Towards a Neural Network determination of nuclear PDFs

Rabah Abdul Khalek
Departement of Physics and Astronomy, Vrije Universiteit Amsterdam
Nikhef Theory Group

Diffraction and low-x 2018
26 Aug - 1 Sep

Physics Motivation

- **Simulations** of relativistic nucleus-nucleus reactions [LHC, RHIC, etc.]

- Understanding the **EMC effect.**

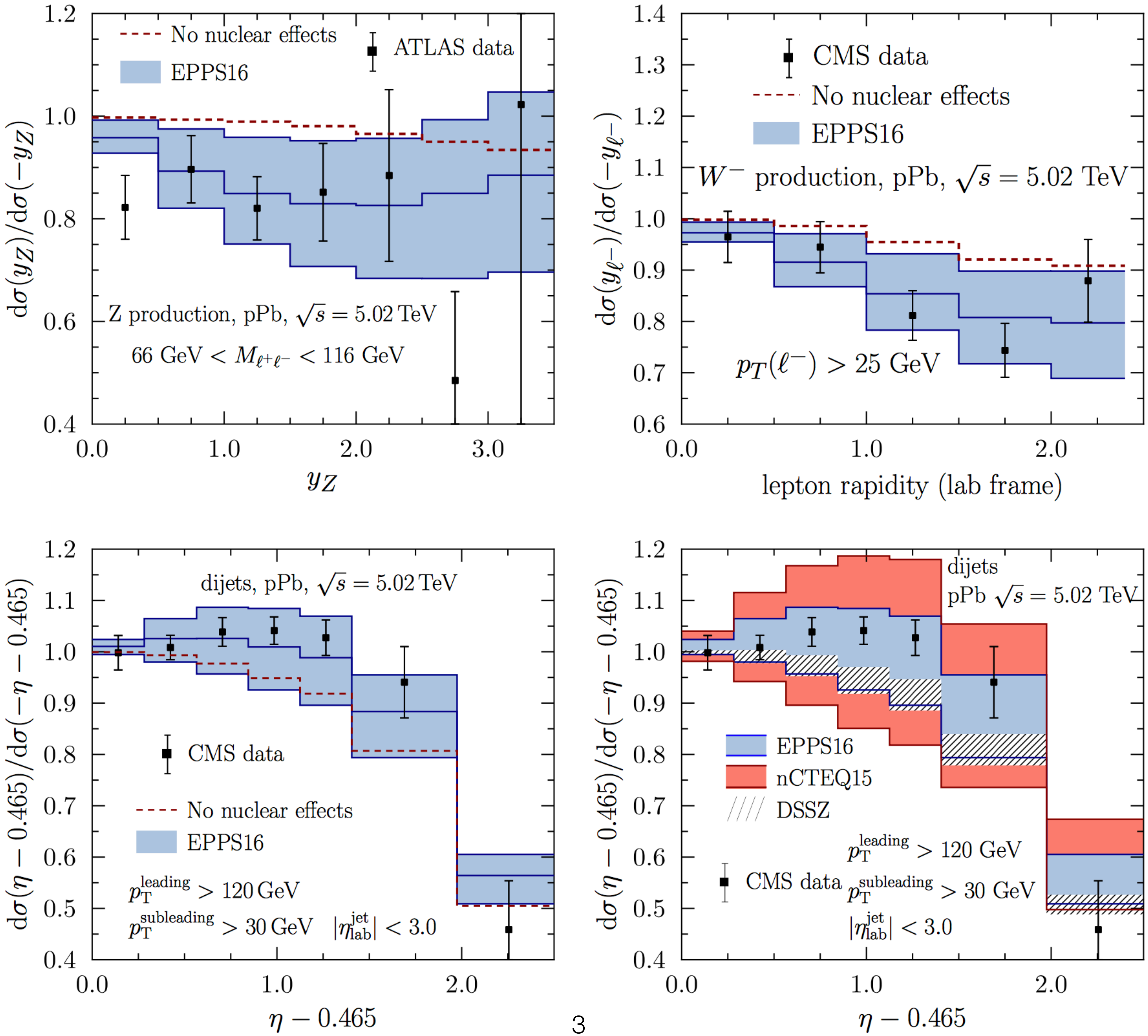
- **Lead, Xenon PDF** for the LHC heavy-ion program.

