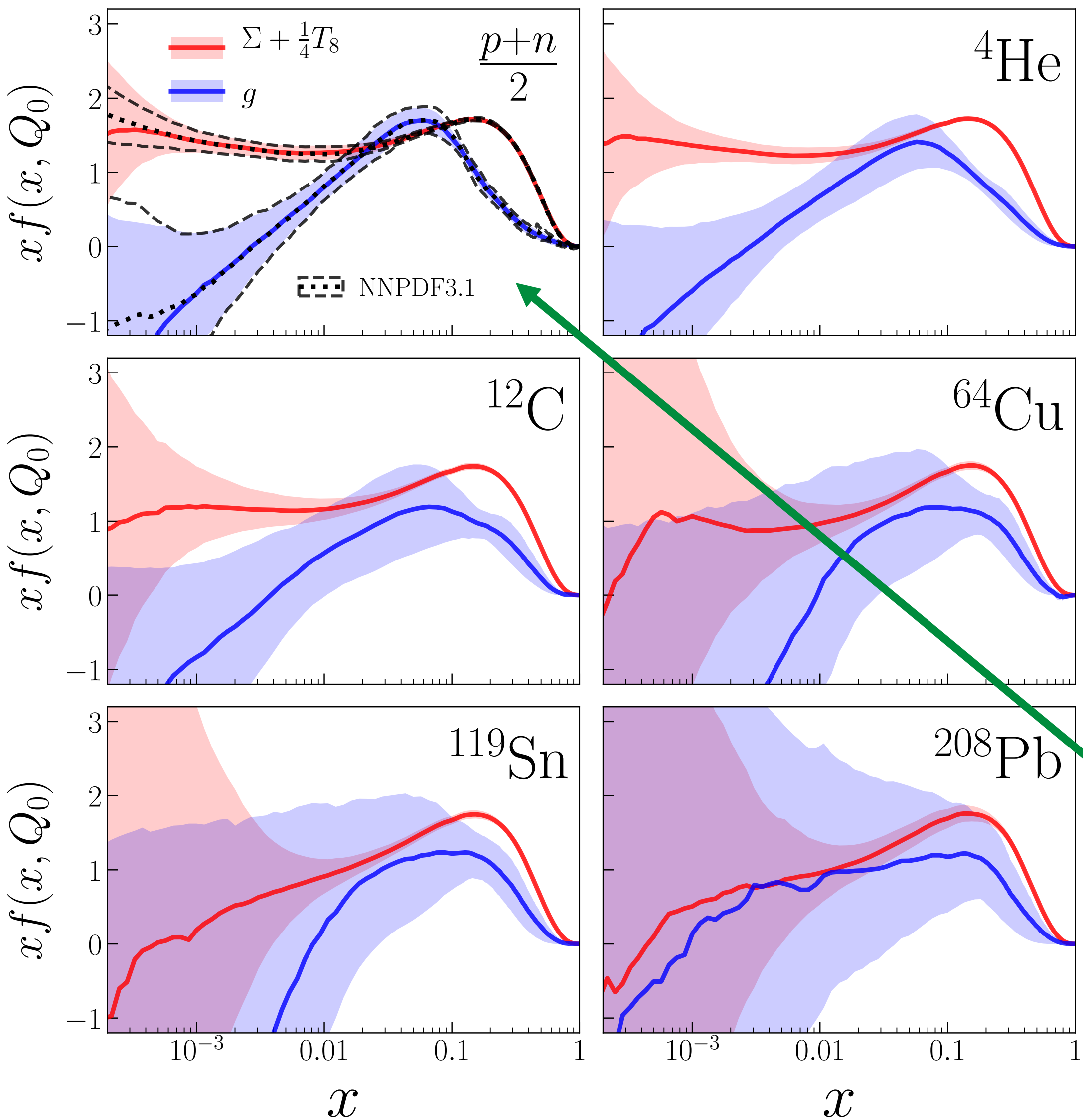
Uncertainties computed as 90% CL range

Only linear combination of quark singlet and octet distributions constrained by NC DIS

$$\Sigma = \sum_i^{n_f}(f_i + \bar{f}_i) = \sum_i^{n_f} f_i^+$$

$$T_8 = u^+ + d^+ - 2s^+$$

nNNPDF1.0 Results



Uncertainties computed as 90% CL range

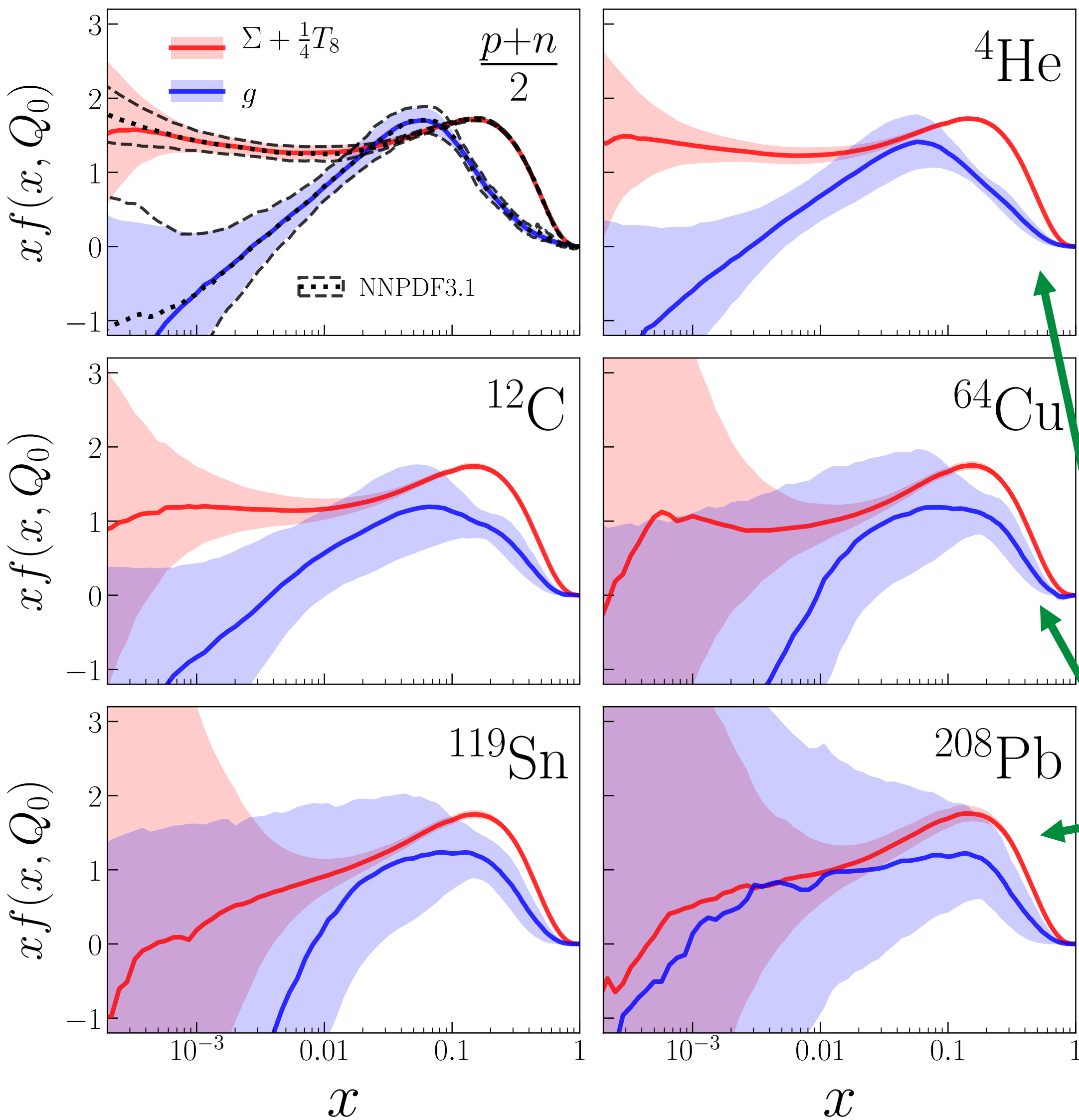
Only linear combination of quark singlet and octet distributions constrained by NC DIS

$$\Sigma = \sum_i^{n_f} (f_i + \bar{f}_i) = \sum_i^{n_f} f_i^+$$

$$T_8 = u^+ + d^+ - 2s^+$$

NNPDF3.1 central value and uncertainties reproduced

nNNPDF1.0 Results



Uncertainties computed as 90% CL range

Only linear combination of quark singlet and octet distributions constrained by NC DIS

