A Monte Carlo Analysis of Nuclear PDFs with Neural Networks

[arXiv:1904.00018]

Jacob Ethier

In collaboration with R. Abdul Khalek and J. Rojo

8^{th} Workshop of the APS Topical Group on Hadronic Physics

April 10^{th}, 2019

VU VRIJE UNIVERSITEIT AMSTERDAM

NNPDF

Nikhef

NWO Netherlands Organisation for Scientific Research

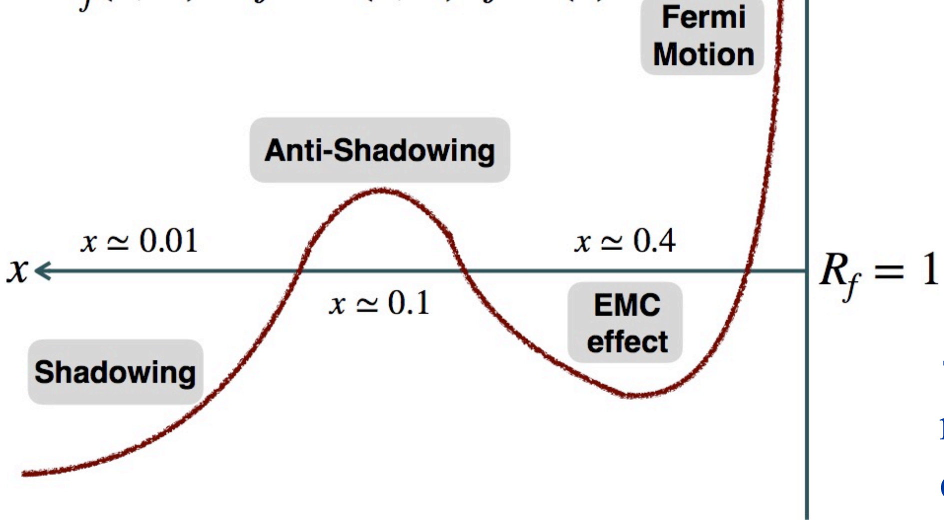
Motivation

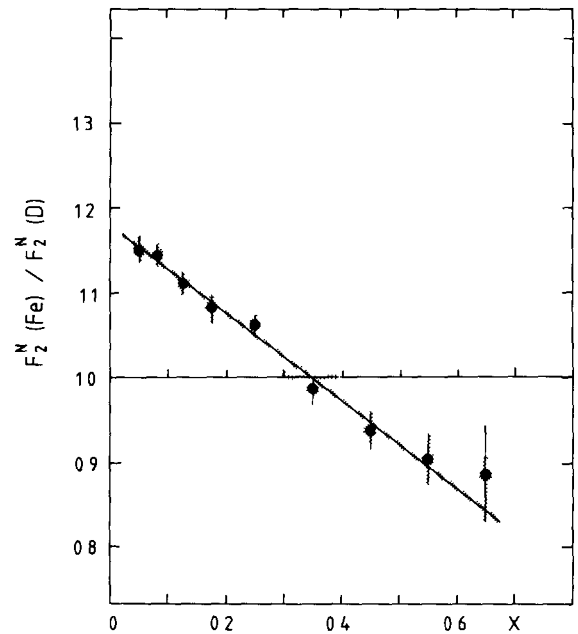
- Significant deviations from unity in measurements of nuclear to nucleon structure function ratios (EMC effect)

→ Structure functions of the nucleus cannot be described as a sum of the free nucleon structure functions (up to Fermi motion corrections)

- Origin of EMC effect still not well understood

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





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

→ Determining the modification of free nucleon PDFs in nuclei can provide crucial insight to various nuclear effects

Motivation

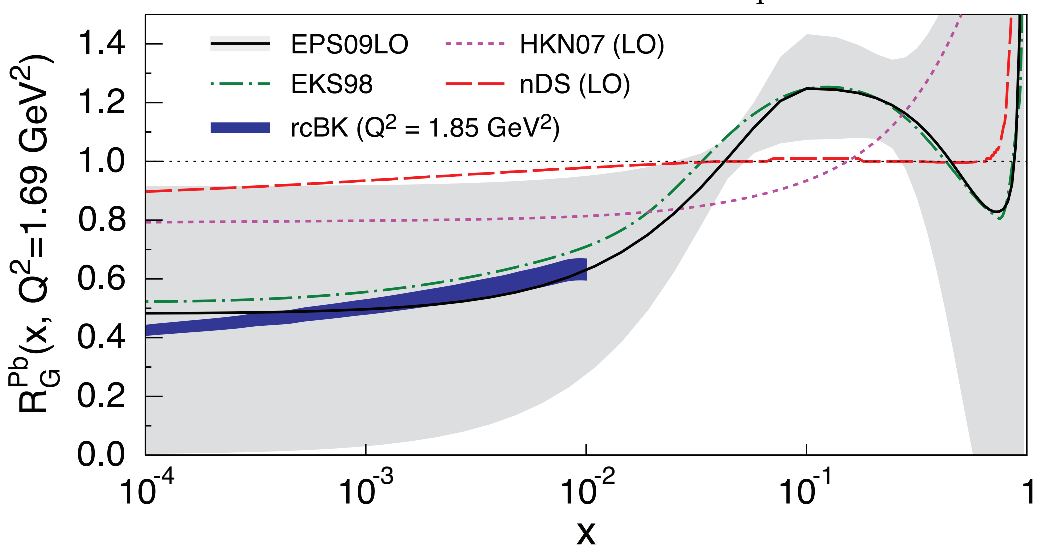
- nPDFs can reveal onset of non-linear evolution effects at low x and Q^2

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

- Can be determined from ratio of nuclear to proton PDFs:

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

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



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

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

Global QCD Analyses

- Method to extract information of the nonperturbative dynamics associated with nuclei structure within QCD

Factorization → separation of short and long distance physics in pQCD expressions of experimental observables, e.g. for electron-proton scattering

Unpolarized deep inelastic scattering (DIS) observable

$\ell + (p, d) \rightarrow \ell' + X$

$$d\sigma(x, Q^2) \simeq \sum_f \int_x^1 \frac{d\xi}{\xi} \underbrace{f\left(\frac{x}{\xi}, Q^2\right)}_{\text{Parton distribution function (PDF)}} \underbrace{d\hat{\sigma}_f(\xi, Q^2)}_{\text{Hard scattering coefficient}}$$

- Collinear factorization → distributions depend on the fraction of longitudinal proton momentum

- The same formalism can be used to study modification of PDFs in nuclei

$$f^{(p)}(x, Q^2) \rightarrow f^{(p/A)}(x, Q^2, A)$$

- Nuclear PDFs are extracted from global data of lepton-nucleus and hadron-nucleus collisions using QCD factorization

Global QCD Analyses

- Standard determination of nonperturbative functions form global QCD analyses:

  → Objects are parameterized $xf(x) = Nx^a(1-x)^b(1+c\sqrt{x}+dx)$

  → Parameters are optimized with a least-squares fit

$$\chi^2 = \sum_e^{N_{exp}} \sum_i^{N_{data}} \frac{(D_i^e - T_i)^2}{(\sigma_i^e)^2}$$

- Issues with performing single chi-squared minimization

  → Uncertainties computed by a Hessian method introduce tolerance criteria (uncertainties inflated by arbitrary factor)

  → Parameters difficult to constrain are typically fixed

  → Highly non-linear chi-squared function = many local minima that a single fit can be trapped

- Nuclear PDFs require proton PDF as boundary condition → typically taken from previous global analyses (which often make different theoretical/methodological choices!)