- **Combining data** across different nuclear targets and provide maximum information for the proton PDFs especially **strangeness**.

- Heavy nuclear targets for νA DIS measurements (statistics) Important data for the **separation of PDFs flavors**.

nPDF in LHC pPb observables

[1612.05741]



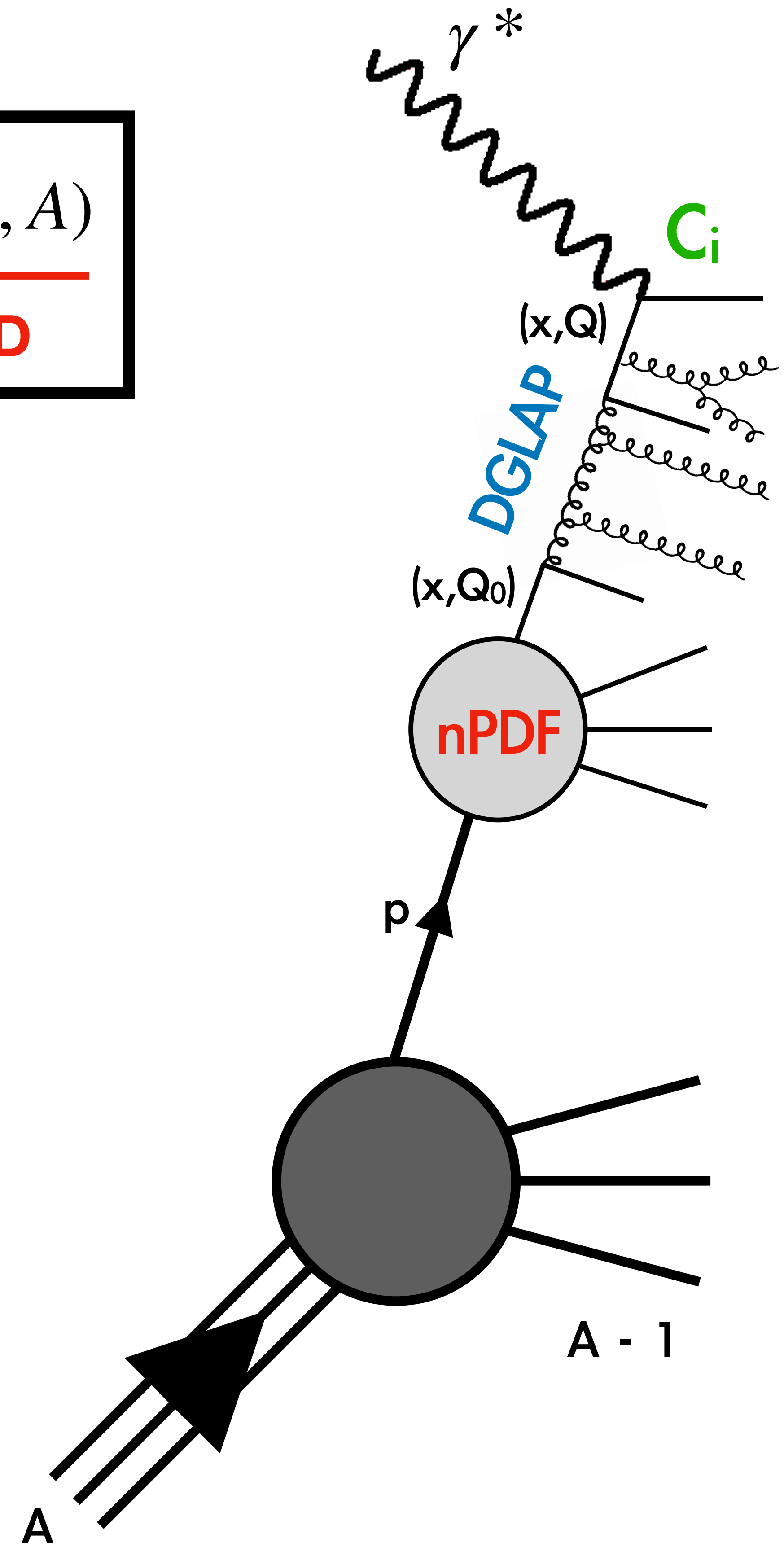
Data

	EPS09	DSSZ12	KA15	NCTEQ15	EPPS16	nNNPDF1.0
Order in α_s	LO & NLO	NLO	NNLO	NLO	NLO	NLO
Neutral current DIS ℓ+A/ℓ+d	✓	✓	✓	✓	✓	✓
Drell-Yan dilepton p+A/p+d	✓	✓	✓	✓	✓	
RHIC pions d+Au/p+p	✓	✓		✓	✓	
Neutrino-nucleus DIS		✓			✓	
Drell-Yan dilepton π+A					✓	
LHC p+Pb jet data					✓	
LHC p+Pb W, Z data					✓	
Q cut in DIS	1.3 GeV	1 GeV	1 GeV	2 GeV	1.3 GeV	1.3 GeV
datapoints	929	1579	1479	708	1811	605
free parameters	15	25	16	17	20	70
error analysis	Hessian	Hessian	Hessian	Hessian	Hessian	MC
error tolerance $\Delta\chi^2$	50	30	not given	35	52	
Free proton baseline PDFs	CTEQ6.1	MSTW2008	JR09	CTEQ6M-like	CT14NLO	NNPDF3.1
Heavy-quark effects		✓		✓	✓	✓
Flavor separation				some	✓	some
Reference	[JHEP 0904 065]	[PR D85 074028]	[PR D93, 014026]	[PR D93 085037]	[EPJ C77 163]	Preliminary

Nuclear NC Inclusive DIS

EM F_2 ($Q < M_Z$)

$$F_2(x, Q^2, A) = \sum_i^{n_f} \sum_j^{n_f} \underbrace{C_i(x, Q^2)}_{\text{pQCD}} \otimes \underbrace{\Gamma_{ij}(x, Q^2)}_{\text{DGLAP}} \otimes \underbrace{q_j(x, Q_0^2, A)}_{\text{npQCD}}$$