$$\Sigma = \sum_i^{n_f}(f_i + \bar{f}_i) = \sum_i^{n_f} f_i^+$$

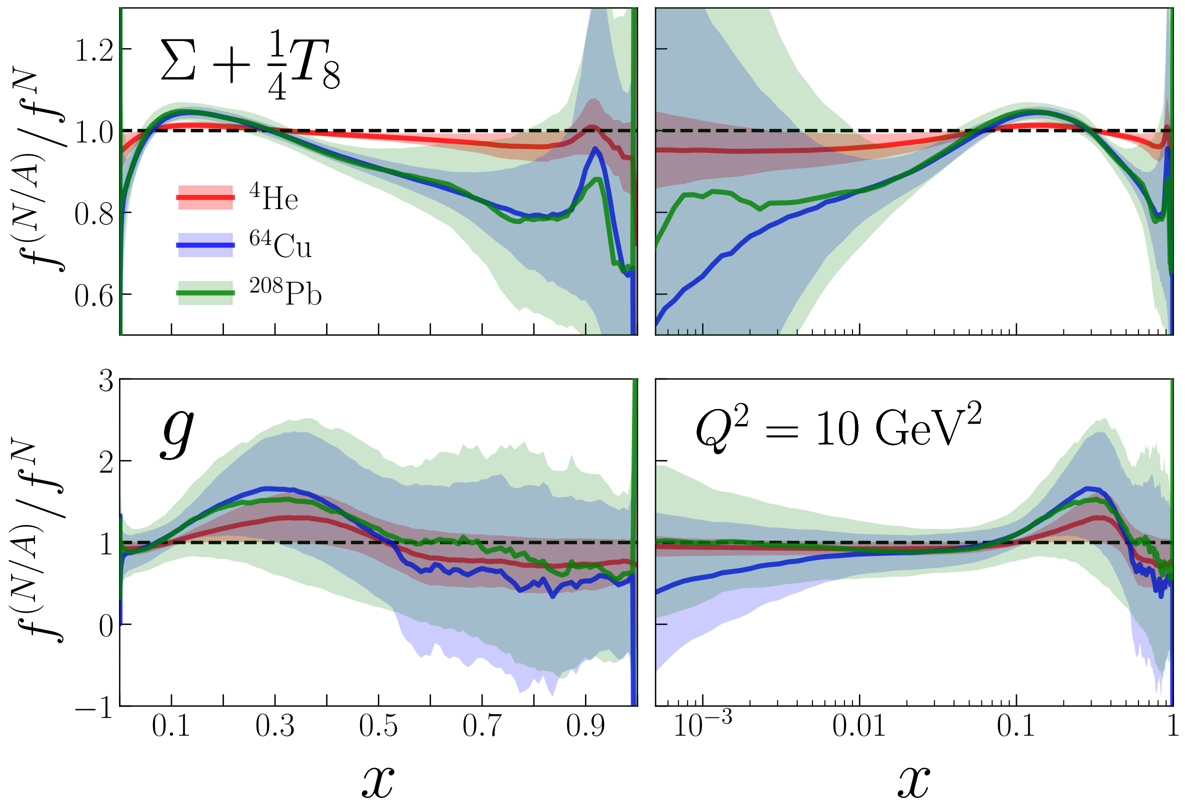
$$T_8 = u^+ + d^+ - 2s^+$$

Increasing uncertainties with A – effect of boundary condition

nNNPDF1.0 Results

Ratio to A=1 result – correlations between nPDFs included

$$R_f^{(k)} = \frac{f^{(N/A)(k)}(x, Q^2, A)}{f^{(N)(k)}(x, Q^2)}$$

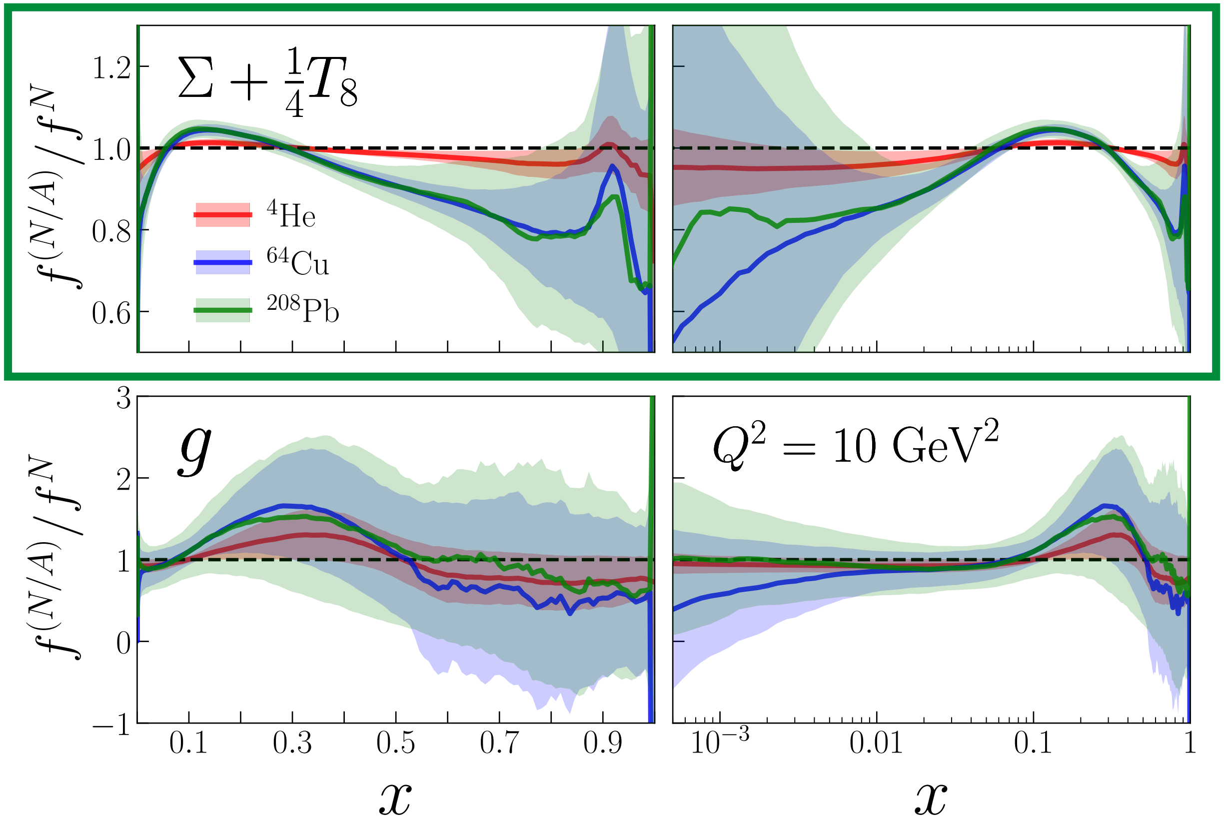


nNNPDF1.0 Results

Ratio to A=1 result – correlations between nPDFs included

$$R_f^{(k)} = \frac{f^{(N/A)(k)}(x, Q^2, A)}{f^{(N)(k)}(x, Q^2)}$$

Nuclear effects visible in quark combination – negligible for A=4



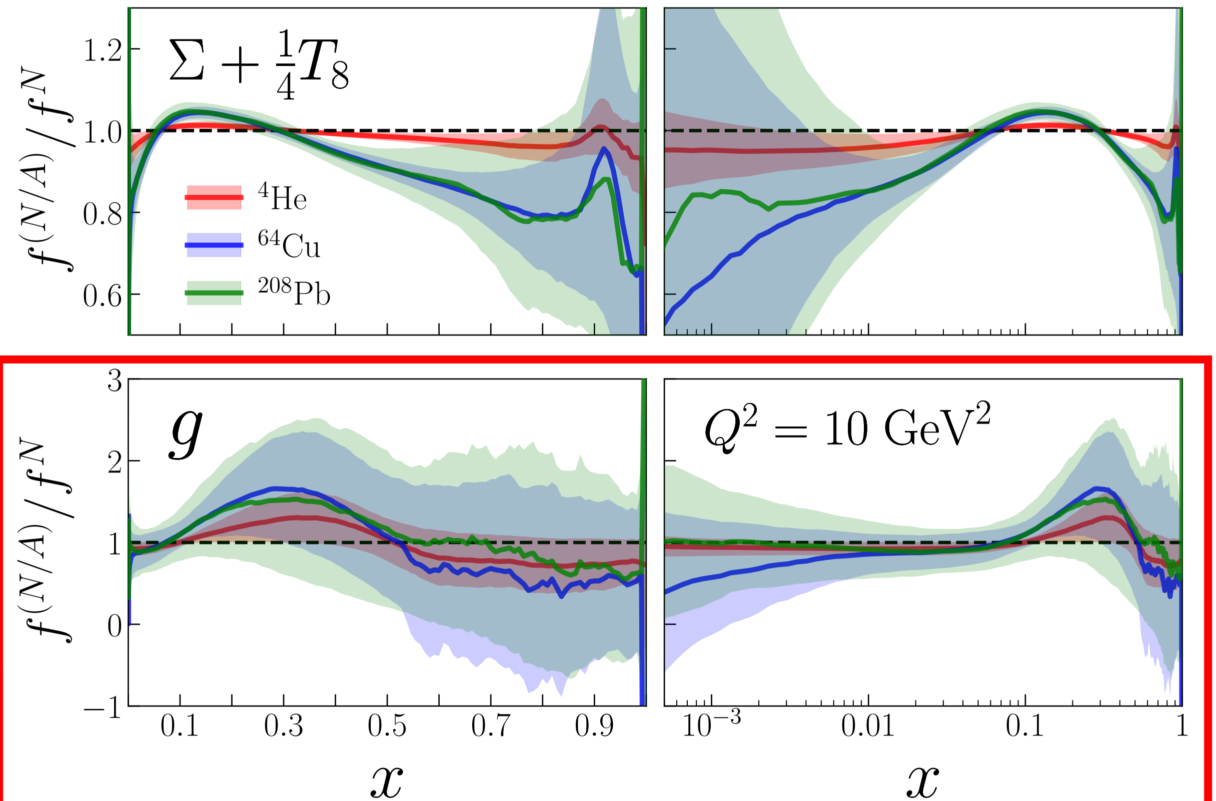
nNNPDF1.0 Results

Ratio to A=1 result – correlations between nPDFs included

$$R_f^{(k)} = \frac{f^{(N/A)(k)}(x, Q^2, A)}{f^{(N)(k)}(x, Q^2)}$$

Nuclear effects visible in quark combination – negligible for A=4