  → Need a consistent theoretical framework and robust fitting procedure to reliably determine nuclear PDF central values and uncertainties

nNNPDF1.0 Analysis

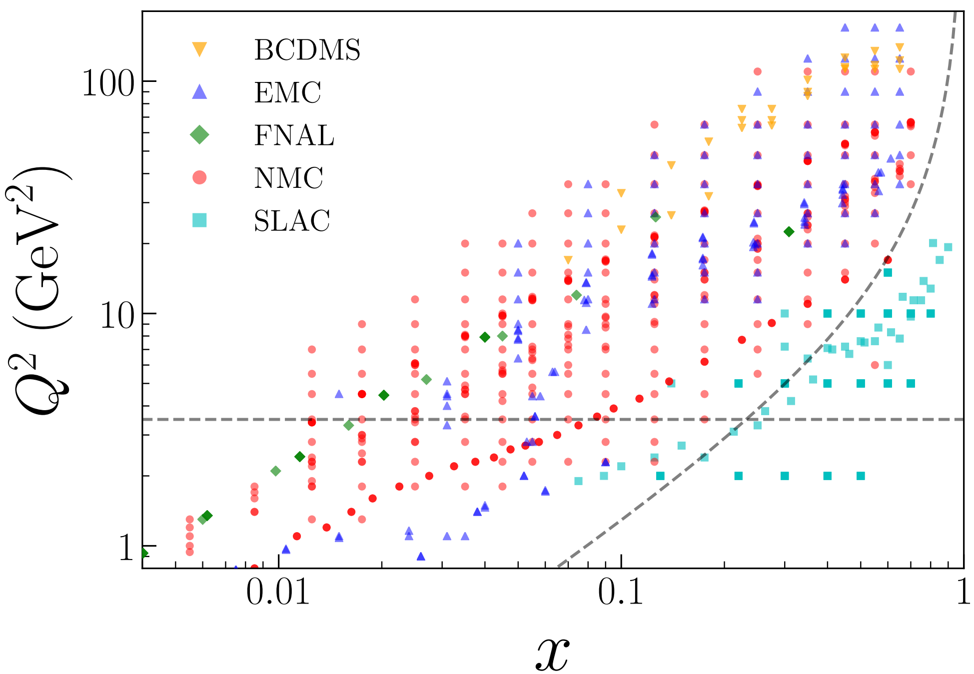
- Includes all available neutral current DIS data from CERN, SLAC, and FNAL experiments

  → 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**

nNNPDF1.0 Fitting Methodology

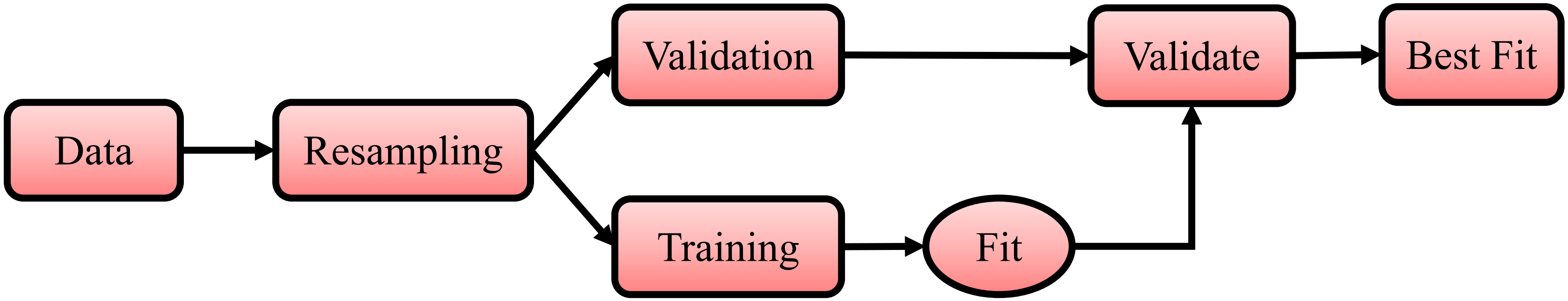
- Based on NNPDF framework:

  → PDFs modeled with neural networks (removes parameterization bias)

  → Monte Carlo sampling of parameter space

  → Gaussian smearing of experimental data

  → Cross-validation with early stopping (prevents overfitting)



→ Many fits are performed to obtain sample distribution that represents the parent probability distribution of nPDFs

nPDF Optimization

- Minimize the cost function:

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

nPDF Optimization

- Minimize the cost function:

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

Experimental measurements

nPDF Optimization

- Minimize the cost function:

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

Experimental measurements

Theoretical predictions (functions of the parameterized PDFs)

nPDF Optimization

- Minimize the cost function:

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

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

Experimental measurements

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

- Minimize the cost function:

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

Experimental measurements

Theoretical predictions (functions of the parameterized PDFs)

Covariance matrix – encodes all uncorrelated and correlated experimental uncertainties

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

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

nPDF Optimization

- Minimize the cost function:

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

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

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

nPDF Optimization

- Minimize the cost function:

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

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

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

→ Free nucleon PDFs must be reproduced at A=1

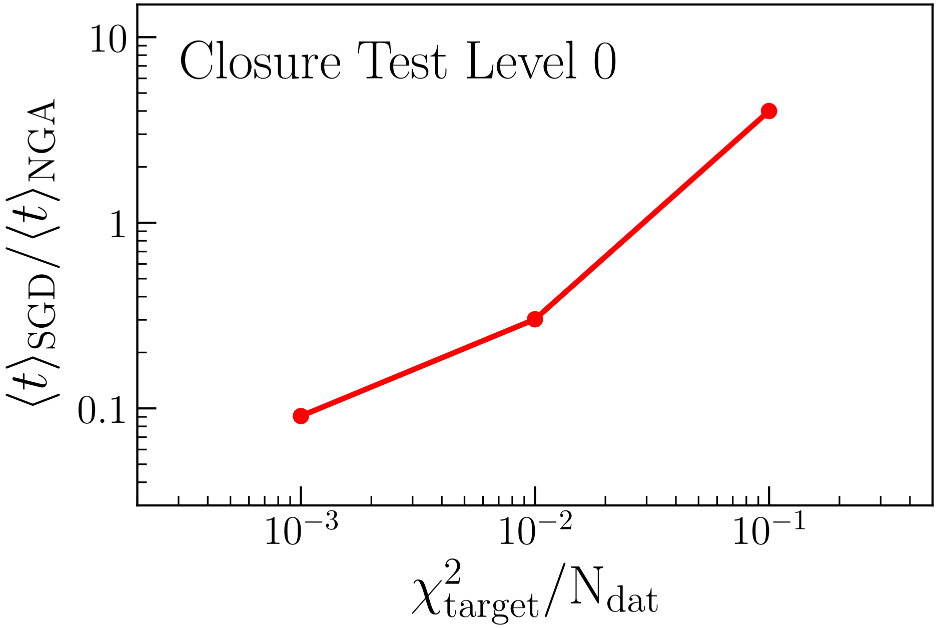
$$f(x, Q, A=1) = \frac{1}{2}\left[f_p(x, Q^2) + f_n(x, Q^2)\right]$$

→ 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!

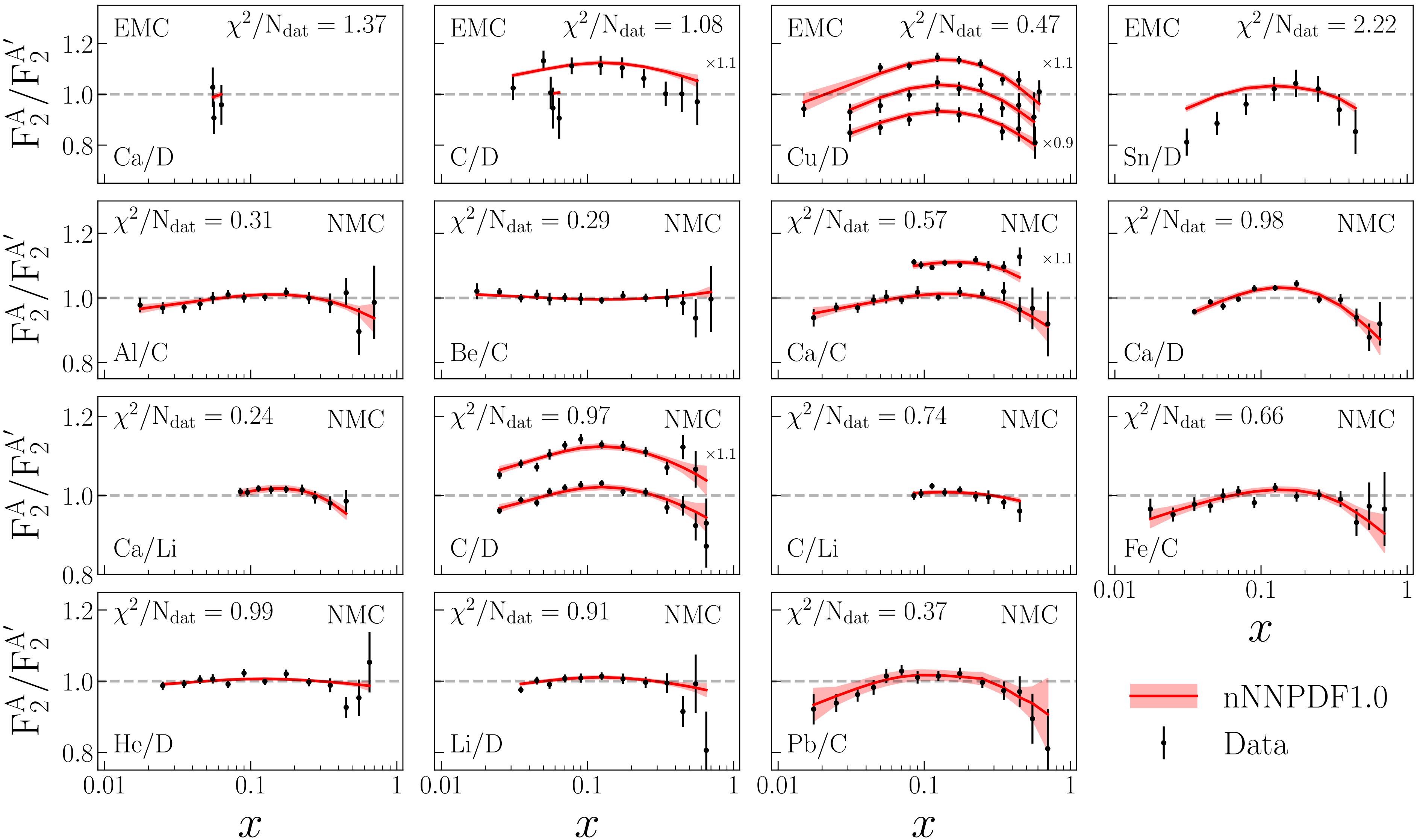
Optimization with TensorFlow

- Open source machine learning software library developed by Google's AI organization
- Highly optimized numerical computations + machine learning tools
- Gradients of cost function (chi-squared) computed by reverse-mode automatic differentiation