X

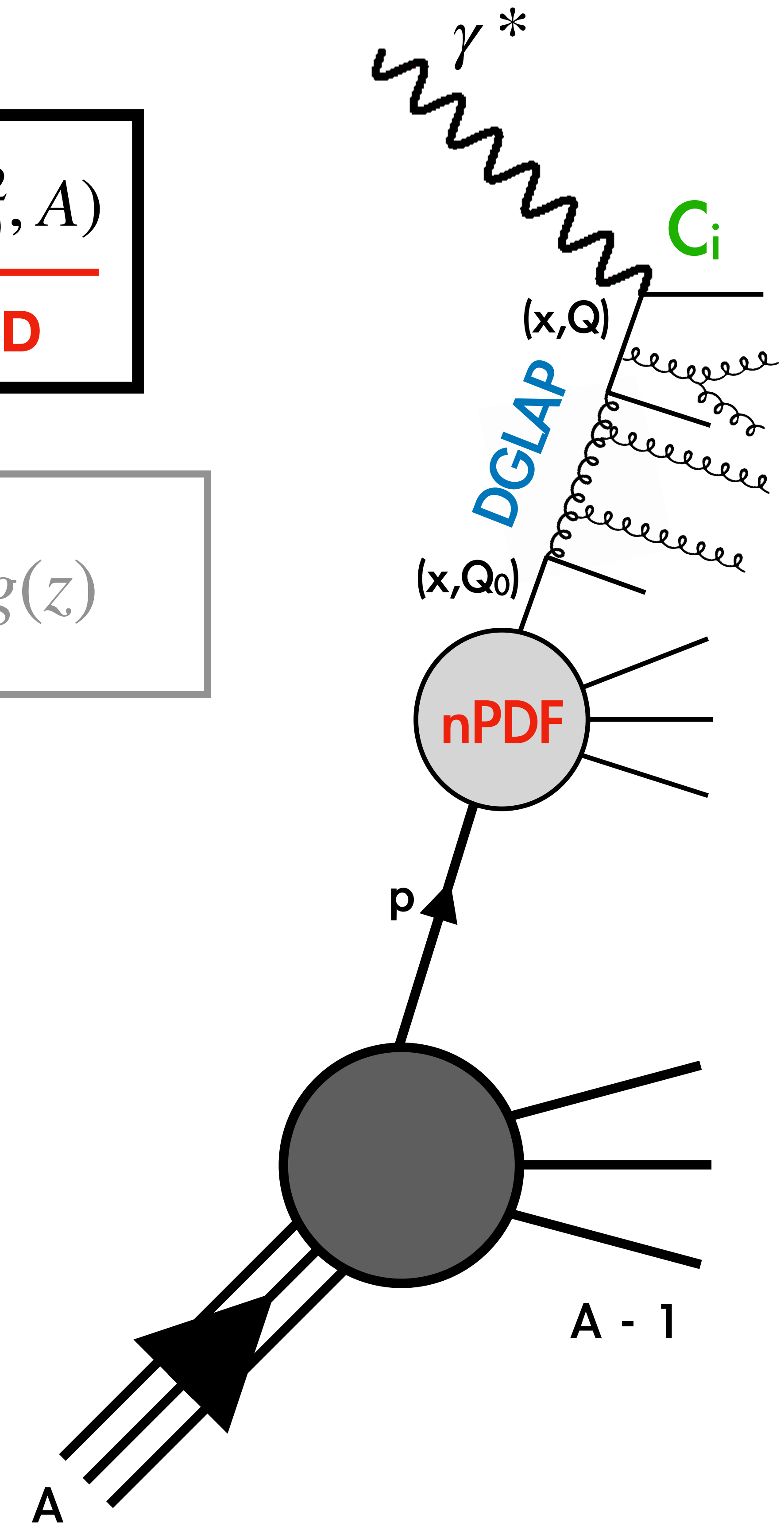
Nuclear NC Inclusive DIS

EM F2 (Q < MZ)

$$F_2(x, Q^2, A) = \sum_i^{n_f} \sum_j^{n_f} \underbrace{C_i(x, Q^2)}_{\text{pQCD}} \otimes \underbrace{\Gamma_{ij}(x, Q^2)}_{\text{DGLAP}} \otimes \underbrace{q_j(x, Q_0^2, A)}_{\text{npQCD}}$$

Convolution

$$f \otimes g = \int_x^1 \frac{dz}{z} f\left(\frac{x}{z}\right) g(z)$$



X

Nuclear NC Inclusive DIS

EM F_2 ($Q < M_Z$)