Larger uncertainties for gluon distribution – consistent with unity

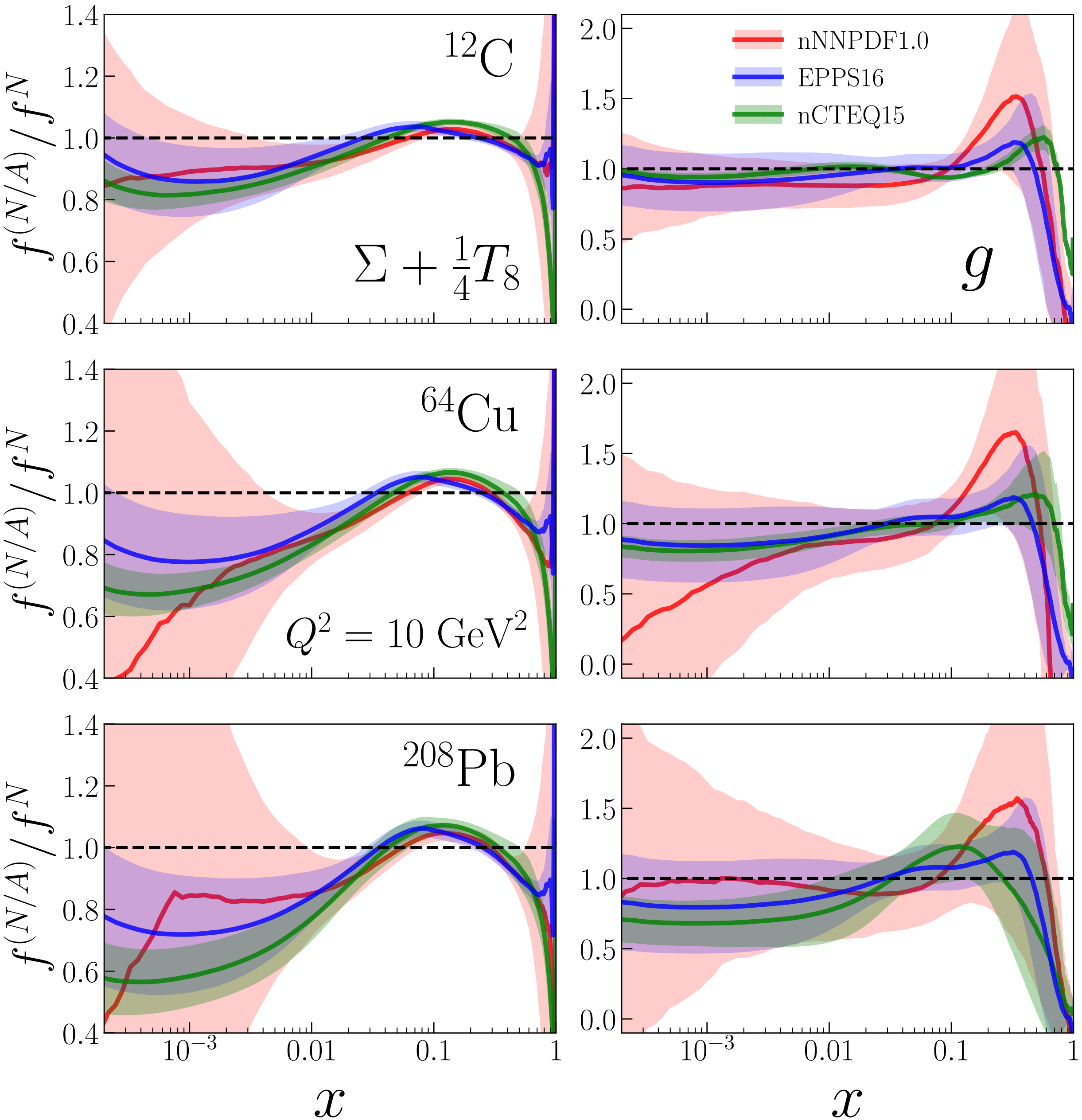


Nuclear PDFs

All distributions normalized by nNNPDF1.0 A=1 distribution

90% CL computed with Hessian method for nCTEQ and EPPS uncertainties

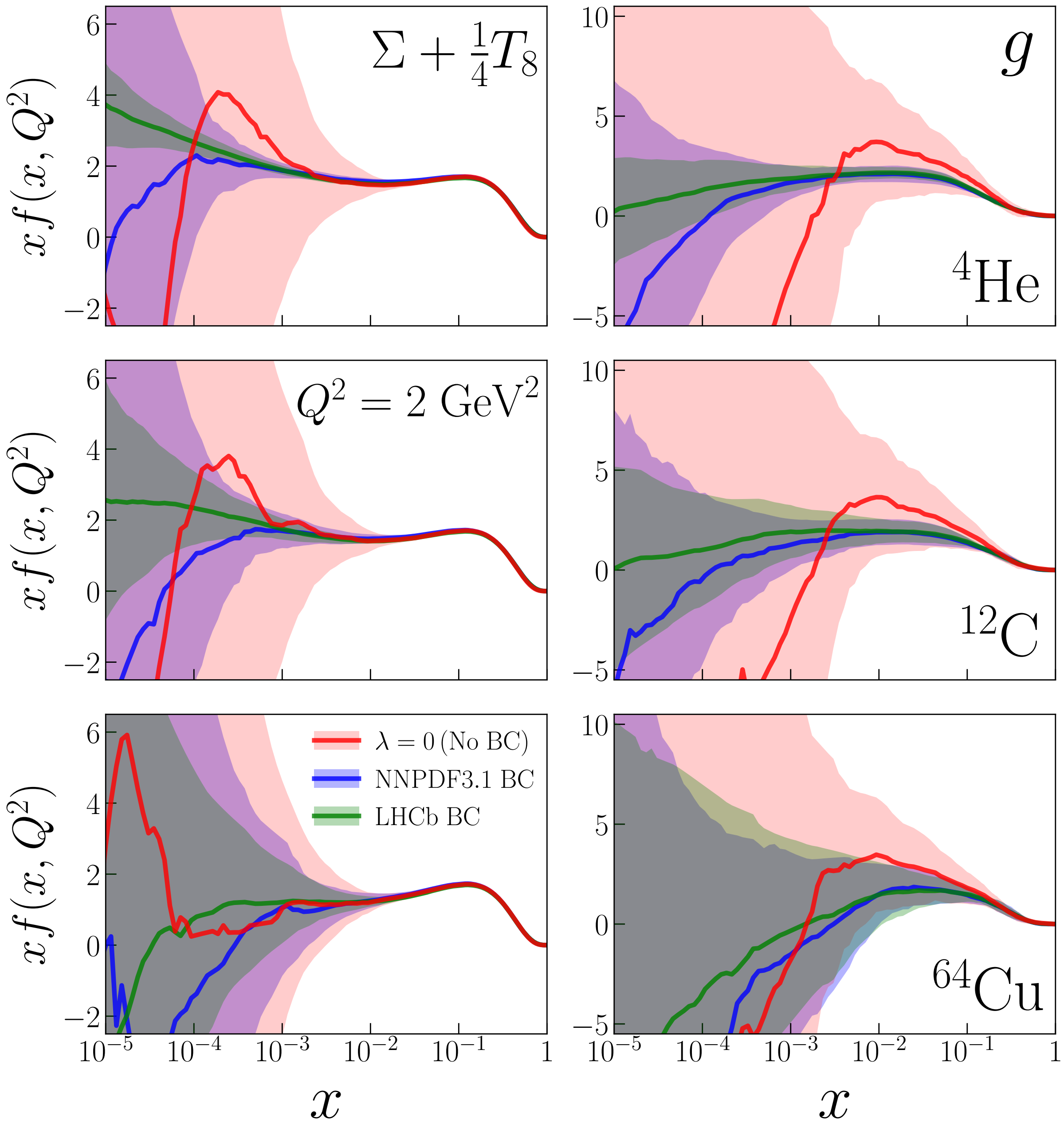
Significant differences in uncertainties



Nuclear PDFs

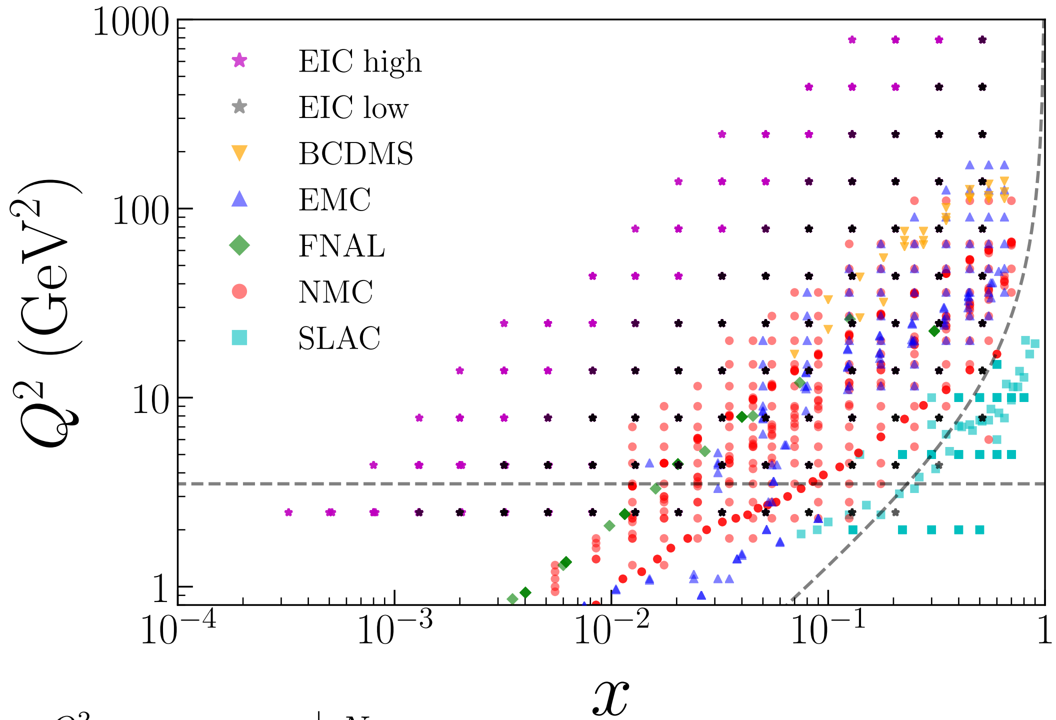
Can test other boundary conditions – NNPDF3.0+LHCb PDF set with smaller uncertainties at low x

Remarkable impact from boundary condition choice – proton PDF constraints relevant for low-A nPDF extraction!



Impact of the EIC

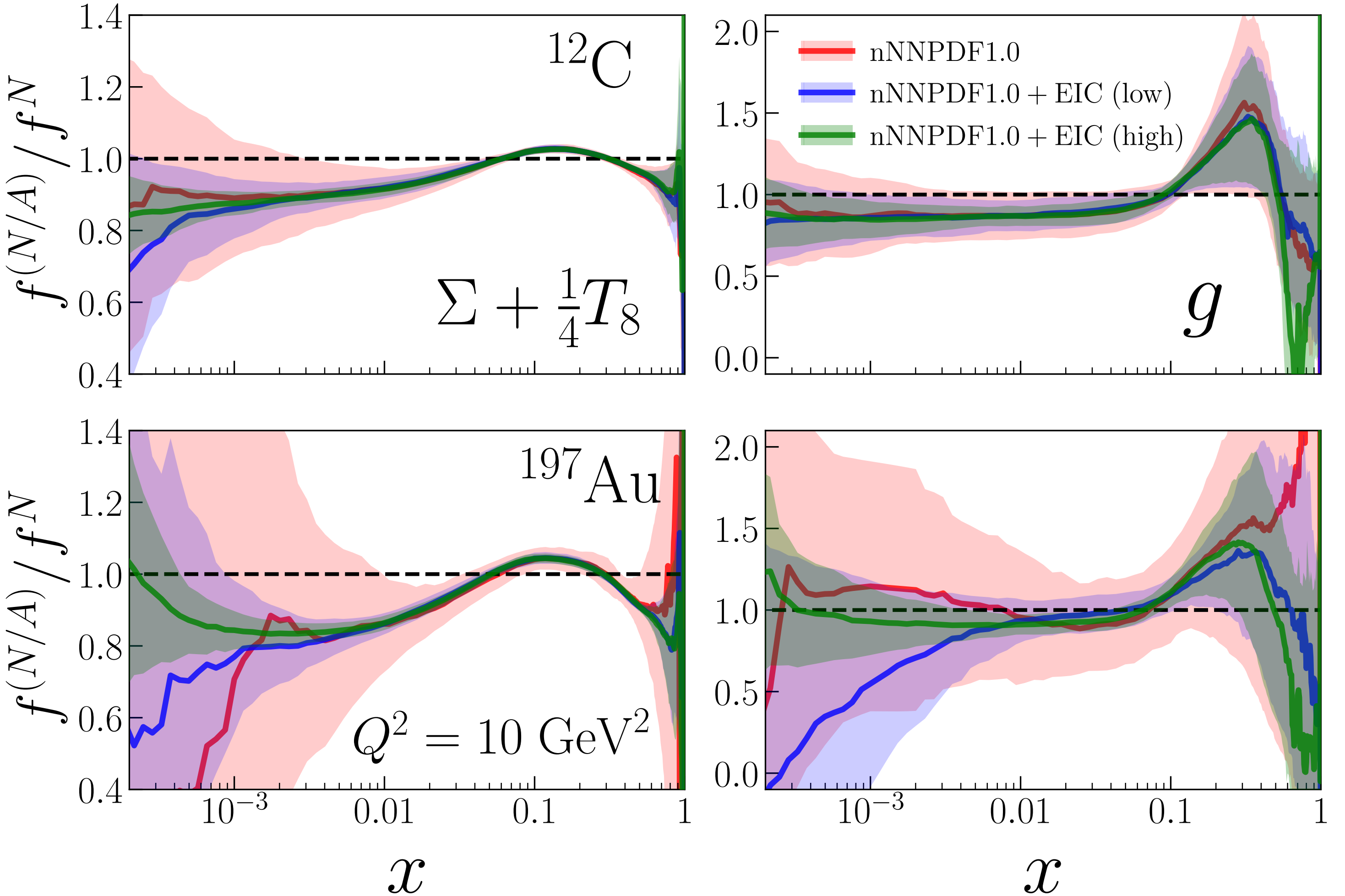
- Analysis of EIC pseudodata – extended kinematic coverage
- Two scenarios: low energy (5 GeV) vs high energy (20 GeV) electron beam