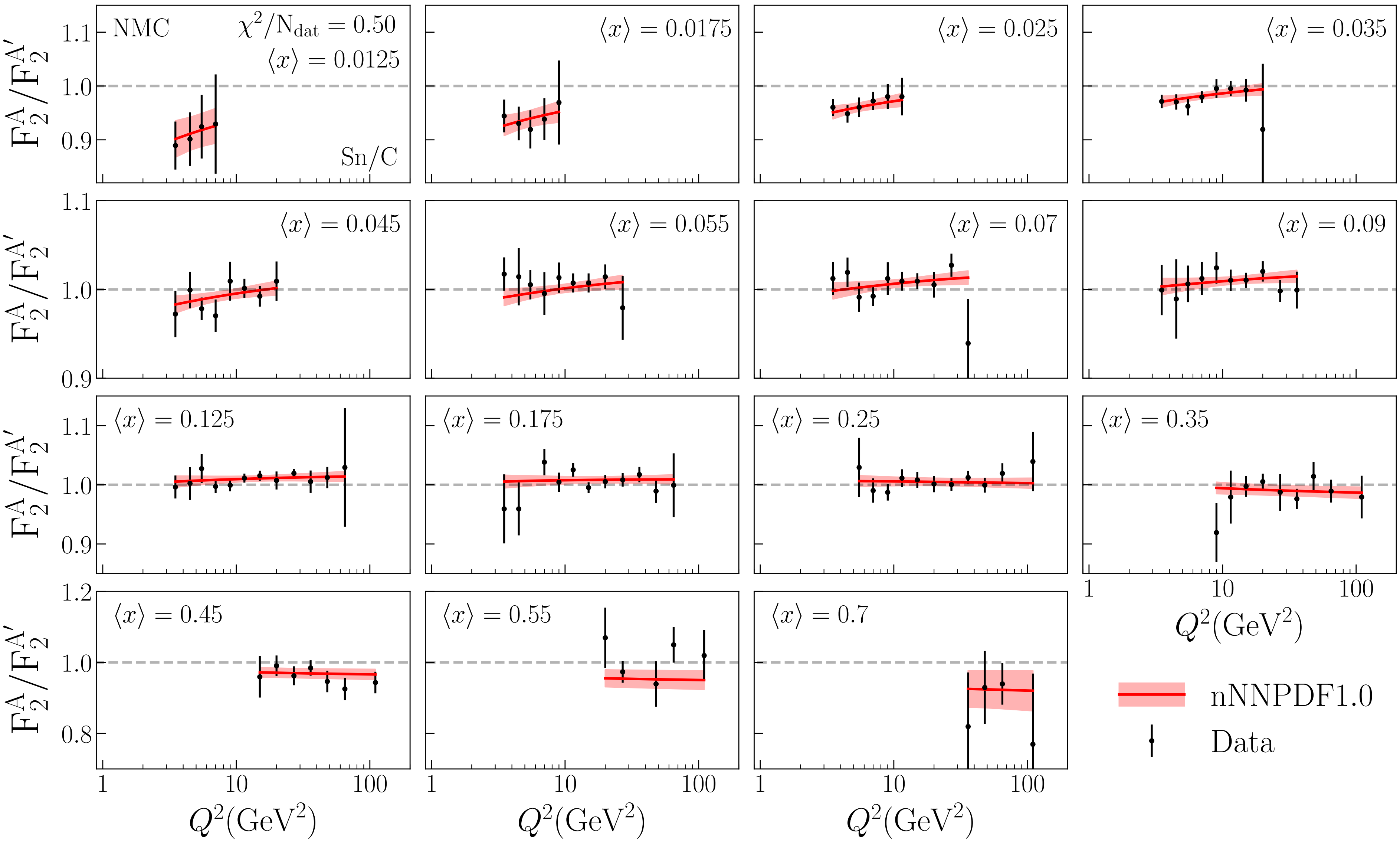


- Parameters are optimized by the Adapative Moment Estimation (ADAM) algorithm – improved stochastic gradient descent (SGD)
- Improved performance over numeric genetic algorithm (NGA) optimization used in previous NNPDF analyses