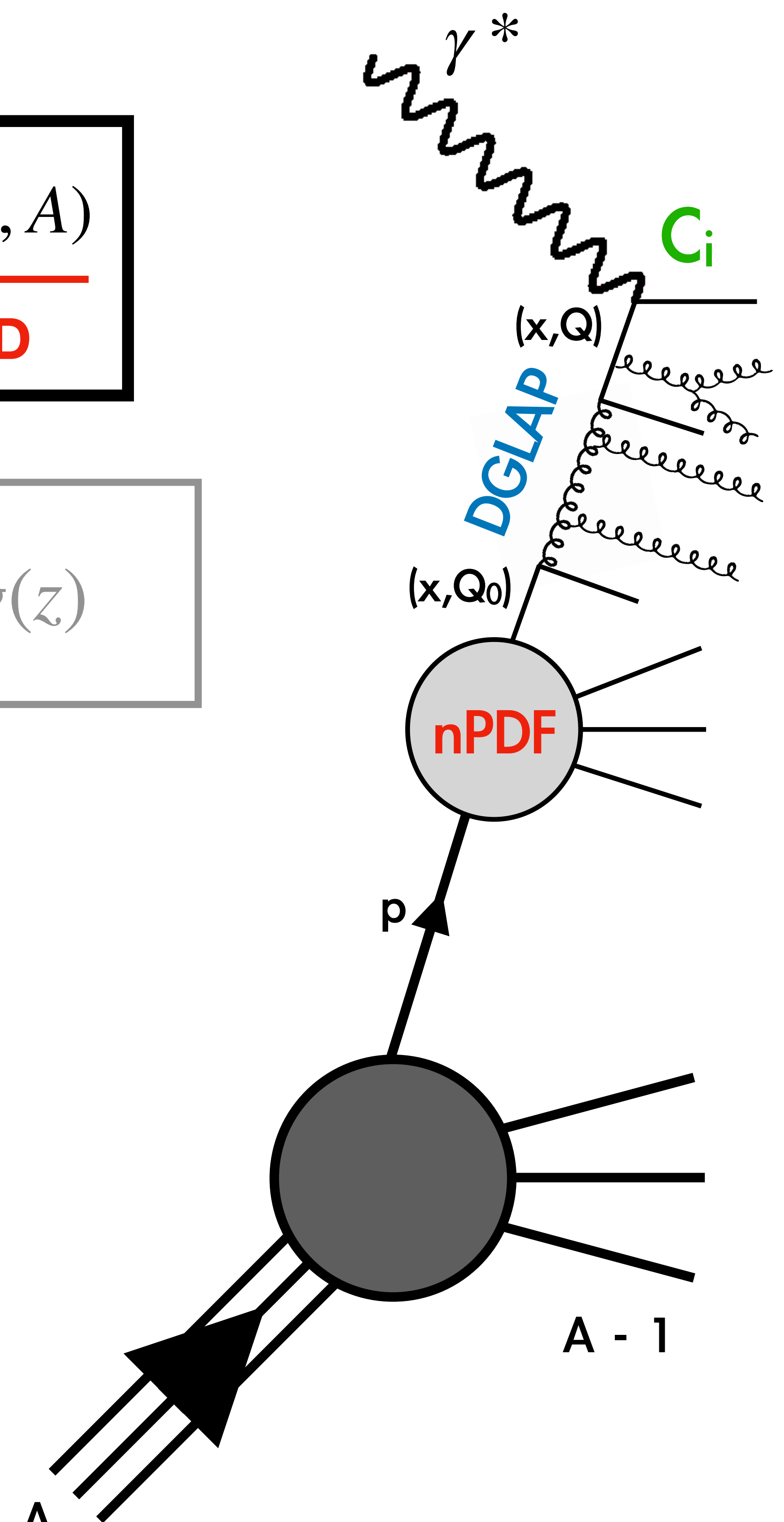
$$F_2(x,Q^2,A) = \sum_i^{n_f}\sum_j^{n_f} \underbrace{C_i(x,Q^2)}_{\text{pQCD}} \otimes \underbrace{\Gamma_{ij}(x,Q^2)}_{\text{DGLAP}} \otimes \underbrace{q_j(x,Q_0^2,A)}_{\text{npQCD}}$$

Coefficient Functions

$$C_i = C_i^{(0)} + \frac{\alpha_s}{4\pi} C_i^{(1)} + O(\alpha_s^2)$$

Convolution

$$f \otimes g = \int_x^1 \frac{dz}{z} f\left(\frac{x}{z}\right) g(z)$$



x

Nuclear NC Inclusive DIS

EM F2 (Q < MZ)

$$F_2(x, Q^2, A) = \sum_i^{n_f} \sum_j^{n_f} \underbrace{C_i(x, Q^2)}_{\text{pQCD}} \otimes \underbrace{\Gamma_{ij}(x, Q^2)}_{\text{DGLAP}} \otimes \underbrace{q_j(x, Q_0^2, A)}_{\text{npQCD}}$$