Scenario	A	E_e	E_A/A	$Q^2_{\max}$	$x_{\min}$	$N_{\rm dat}$
eRHIC_5x50C	12	5 GeV	50 GeV	440 GeV^2	0.003	50
eRHIC_5x75C	12	5 GeV	75 GeV	440 GeV^2	0.002	57
eRHIC_5x100C	12	5 GeV	100 GeV	780 GeV^2	0.001	64
eRHIC_5x50Au	197	5 GeV	50 GeV	440 GeV^2	0.003	50
eRHIC_5x75Au	197	5 GeV	75 GeV	440 GeV^2	0.002	57
eRHIC_5x100Au	197	5 GeV	100 GeV	780 GeV^2	0.001	64
eRHIC_20x50C	12	20 GeV	50 GeV	780 GeV^2	0.0008	75
eRHIC_20x75C	12	20 GeV	75 GeV	780 GeV^2	0.0005	79
eRHIC_20x100C	12	20 GeV	100 GeV	780 GeV^2	0.0003	82
eRHIC_20x50Au	197	20 GeV	50 GeV	780 GeV^2	0.0008	75
eRHIC_20x75Au	197	20 GeV	75 GeV	780 GeV^2	0.0005	79
eRHIC_20x100Au	197	20 GeV	100 GeV	780 GeV^2	0.0003	82

- Pseudodata constructed with nNNPDF1.0 PDF sets for carbon and gold nuclei
- Uncertainty projections from analysis of E.C. Aschenaur et al. [arXiv:1708.05654]

Impact of the EIC



Significant reduction of nPDF uncertainties at low-x for large A – particularly for higher energy option

Towards nNNPDF2.0

- Experimental data:

  → CC DIS data from CHORUS (Pb) & NuTeV (Fe)

  → W/Z production from LHC

- Methodology:

  → Positivity constraints (!): $\chi^2 \to \chi^2 + \lambda \max\big(0, -\mathcal{O}(x, Q^2)\big)$

  where observables O (F2, cross-sections, etc) computed for grid in x and Q^2

  → Boundary condition: NNPDF3.1 <u>without heavy nuclear target data</u>

  → Valence sum rules

Results on remaining slides are PRELIMINARY

Based on 100 replica fit

nNNPDF2.0 Preliminary Results

Dataset	χ^2	N_{data}	χ^2/N_{data}
nNMCALC	4.991	14	0.357
nNMCBEC	3.784	14	0.27
nNMCCAC	10.119	23	0.44
nNMCCALI	2.119	9	0.235
nNMCCLI	10.025	9	1.114
nNMCFEC	9.511	14	0.679
nNMCPBC	13.251	14	0.946
nNMCSNC	70.293	119	0.591
nBCDMSFED	49.685	16	3.105
nBCDMSND	24.01	9	2.668
nSLACAGD	1.366	2	0.683
nSLACALD	2.436	3	0.812
nSLACAUD	3.852	3	1.284
nSLACBED	5.29	3	1.763
nSLACCAD	2.011	2	1.006
nSLACCD	0.573	2	0.287
nSLACFED	19.97	8	2.496
nSLACHED	1.741	3	0.58
nNMCCAD	21.064	12	1.755
nNMCCD	56.053	26	2.156
nNMCLID	13.622	12	1.135
nNMCHED	14.537	13	1.118
nEMCCAD	4.212	3	1.404
nEMCCD	14.495	12	1.208
nEMCCUD	16.476	27	0.61
nEMCFED	53.479	58	0.922
nEMCSND	17.944	8	2.243
nFNALCD	2.584	3	0.861
nFNALCAD	3.649	3	1.216
nFNALPBD	6.895	3	2.298
nFNALXED	1.483	4	0.371
CHORUSNBPb	376.367	423	0.89
CHORUSNUPb	419.117	423	0.991
NTVNUDMNFe	8.627	39	0.221
NTVNBDMNFe	26.467	37	0.715
Total:	1292.098	1373	0.941

nNNPDF2.0 Preliminary Results

- Excellent agreement with CC data (chi-squareds similar to proton fits)

Dataset	χ^2	N_{data}	χ^2/N_{data}
nNMCALC	4.991	14	0.357
nNMCBEC	3.784	14	0.27
nNMCCAC	10.119	23	0.44
nNMCCALI	2.119	9	0.235
nNMCCLI	10.025	9	1.114
nNMCFEC	9.511	14	0.679
nNMCPBC	13.251	14	0.946
nNMCSNC	70.293	119	0.591
nBCDMSFED	49.685	16	3.105
nBCDMSND	24.01	9	2.668
nSLACAGD	1.366	2	0.683
nSLACALD	2.436	3	0.812
nSLACAUD	3.852	3	1.284
nSLACBED	5.29	3	1.763
nSLACCAD	2.011	2	1.006
nSLACCD	0.573	2	0.287
nSLACFED	19.97	8	2.496
nSLACHED	1.741	3	0.58
nNMCCAD	21.064	12	1.755
nNMCCD	56.053	26	2.156
nNMCLID	13.622	12	1.135
nNMCHED	14.537	13	1.118
nEMCCAD	4.212	3	1.404
nEMCCD	14.495	12	1.208
nEMCCUD	16.476	27	0.61
nEMCFED	53.479	58	0.922
nEMCSND	17.944	8	2.243
nFNALCD	2.584	3	0.861
nFNALCAD	3.649	3	1.216
nFNALPBD	6.895	3	2.298
nFNALXED	1.483	4	0.371
CHORUSNBPb	376.367	423	0.89
CHORUSNUPb	419.117	423	0.991
NTVNUDMNFe	8.627	39	0.221
NTVNBDMNFe	26.467	37	0.715
Total:	1292.098	1373	0.941

CHORUSNBPb	376.367	423	0.89
CHORUSNUPb	419.117	423	0.991
NTVNUDMNFe	8.627	39	0.221
NTVNBDMNFe	26.467	37	0.715

nNNPDF2.0 Preliminary Results

- Excellent agreement with CC data (chi-squareds similar to proton fits)

nBCDMSFED	49.685	16	3.105

nNNPDF1.0 : (0.6)

- Possible CC and NC tension?

nSLACFED	19.97	8	2.496

(1.38)

nFNALPBD	6.895	3	2.298

(1.66)

CHORUSNBPb	376.367	423	0.89
CHORUSNUPb	419.117	423	0.991
NTVNUDMNFe	8.627	39	0.221
NTVNBDMNFe	26.467	37	0.715

Dataset	χ^2	N_{data}	χ^2/N_{data}
nNMCALC	4.991	14	0.357
nNMCBEC	3.784	14	0.27
nNMCCAC	10.119	23	0.44
nNMCCALI	2.119	9	0.235
nNMCCLI	10.025	9	1.114
nNMCFEC	9.511	14	0.679
nNMCPBC	13.251	14	0.946
nNMCSNC	70.293	119	0.591
nBCDMSFED	49.685	16	3.105
nBCDMSND	24.01	9	2.668
nSLACAGD	1.366	2	0.683
nSLACALD	2.436	3	0.812
nSLACAUD	3.852	3	1.284
nSLACBED	5.29	3	1.763
nSLACCAD	2.011	2	1.006
nSLACCD	0.573	2	0.287
nSLACFED	19.97	8	2.496
nSLACHED	1.741	3	0.58
nNMCCAD	21.064	12	1.755
nNMCCD	56.053	26	2.156
nNMCLID	13.622	12	1.135
nNMCHED	14.537	13	1.118
nEMCCAD	4.212	3	1.404
nEMCCD	14.495	12	1.208
nEMCCUD	16.476	27	0.61
nEMCFED	53.479	58	0.922
nEMCSND	17.944	8	2.243
nFNALCD	2.584	3	0.861
nFNALCAD	3.649	3	1.216
nFNALPBD	6.895	3	2.298
nFNALXED	1.483	4	0.371
CHORUSNBPb	376.367	423	0.89
CHORUSNUPb	419.117	423	0.991
NTVNUDMNFe	8.627	39	0.221
NTVNBDMNFe	26.467	37	0.715
Total:	1292.098	1373	0.941

nNNPDF2.0 Preliminary Results

- Excellent agreement with CC data (chi-squareds similar to proton fits)

nBCDMSFED	49.685	16	3.105

nNNPDF1.0 : (0.6)

- Possible CC and NC tension?

nSLACFED	19.97	8	2.496

(1.38)

nFNALPBD	6.895	3	2.298

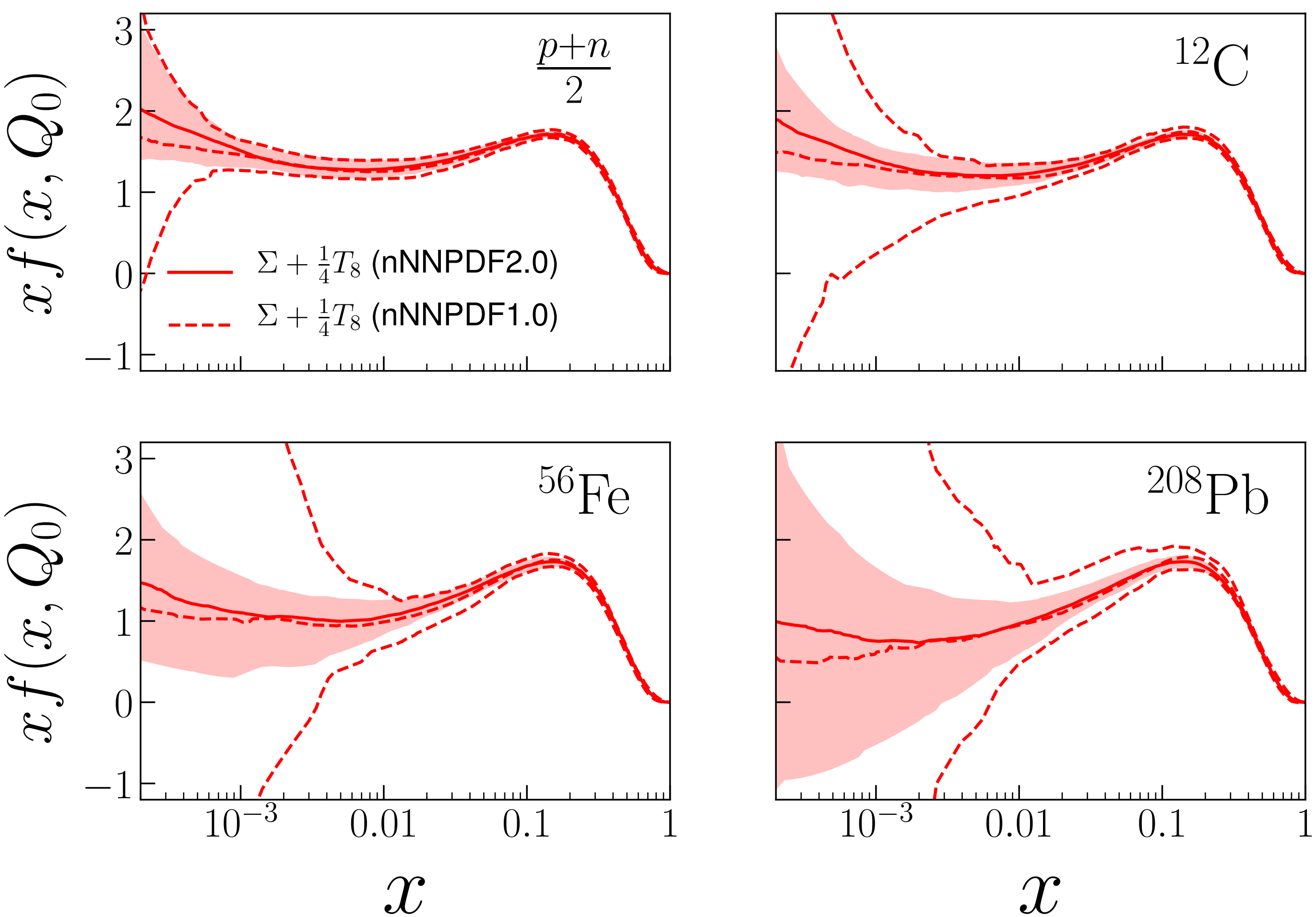
(1.66)