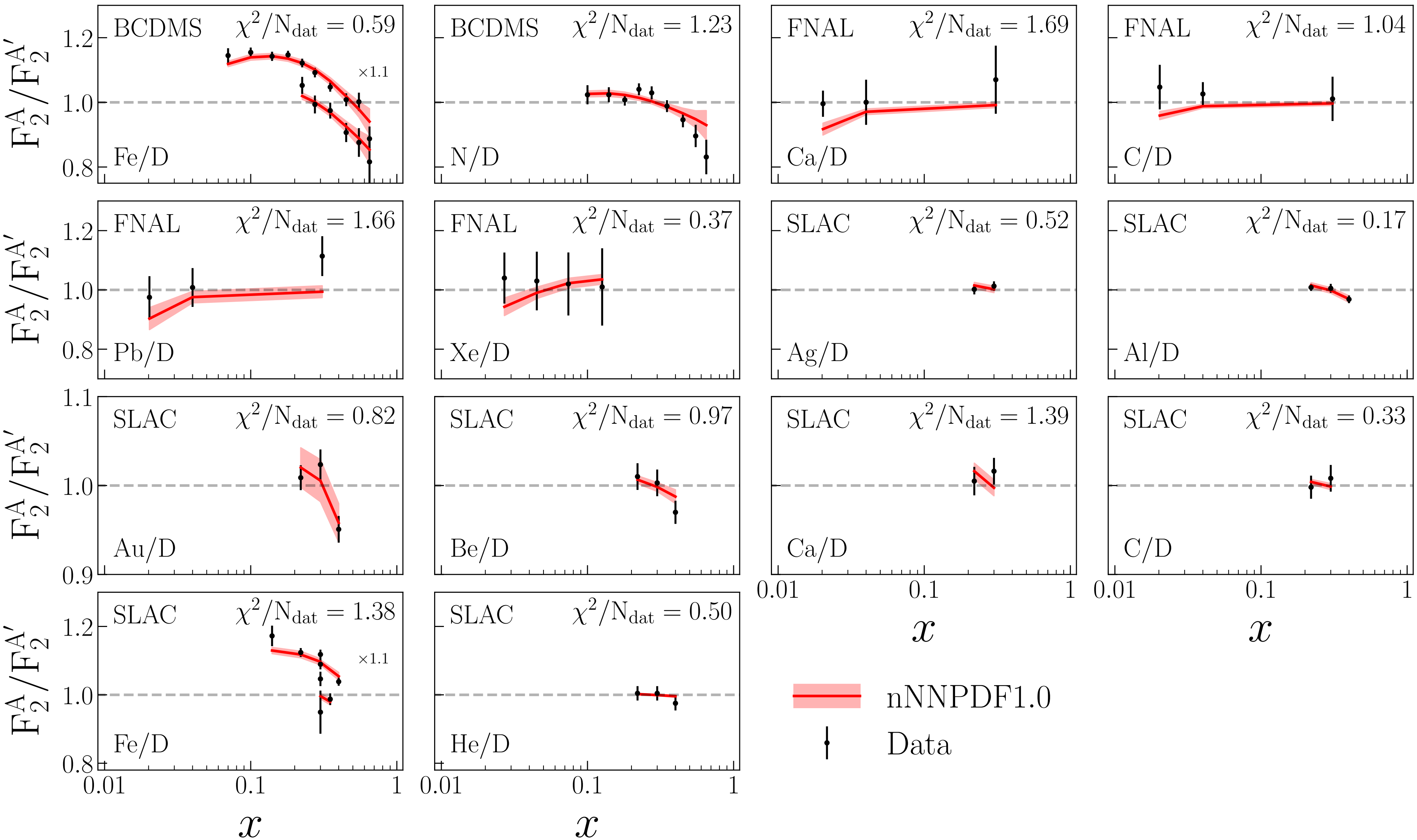
Data vs Theory – EMC + NMC



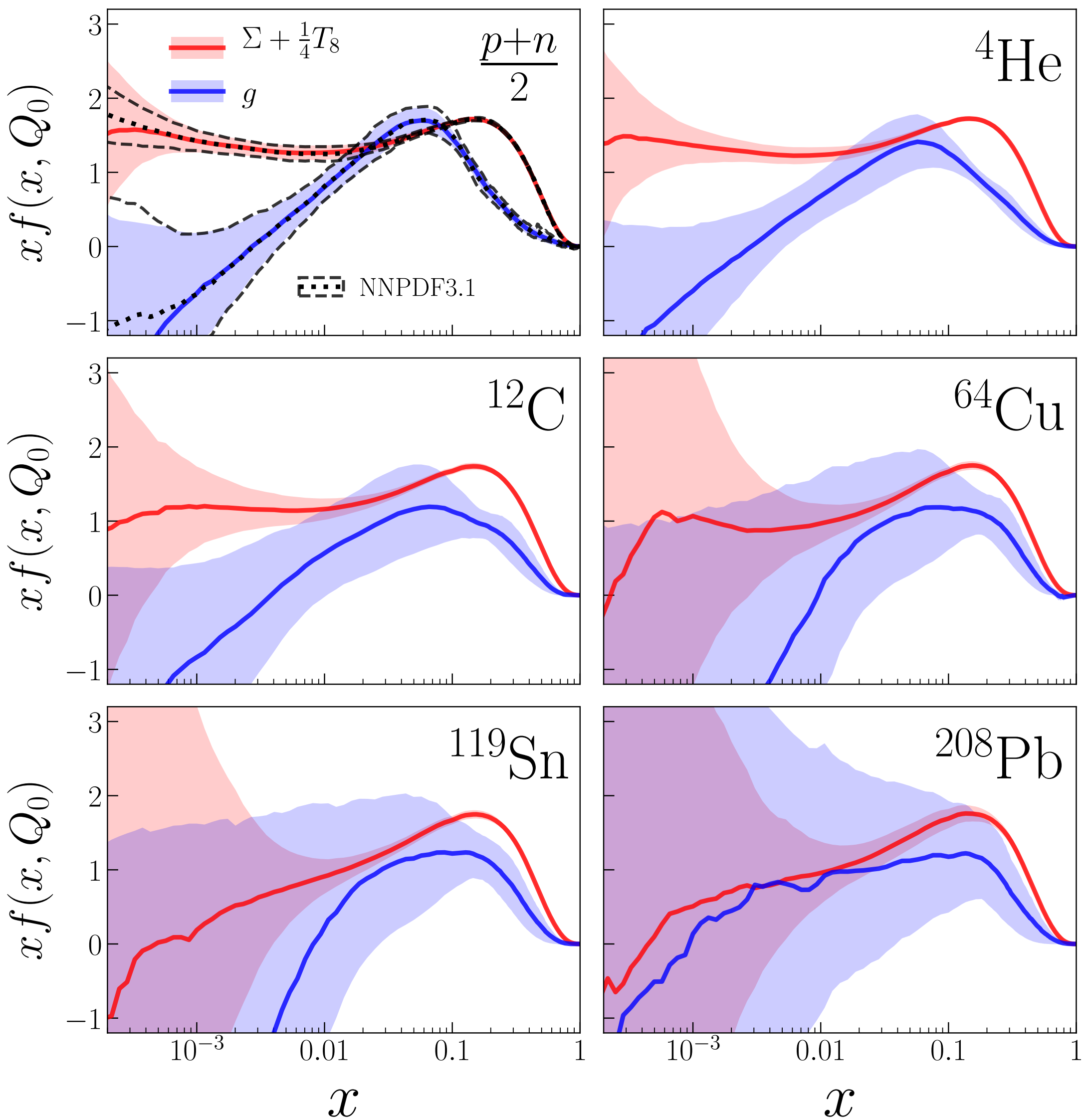
Data vs Theory – NMC Sn/C



Data vs Theory – BCDMS + SLAC + FNAL



Nuclear PDFs



Uncertainties computed as 90% CL range – central value taken to be midpoint

For every value of x:

$$f_1 \leq f_2 \leq \ldots \leq f_{N_{\rm rep}-1} \leq f_{N_{\rm rep}}$$

90% CL range:

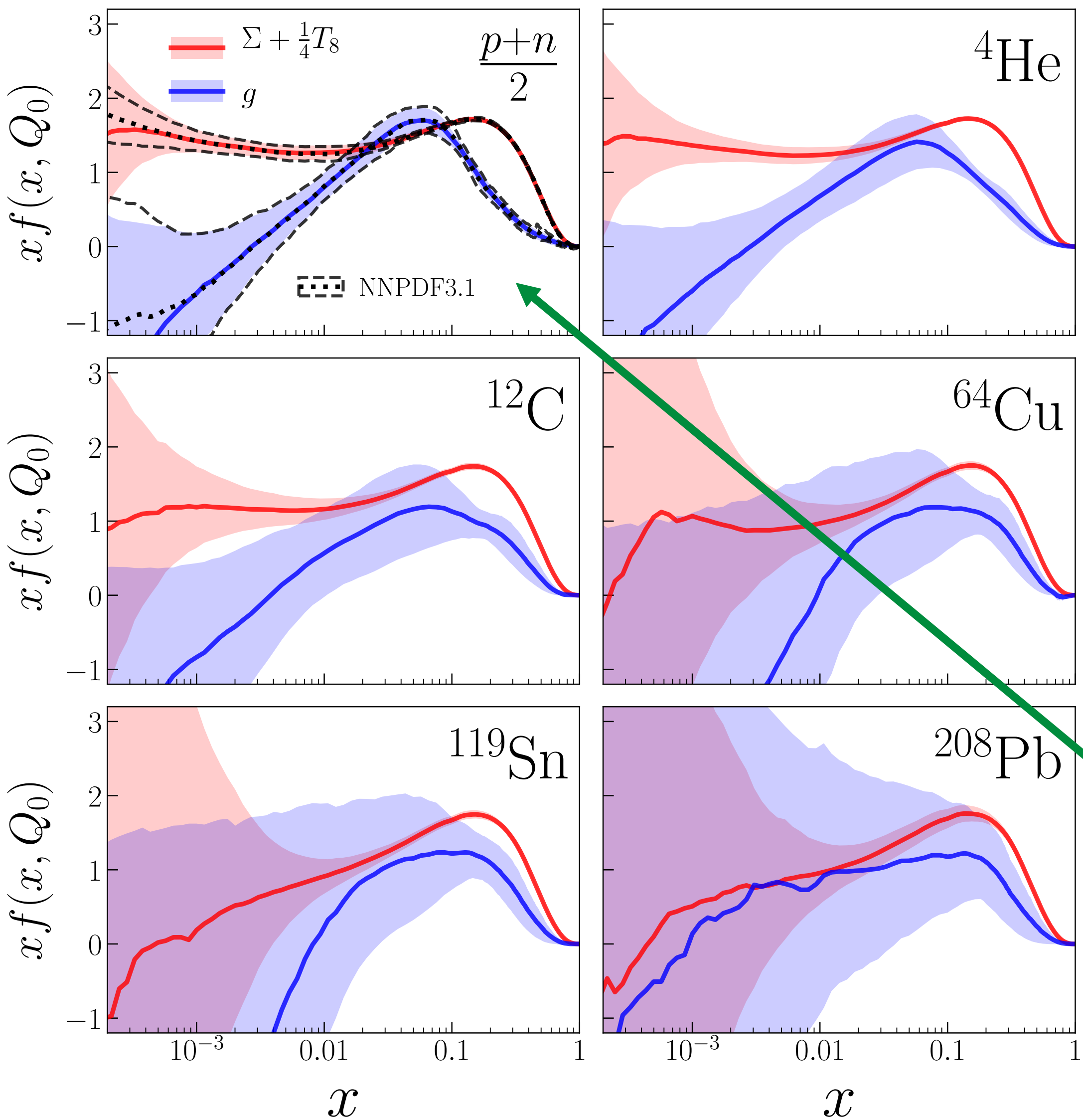
$$\left[f_{0.05\,N_{\rm rep}}, f_{0.95\,N_{\rm rep}}\right]$$

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

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

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

Nuclear PDFs



Uncertainties computed as 90% CL range – central value taken to be midpoint

For every value of x:

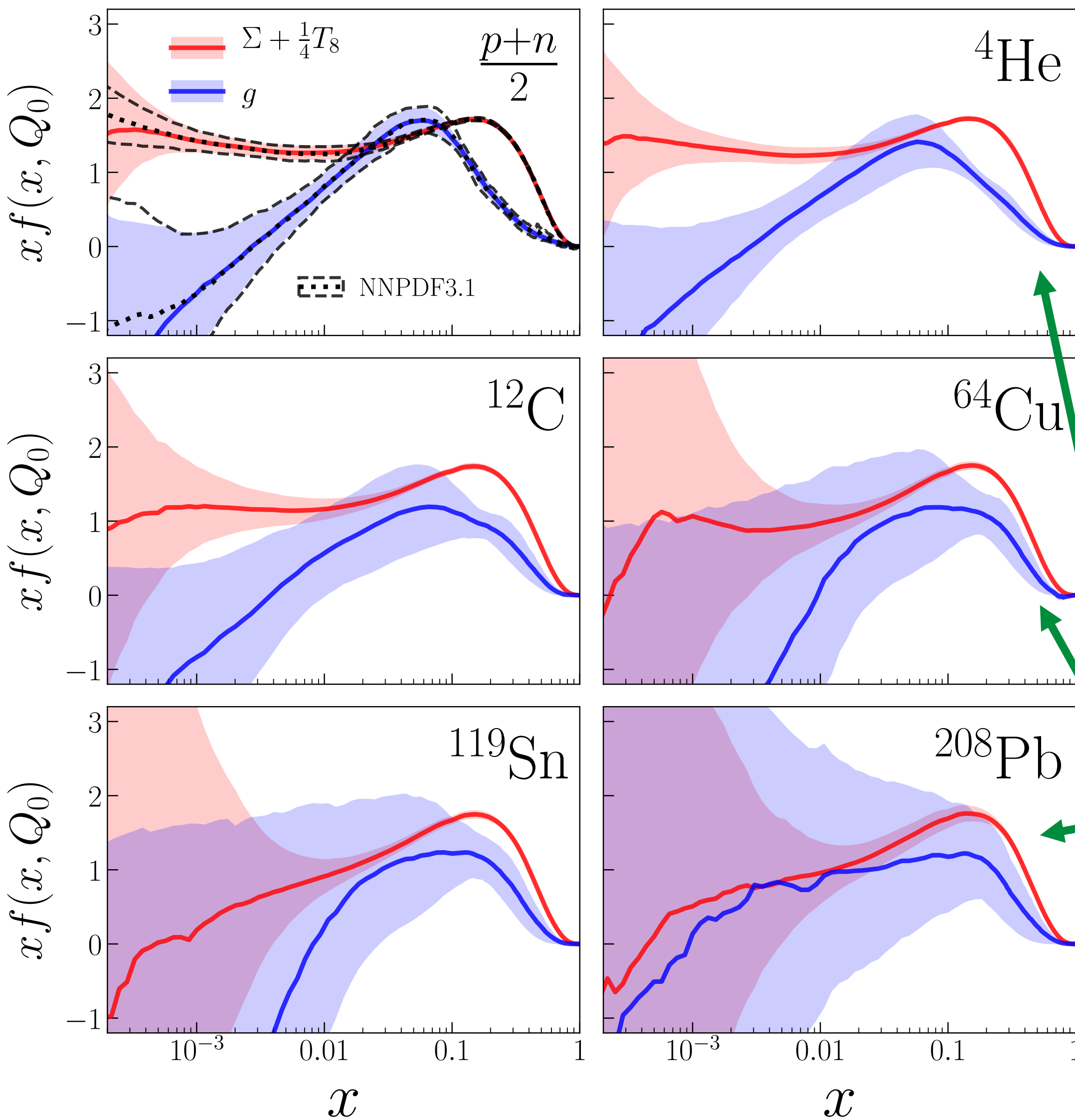
$$f_1 \leq f_2 \leq \ldots \leq f_{N_{\rm rep}-1} \leq f_{N_{\rm rep}}$$

90% CL range:

$$\left[f_{0.05\, N_{\rm rep}}, f_{0.95\, N_{\rm rep}} \right]$$

NNPDF3.1 central value and uncertainties reproduced!

Nuclear PDFs



Uncertainties computed as 90% CL range – central value taken to be midpoint

For every value of x:

$$f_1 \leq f_2 \leq \ldots \leq f_{N_{\rm rep}-1} \leq f_{N_{\rm rep}}$$

90% CL range:

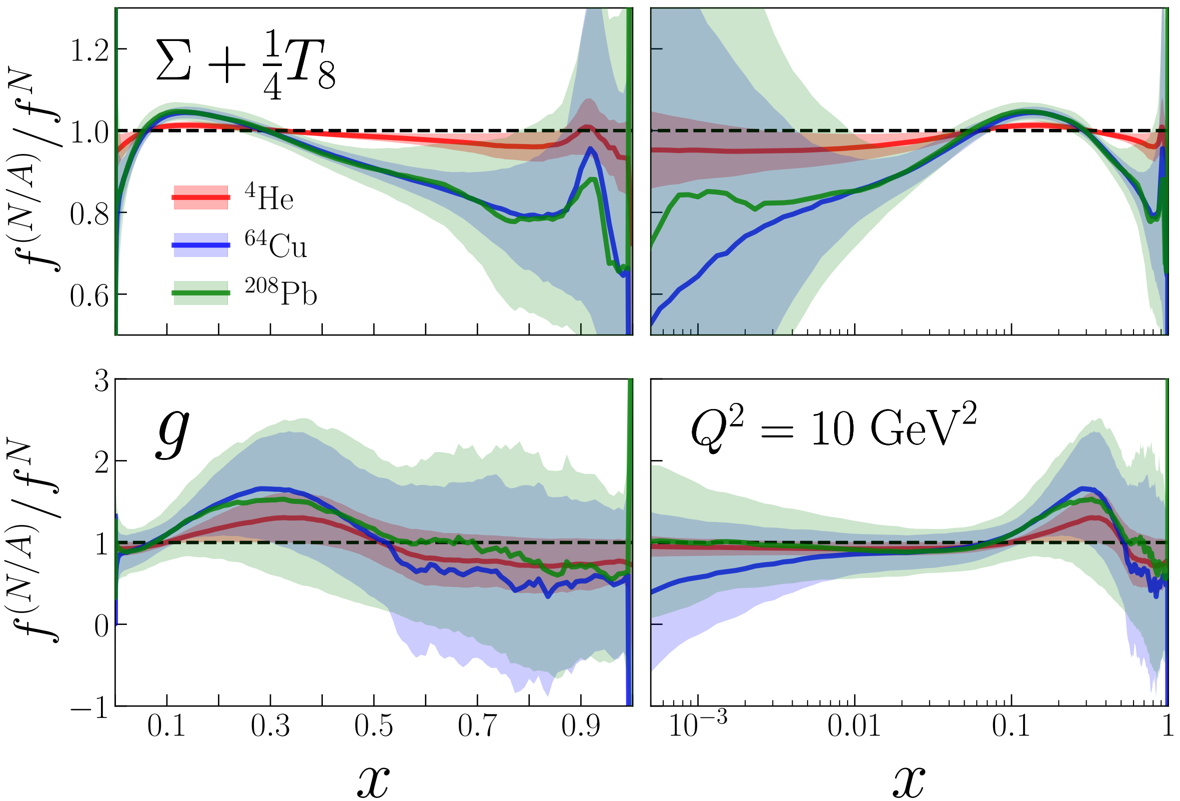
$$\left[f_{0.05\,N_{\rm rep}}, f_{0.95\,N_{\rm rep}}\right]$$