- Good overall chi-squared

CHORUSNBPb	376.367	423	0.89
CHORUSNUPb	419.117	423	0.991
NTVNUDMNFe	8.627	39	0.221
NTVNBDMNFe	26.467	37	0.715

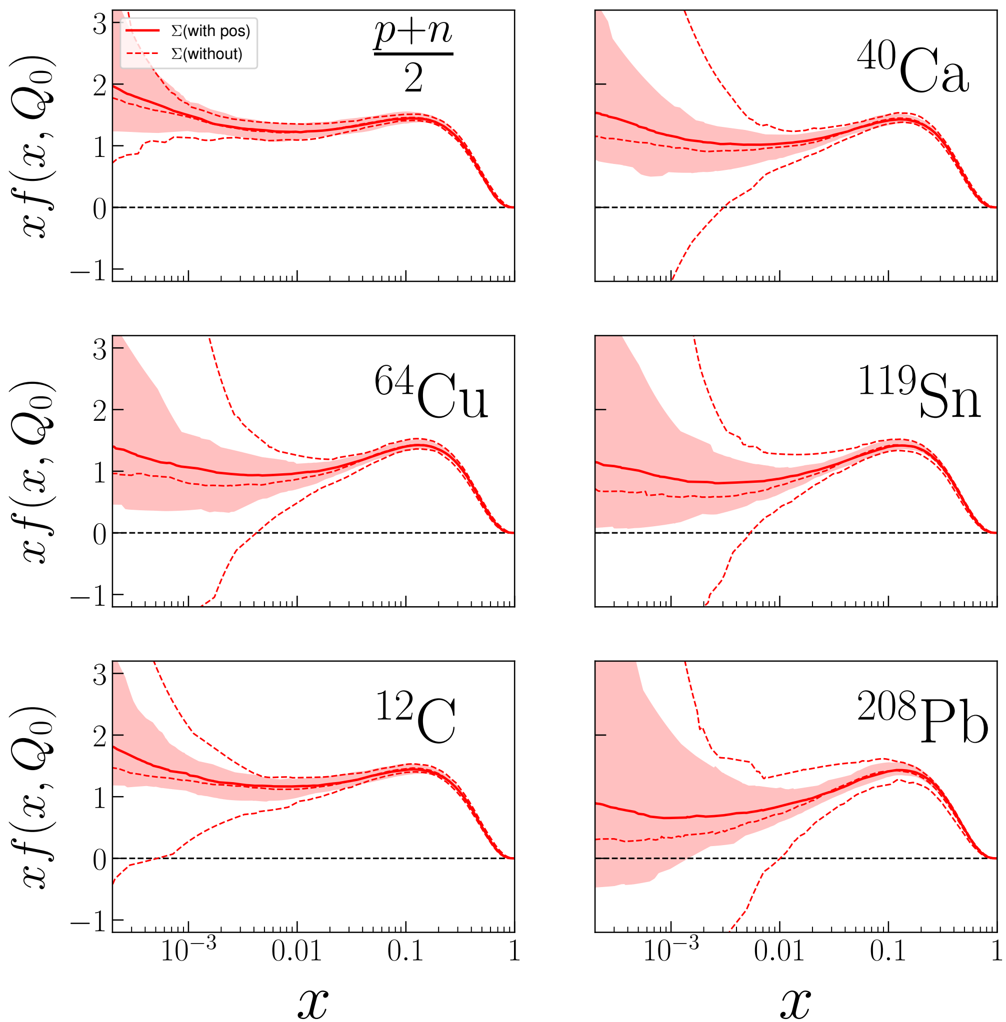
Dataset	χ^2	N_{data}	χ^2/N_{data}
nNMCALC	4.991	14	0.357
nNMCBEC	3.784	14	0.27
nNMCCAC	10.119	23	0.44
nNMCCALI	2.119	9	0.235
nNMCCLI	10.025	9	1.114
nNMCFEC	9.511	14	0.679
nNMCPBC	13.251	14	0.946
nNMCSNC	70.293	119	0.591
nBCDMSFED	49.685	16	3.105
nBCDMSND	24.01	9	2.668
nSLACAGD	1.366	2	0.683
nSLACALD	2.436	3	0.812
nSLACAUD	3.852	3	1.284
nSLACBED	5.29	3	1.763
nSLACCAD	2.011	2	1.006
nSLACCD	0.573	2	0.287
nSLACFED	19.97	8	2.496
nSLACHED	1.741	3	0.58
nNMCCAD	21.064	12	1.755
nNMCCD	56.053	26	2.156
nNMCLID	13.622	12	1.135
nNMCHED	14.537	13	1.118
nEMCCAD	4.212	3	1.404
nEMCCD	14.495	12	1.208
nEMCCUD	16.476	27	0.61
nEMCFED	53.479	58	0.922
nEMCSND	17.944	8	2.243
nFNALCD	2.584	3	0.861
nFNALCAD	3.649	3	1.216
nFNALPBD	6.895	3	2.298
nFNALXED	1.483	4	0.371
CHORUSNBPb	376.367	423	0.89
CHORUSNUPb	419.117	423	0.991
NTVNUDMNFe	8.627	39	0.221
NTVNBDMNFe	26.467	37	0.715
Total:	1292.098	1373	0.941

nNNPDF2.0 Preliminary Results



- Agreement with nNNPDF1.0
- Significant reduction in uncertainties

nNNPDF2.0 Preliminary Results

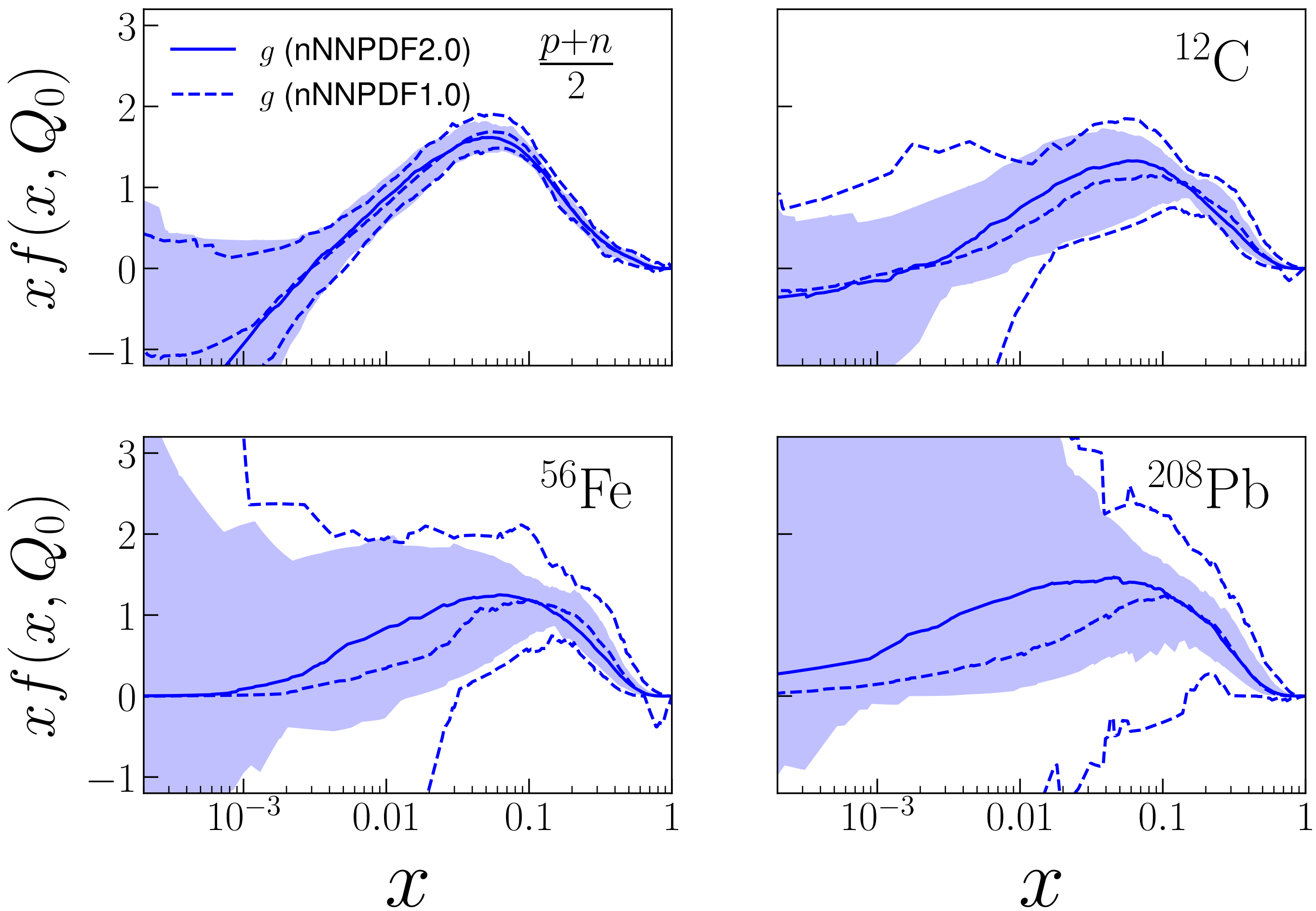


Comparison of 2.0 quark singlet with and without positivity (similar effect in octet distribution)