Increasing uncertainties with A – effect of boundary condition

Nuclear PDFs

Ratio to A=1 result – correlations between nPDFs included

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

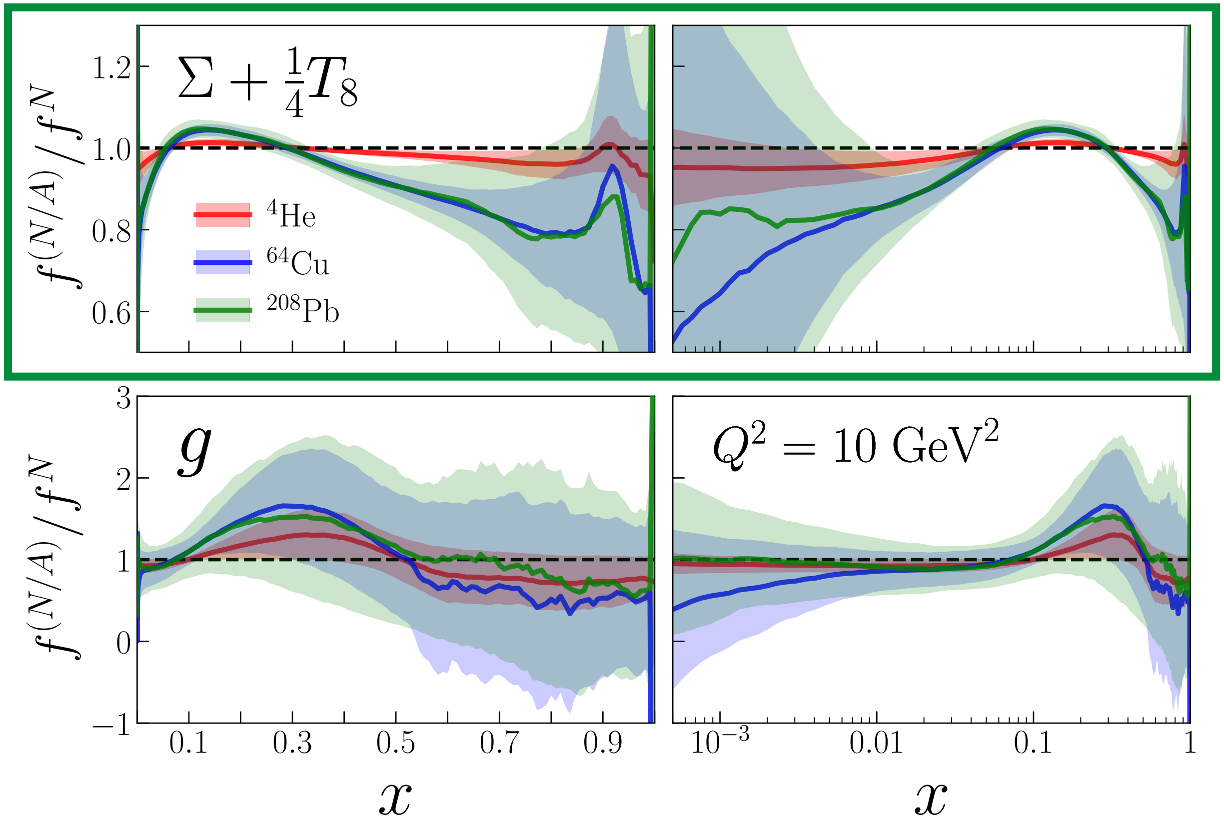


Nuclear PDFs

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



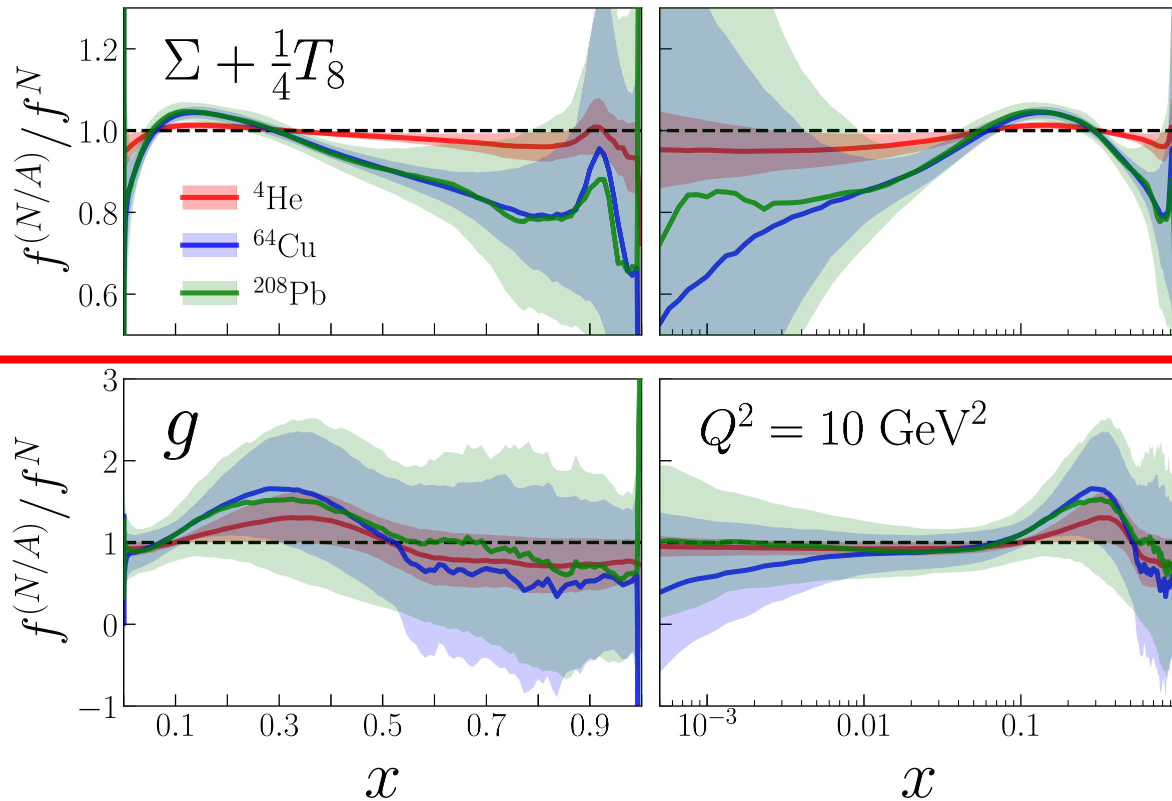
Nuclear PDFs

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

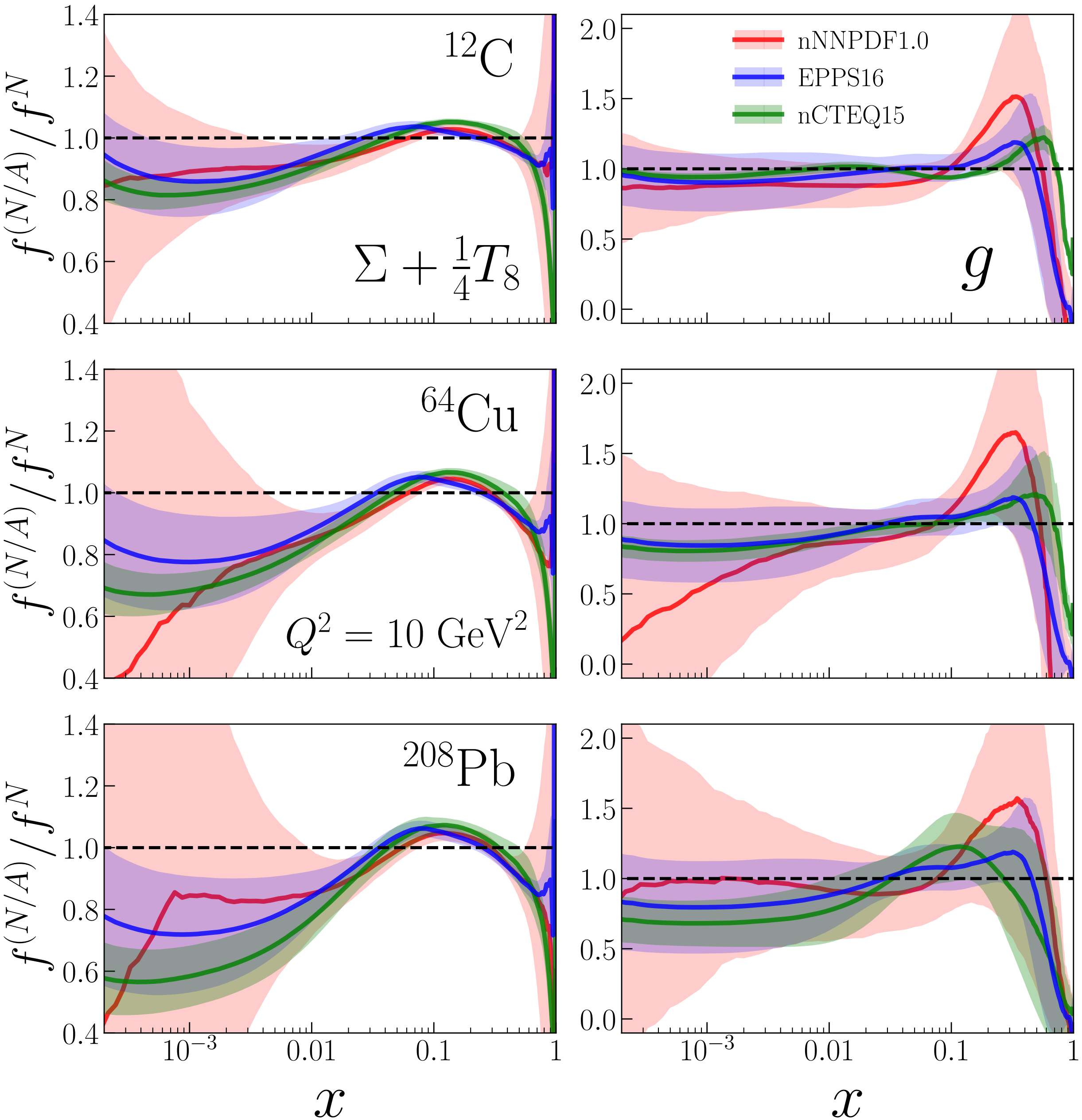


Nuclear PDFs

All distributions normalized by nNNPDF1.0 A=1 distribution