- Agreement with nNNPDF1.0
- Significant reduction in uncertainties

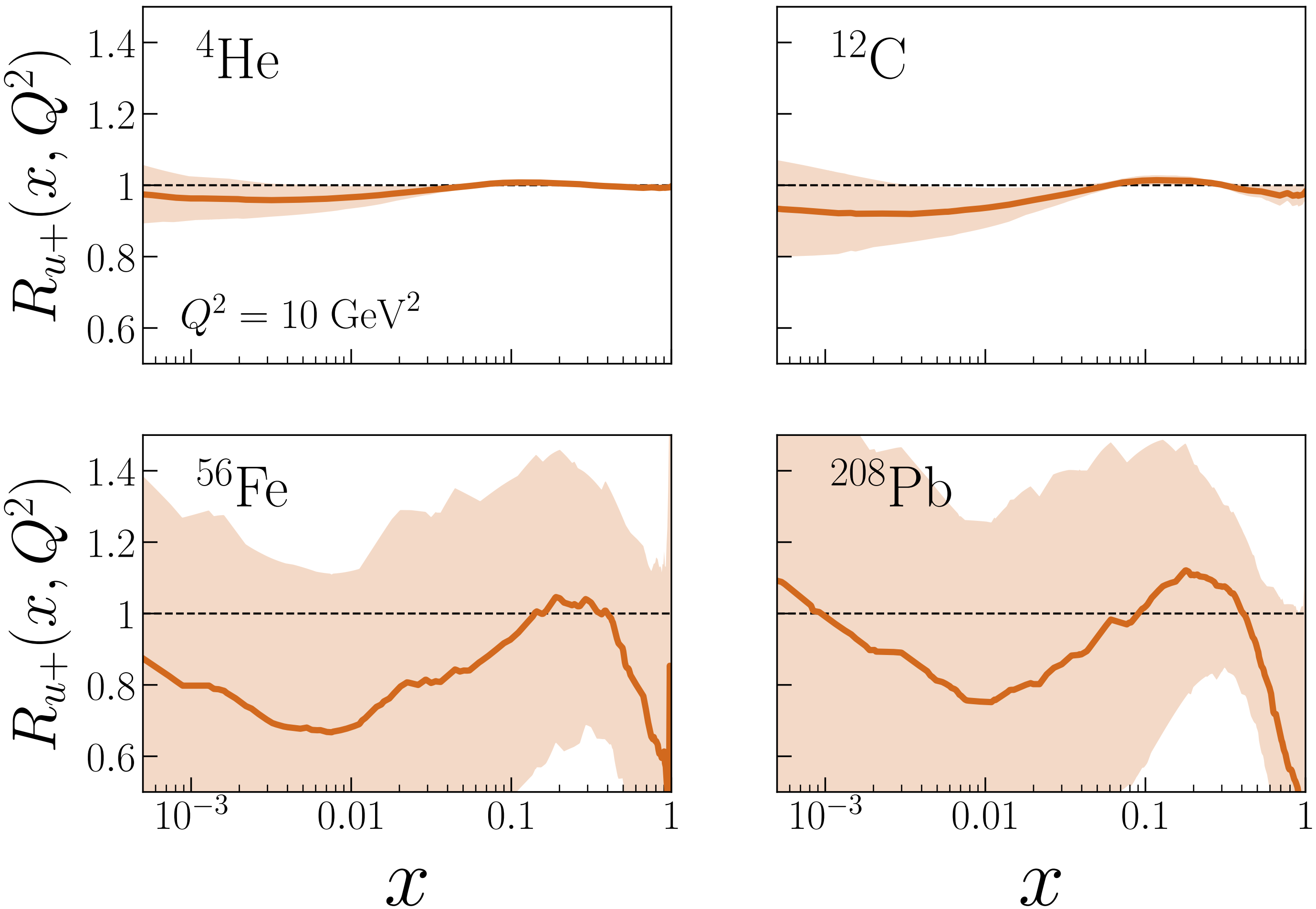
→ Mostly from positivity constraint!

nNNPDF2.0 Preliminary Results



- Agreement with nNNPDF1.0
- Significant reduction in uncertainties

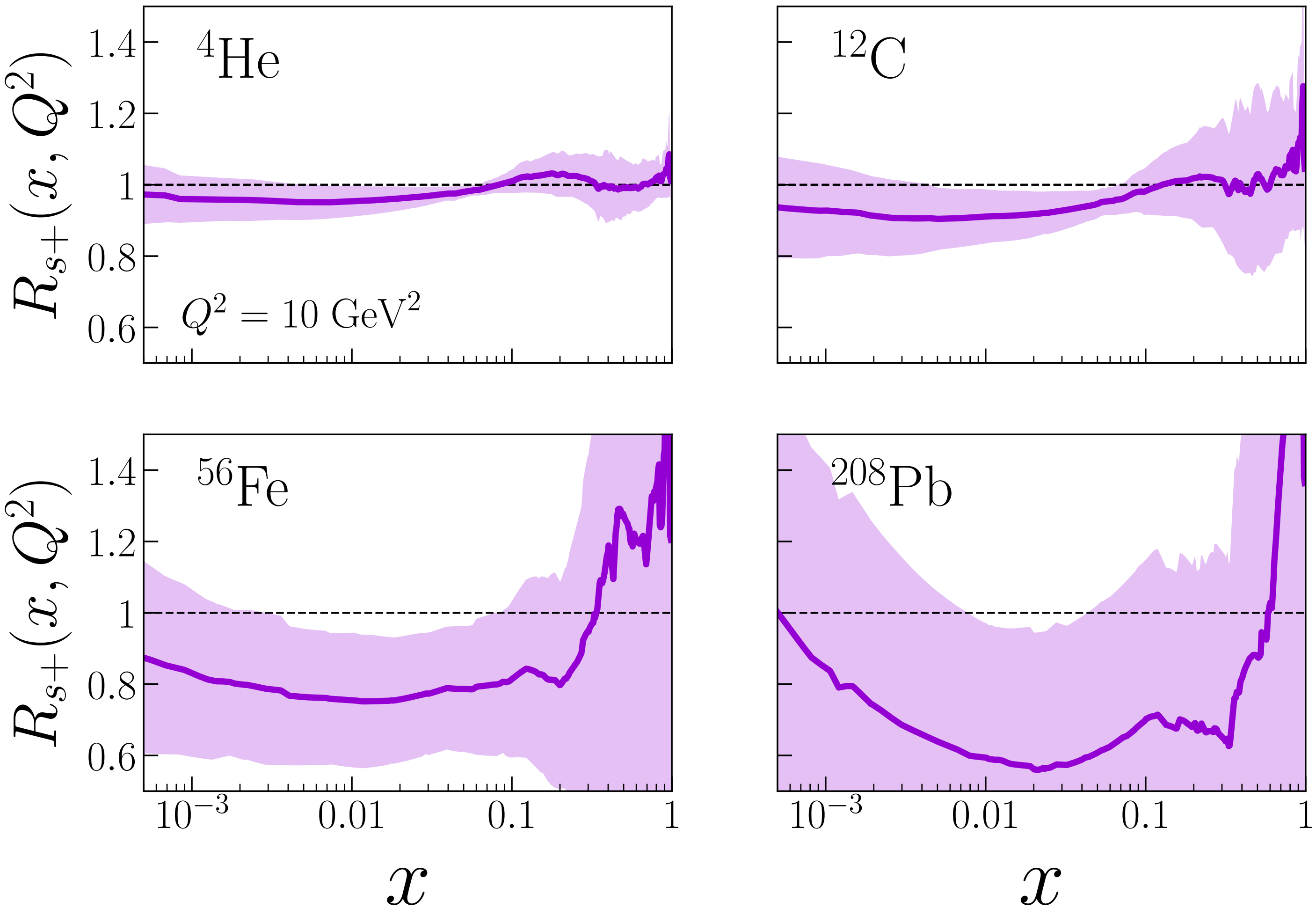
Flavor Dependent Nuclear Effects



$$R_f(x, Q^2) = \frac{f^{N/A}(x, Q^2)}{f^N(x, Q^2)}$$

- Growth of uncertainties with A

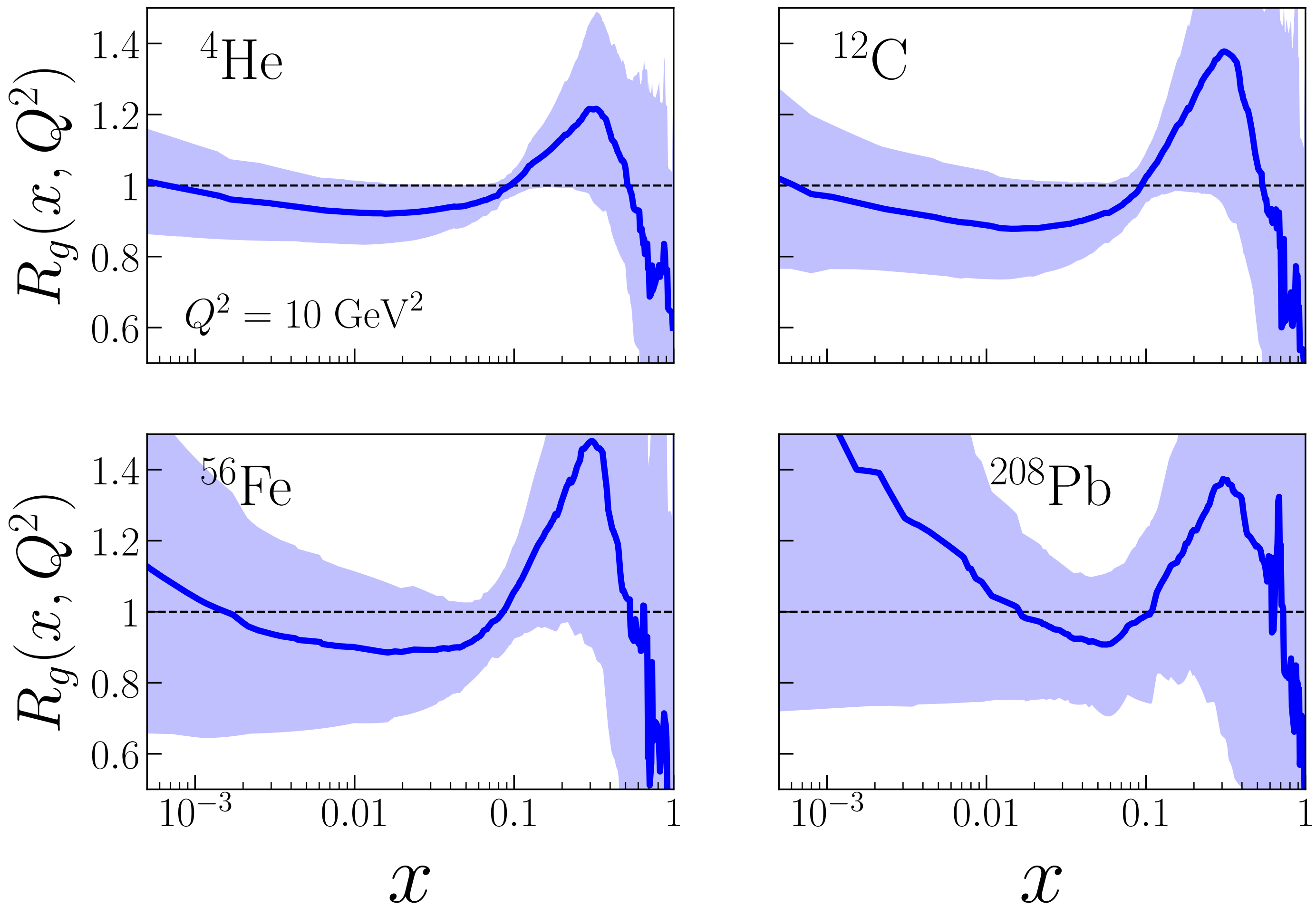
Flavor Dependent Nuclear Effects



$$R_f(x, Q^2) = \frac{f^{N/A}(x, Q^2)}{f^N(x, Q^2)}$$

- Strange suppression for large nuclei

Flavor Dependent Nuclear Effects

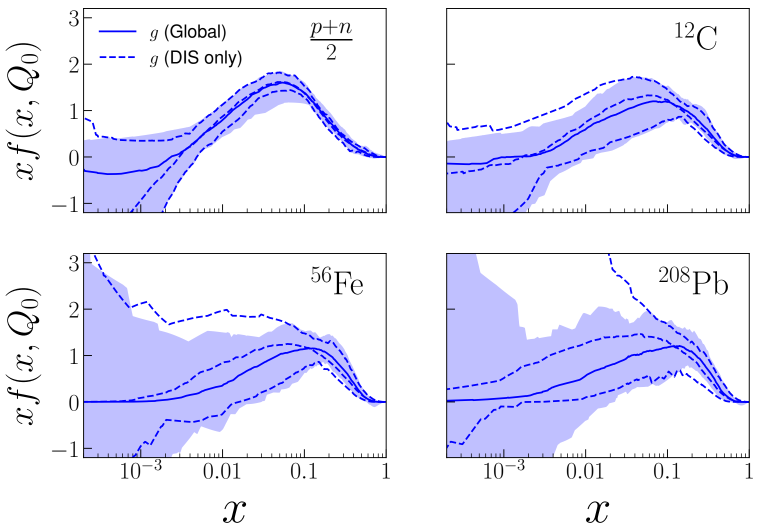


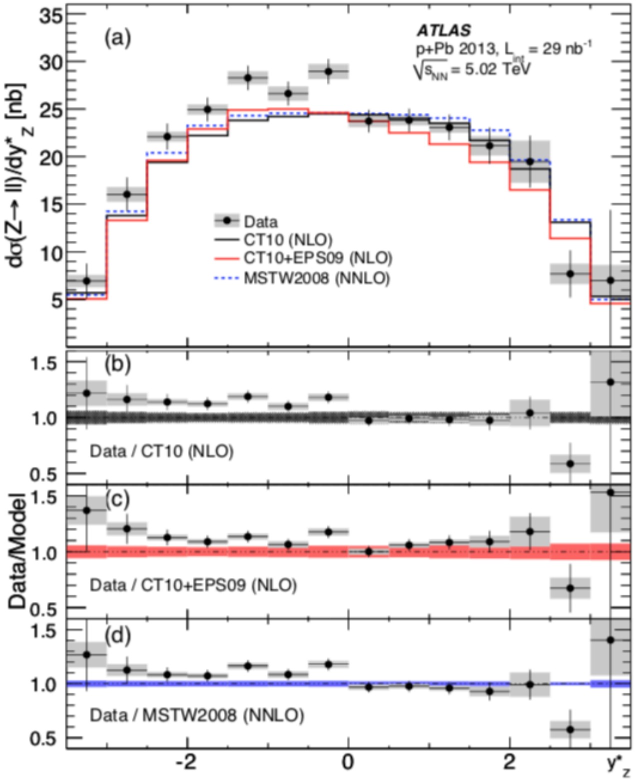
$$R_f(x, Q^2) = \frac{f^{N/A}(x, Q^2)}{f^N(x, Q^2)}$$

- Evidence of gluon shadowing unclear

Impact of ATLAS Z production

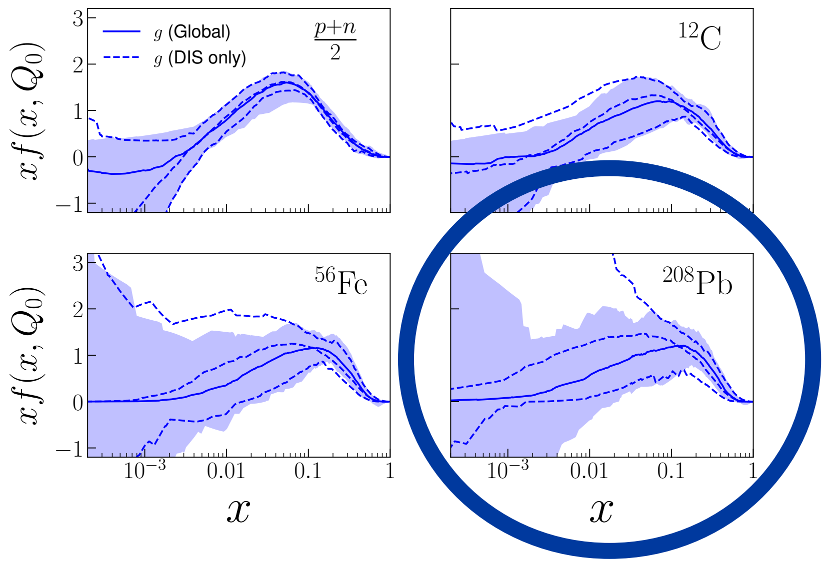
- 14 data points from pPb → Z → dilepton

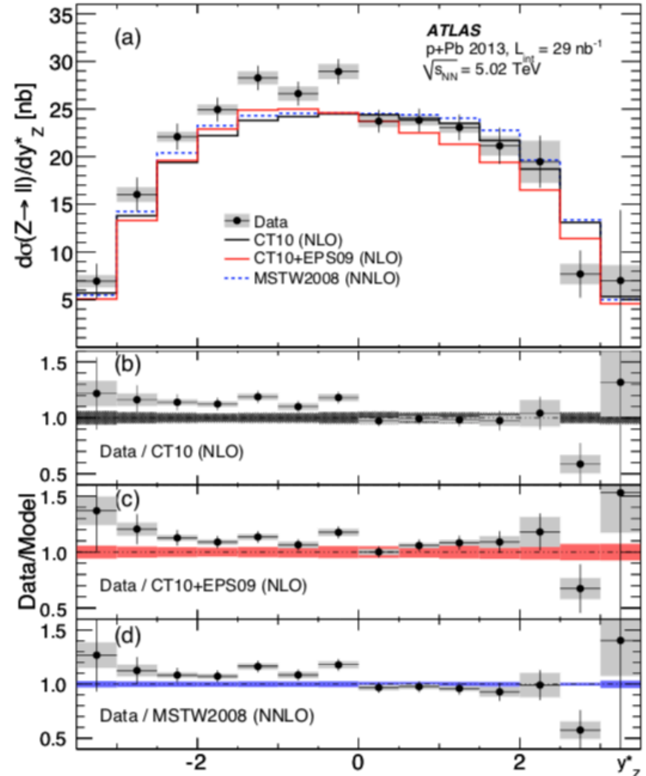




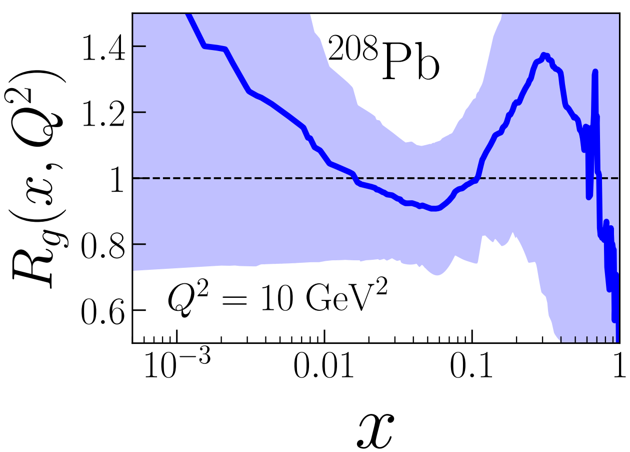
Impact of ATLAS Z production

- 14 data points from pPb → Z → dilepton

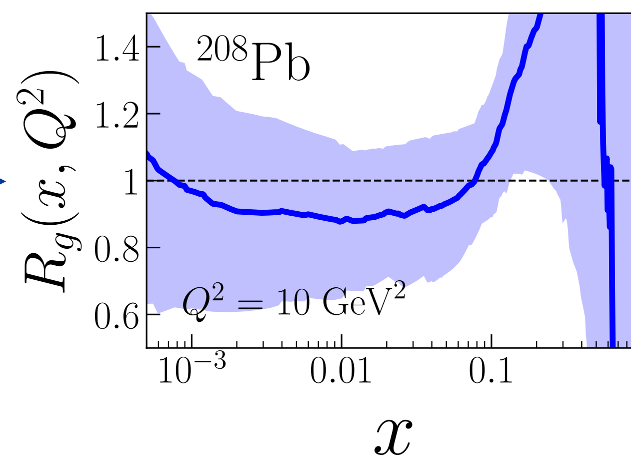




(Prior) $\chi^2/N_{\mathrm{data}} = 2.004$



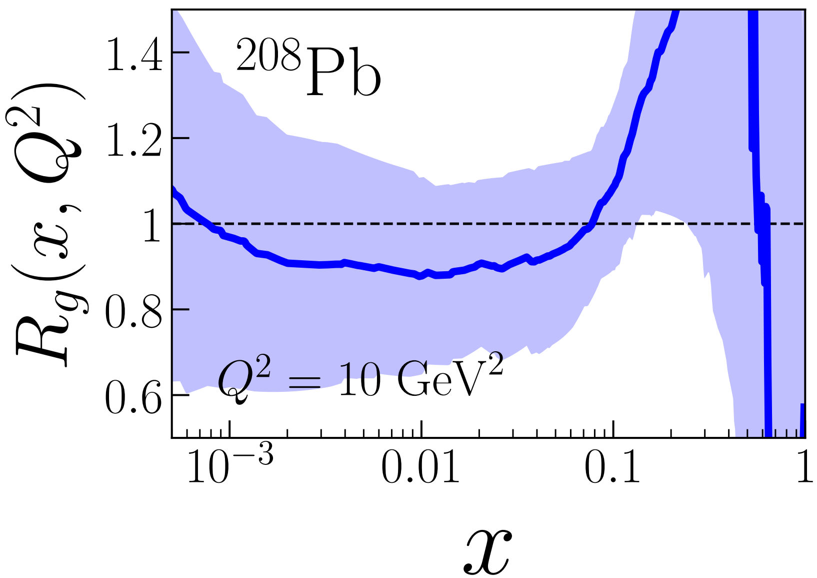
(Post Fit) $\chi^2/N_{\mathrm{data}} = 1.572$