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

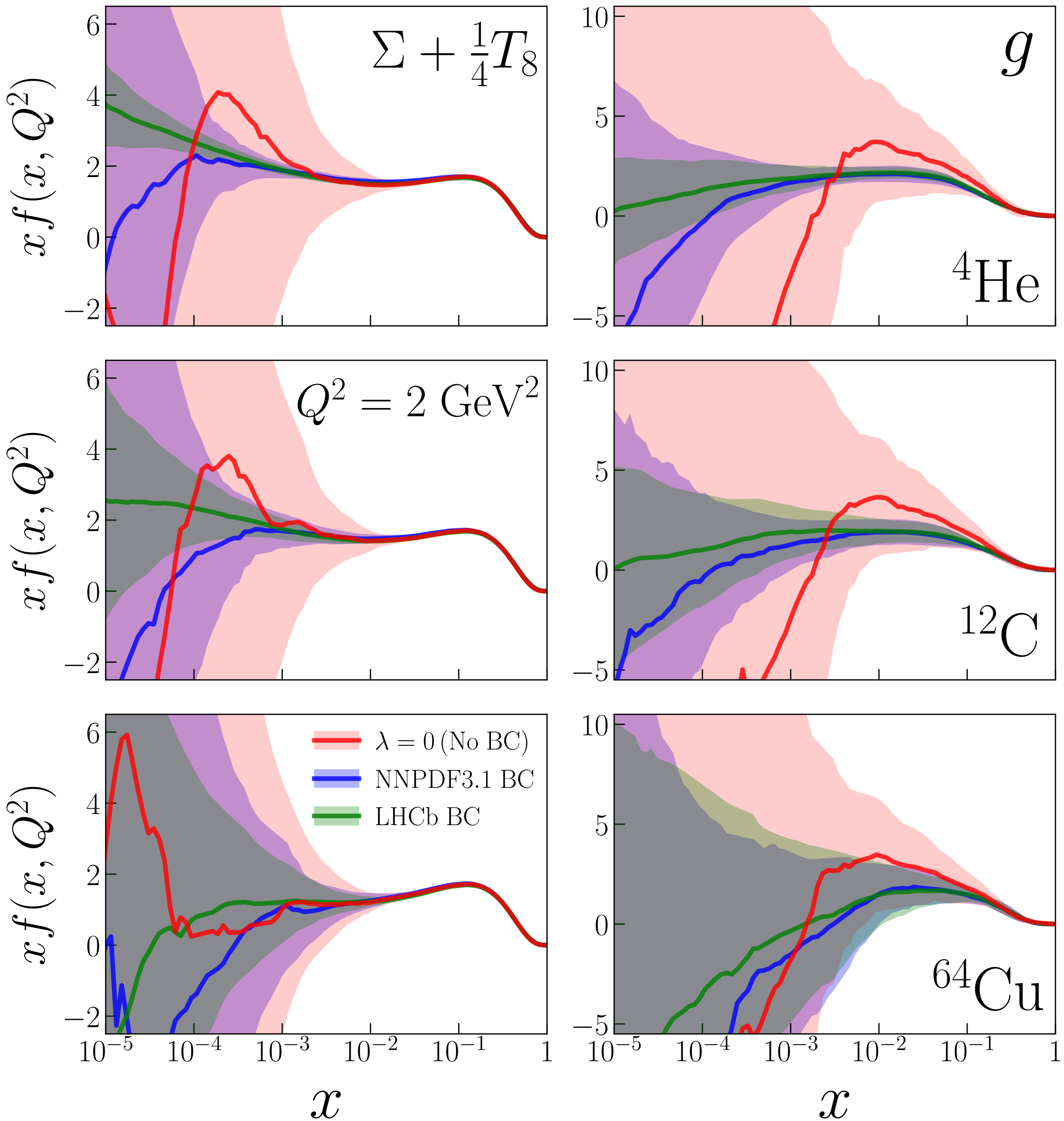
Significant differences in uncertainties!



Nuclear PDFs

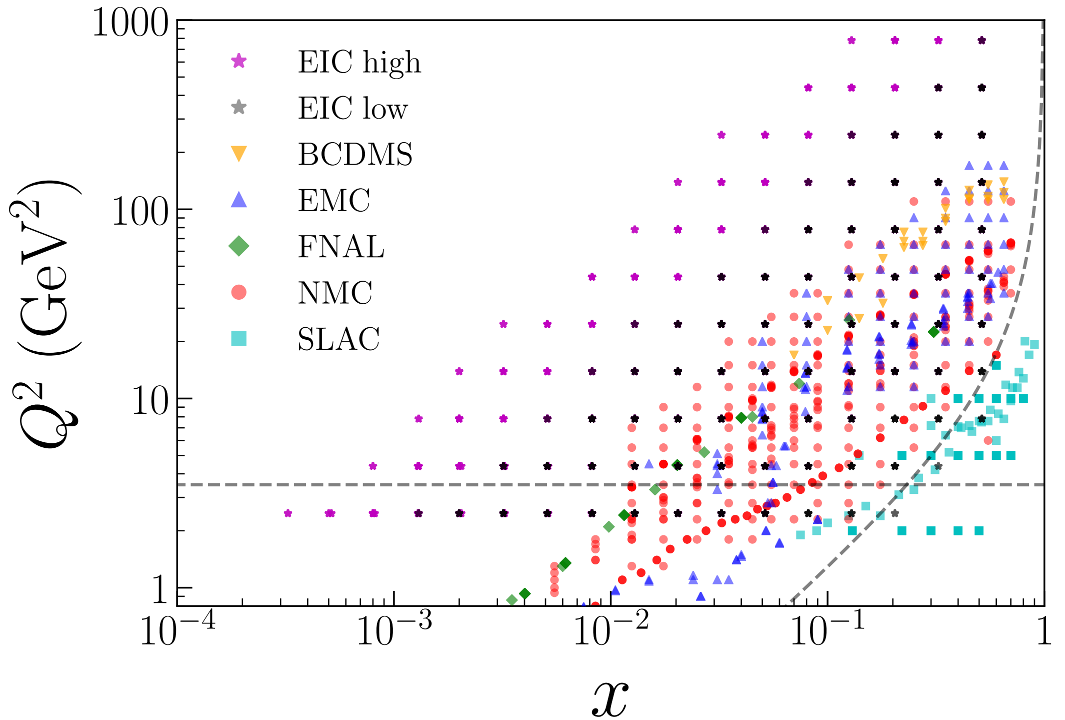
Can test other boundary conditions – NNPDF3.0+LHCb PDF set with smaller uncertainties at low x

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



Impact of the EIC

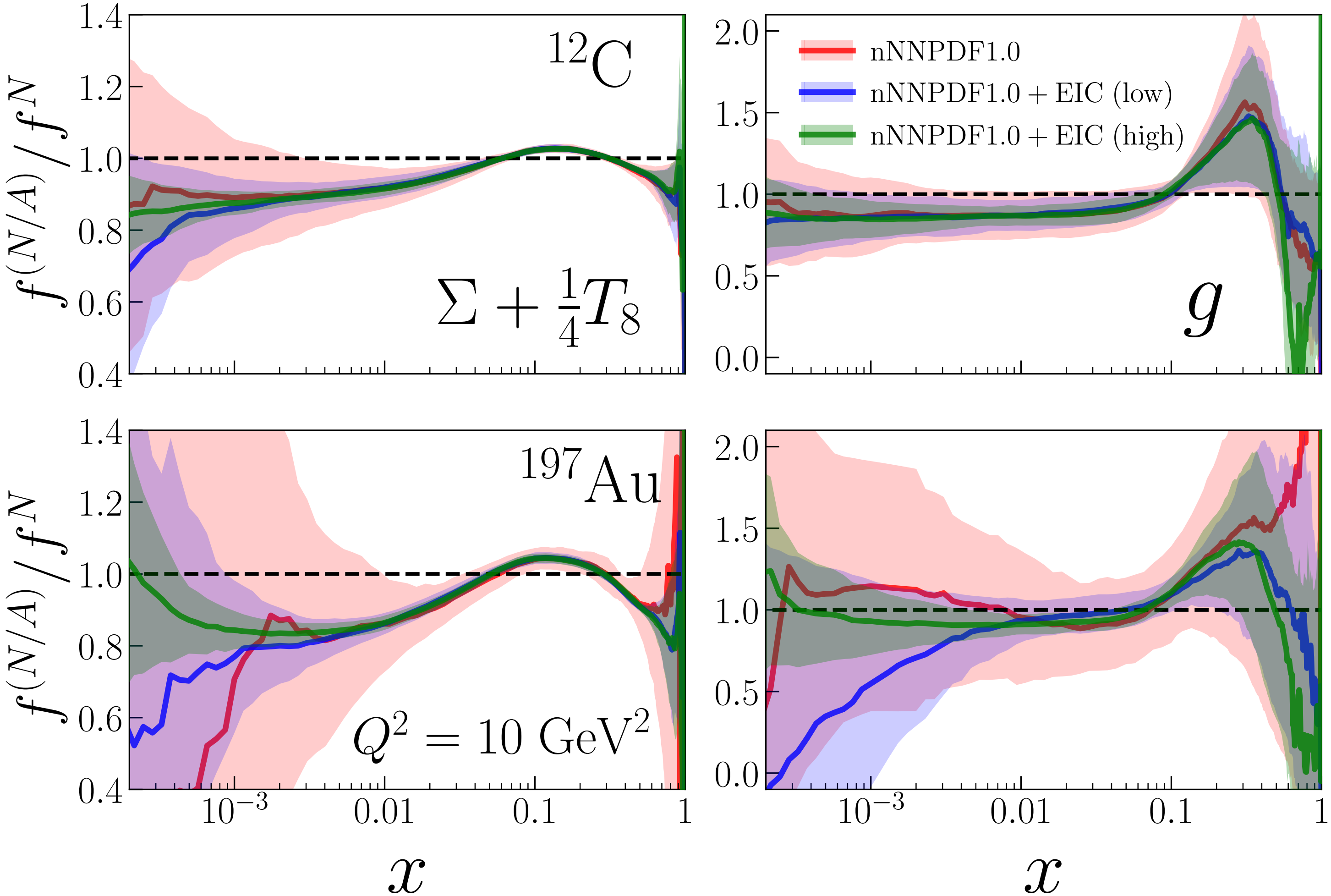
- Analysis of EIC pseudodata – extended kinematic coverage
- Two scenarios: low energy (5 GeV) vs high energy (20 GeV) electron beam



Scenario	A	E_e	E_A/A	$Q^2_{\max}$	$x_{\min}$	N_{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!

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 $x \sim 10^{-4}$
- Available NC DIS data not sensitive to separation of singlet and octet distributions

  → Inclusion of additional observables (new LHC p+Pb observables, CC DIS, Drell-Yan) for flavor separation and uncertainty reduction is needed

BACKUP SLIDES

Neutral Current (NC) DIS – Theory

- Experimental observables given as ratios of cross sections/structure functions

$$\frac{d^2\sigma^{\rm NC}(x,Q^2,A_2)/dxdQ^2}{d^2\sigma^{\rm NC}(x,Q^2,A_1)/dxdQ^2} \simeq \frac{F_2(x,Q^2,A_2)}{F_2(x,Q^2,A_1)} = R_{F_2}\left(x,Q^2,A_1,A_2\right)$$

- Nuclear effects encoded in PDFs so that

$$F_2(x,Q^2,A) = \frac{1}{A}(ZF_2^{(p/A)} + (A-Z)F_2^{(n/A)})$$

(Isoscalar nuclei): $F_2(x,Q^2,A) = \frac{1}{2}(F_2^{(p/A)} + F_2^{(n/A)})$

- In collinear factorization:

$$F_2^{(\rm NLO)}(x,Q^2,A) = C_\Sigma \otimes \Sigma(x,Q^2,A) + C_{T_8} \otimes T_8(x,Q^2,A) + C_g \otimes g(x,Q^2,A)$$

Neutral Current (NC) DIS – Theory

- Experimental observables given as ratios of cross sections/structure functions

$$\frac{d^2\sigma^{\rm NC}(x,Q^2,A_2)/dxdQ^2}{d^2\sigma^{\rm NC}(x,Q^2,A_1)/dxdQ^2} \simeq \frac{F_2(x,Q^2,A_2)}{F_2(x,Q^2,A_1)} = R_{F_2}\left(x,Q^2,A_1,A_2\right)$$

- Nuclear effects encoded in PDFs so that

$$F_2(x,Q^2,A) = \frac{1}{A}(ZF_2^{(p/A)} + (A-Z)F_2^{(n/A)})$$

(Isoscalar nuclei): $F_2(x,Q^2,A) = \frac{1}{2}(F_2^{(p/A)} + F_2^{(n/A)})$

- In collinear factorization:

$$F_2^{(\rm NLO)}(x,Q^2,A) = C_\Sigma \otimes \Sigma(x,Q^2,A) + C_{T_8} \otimes T_8(x,Q^2,A) + C_g \otimes g(x,Q^2,A)$$

Coefficient functions (including relevant charge factors)

Neutral Current (NC) DIS – Theory

- Experimental observables given as ratios of cross sections/structure functions

$$\frac{d^2\sigma^{\rm NC}(x,Q^2,A_2)/dxdQ^2}{d^2\sigma^{\rm NC}(x,Q^2,A_1)/dxdQ^2} \simeq \frac{F_2(x,Q^2,A_2)}{F_2(x,Q^2,A_1)} = R_{F_2}\left(x,Q^2,A_1,A_2\right)$$

- Nuclear effects encoded in PDFs so that

$$F_2(x,Q^2,A) = \frac{1}{A}(ZF_2^{(p/A)} + (A-Z)F_2^{(n/A)})$$

(Isoscalar nuclei): $F_2(x,Q^2,A) = \frac{1}{2}(F_2^{(p/A)} + F_2^{(n/A)})$

- In collinear factorization:

$$F_2^{(\rm NLO)}(x,Q^2,A) = C_\Sigma \otimes \Sigma(x,Q^2,A) + C_{T_8} \otimes T_8(x,Q^2,A) + C_g \otimes g(x,Q^2,A)$$

Coefficient functions (including relevant charge factors)

Nuclear PDFs (in DGLAP evolution basis)

$$\Sigma(x,Q^2,A) \equiv \sum_{i=1}^{n_f=3} f_i^+(x,Q^2,A) \qquad \text{(quark singlet)}\,,$$

$$T_3(x,Q^2,A) \equiv \left(u^+ - d^+\right)(x,Q^2,A) \qquad \text{(quark triplet)}\;,$$

$$T_8(x,Q^2,A) \equiv \left(u^+ + d^+ - 2s^+\right)(x,Q^2,A) \qquad \text{(quark octet)}$$

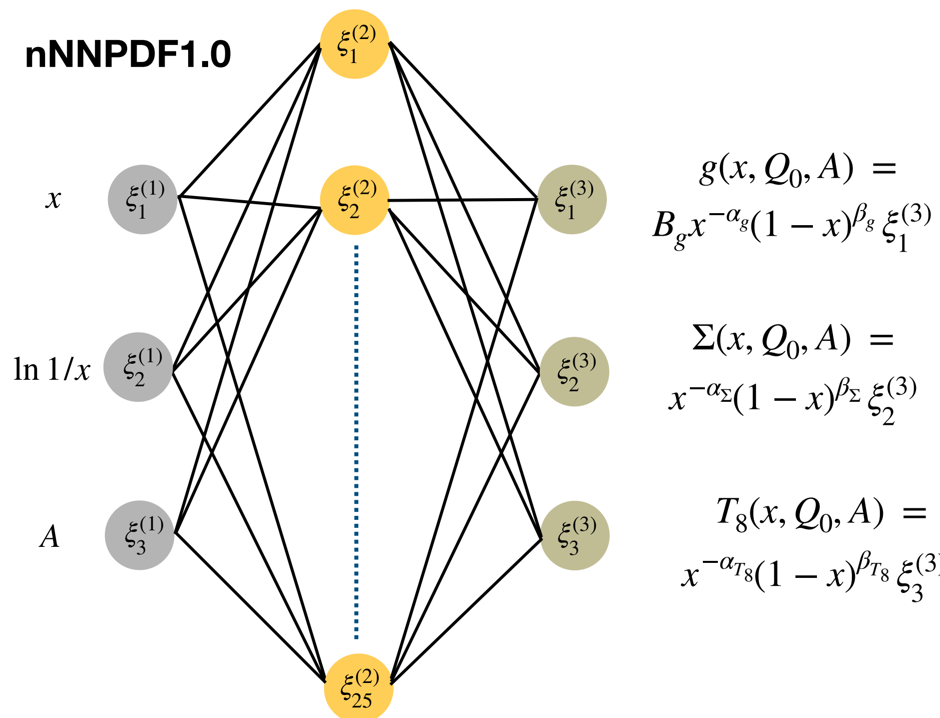
nPDF Parameterization

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

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

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

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



- Exponents treated as free parameters
- Momentum Sum Rule:

$$\int_0^1 dx\, x(\Sigma(x, A) + g(x, A)) = 1 \;\rightarrow\; B_g = \frac{1 - \int_0^1 dx\, x\Sigma(x, A)}{\int_0^1 dx\, x g(x, A)}$$