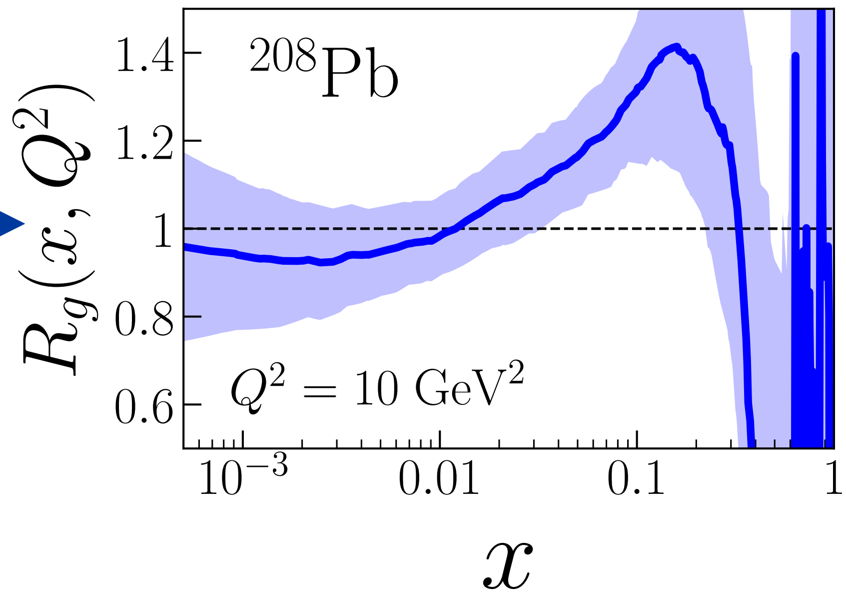
Re-analysis of EIC Impact

- Analysis of high energy EIC pseudo-data (constructed with prelim 2.0)

nNNPDF2.0 prelim



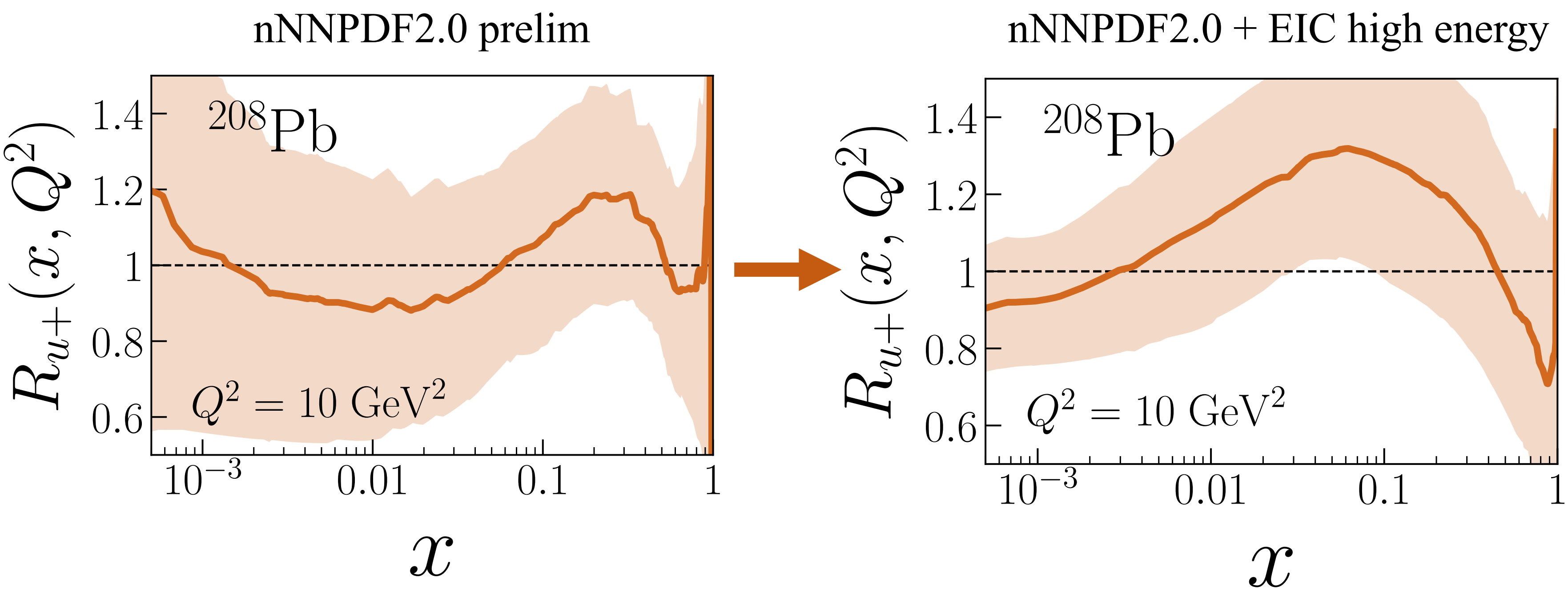
nNNPDF2.0 + EIC high energy



- ~20% reduction in 90% CL levels for Pb ratio at low x!

Re-analysis of EIC Impact

- Analysis of high energy EIC pseudo-data (constructed with prelim 2.0)



- ~20% reduction in 90% CL levels for Pb ratio at low *x*!
- Can help pin down flavor separation at low *x*

Summary and Outlook

- Machine learning + Monte Carlo methods are important for robust extractions of nPDFs and their uncertainties
- Methodology improvements in nuclear PDF analysis:

  → Neural networks optimized with stochastic gradient descent in TensorFlow

- Highlights from first Monte Carlo nPDF fit

  → Significant impact of A=1 boundary condition for low-A nuclei

  → High energy EIC scenario can constrain nPDFs down to $x \sim 10^{-4}$

- Towards nNNPDF2.0

  → Inclusion of additional observables (VB production from LHC p+Pb collisions, CC DIS) for flavor separation and uncertainty reduction

  → New proton BC (not contaminated with heavy nuclear target data)

  → Positivity constraints

Summary and Outlook

- Machine learning + Monte Carlo methods are important for robust extractions of nPDFs and their uncertainties
- Methodology improvements in nuclear PDF analysis:

  → Neural networks optimized with stochastic gradient descent in TensorFlow

- Highlights from first Monte Carlo nPDF fit

  → Significant impact of A=1 boundary condition for low-A nuclei

  → High energy EIC scenario can constrain nPDFs down to $x \sim 10^{-4}$

- Towards nNNPDF2.0

  → Inclusion of additional observables (VB production from LHC p+Pb collisions, CC DIS) for flavor separation and uncertainty reduction

  → New proton BC (not contaminated with heavy nuclear target data)

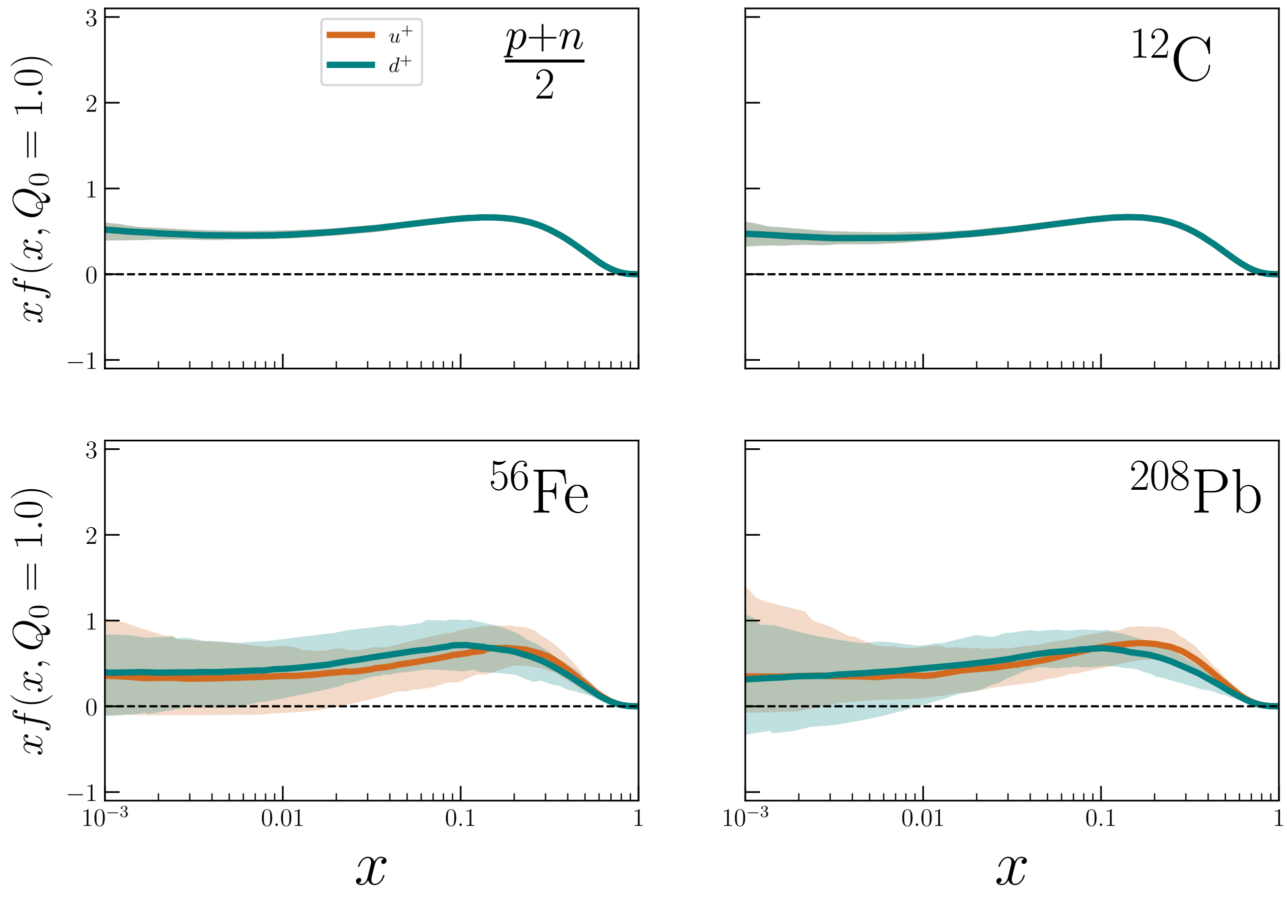
  → Positivity constraints

THANK YOU!

BACKUP SLIDES

Non-Isoscalarity

$$\tilde{T}_3^{(p/A)} \equiv \frac{2Z-A}{A} T_3^{(p/A)} = \frac{2Z-A}{A} \left(u^+ - d^+\right)^{(p/A)} \text{ (= 0 for isoscalar nuclei)}$$



Sum Rules

- Single fit with momentum and valence sum rules

```
[0.0001, 1]:      V SR_C12 =      2.9945339662073343      Momentum SR_C12 =      0.9999974805293064      V3 SR_C12 =
2.838330794514598e-19
[0.001, 1]:       V SR_C12 =      2.9551405609330423      Momentum SR_C12 =      0.9993843277929704      V3 SR_C12 =
1.8566378850048086e-18
[0.01, 1]:        V SR_C12 =      2.737367442451421       Momentum SR_C12 =      0.9868255647804303      V3 SR_C12 =
-3.030281551583433e-19
```

- Single fit without momentum and valence sum rules

```
[0.0001, 1]:      V SR_C12 =      2.937137160084468       Momentum SR_C12 =      1.0160676683661045      V3 SR_C12 =
1.634969168485992e-18
[0.001, 1]:       V SR_C12 =      2.8889274897673056      Momentum SR_C12 =      1.0147116303766643      V3 SR_C12 =
5.84957487568062e-19
[0.01, 1]:        V SR_C12 =      2.6173323416484955      Momentum SR_C12 =      1.0002115750268383      V3 SR_C12 =
9.72761683126722e-19
```

- Violation < 3%