Coefficient Functions

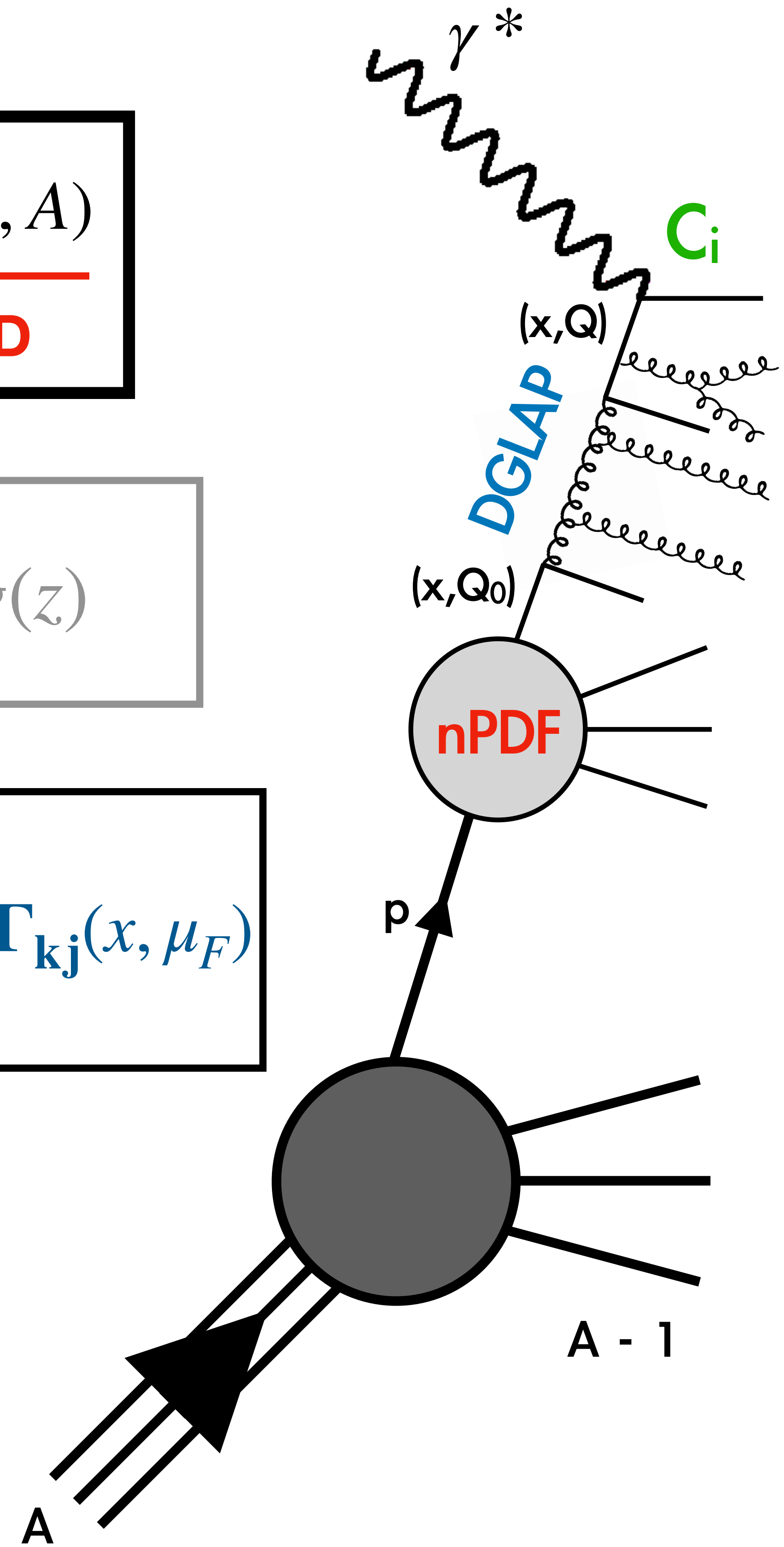
$$C_i = C_i^{(0)} + \frac{\alpha_s}{4\pi} C_i^{(1)} + O(\alpha_s^2)$$

Convolution

$$f \otimes g = \int_x^1 \frac{dz}{z} f(\frac{x}{z}) g(z)$$

DGLAP equation

$$\frac{\partial \Gamma_{\mathbf{ij}}}{\partial \ln \mu_F^2} = \sum_k \frac{\alpha_s(\mu_F)}{4\pi} \left[P_{ik}^{(0)}(x) + \frac{\alpha_s(\mu_F)}{4\pi} P_{ik}^{(1)}(x) + \ldots \right] \otimes \Gamma_{\mathbf{kj}}(x, \mu_F)$$



X

Nuclear NC Inclusive DIS

EM F_2 ($Q < M_Z$)

$$F_2(x, Q^2, A) = \sum_i^{n_f} \sum_j^{n_f} \underbrace{C_i(x, Q^2)}_{\text{pQCD}} \otimes \underbrace{\Gamma_{ij}(x, Q^2)}_{\text{DGLAP}} \otimes \underbrace{q_j(x, Q_0^2, A)}_{\text{npQCD}}$$

Coefficient Functions

$$C_i = C_i^{(0)} + \frac{\alpha_s}{4\pi} C_i^{(1)} + O(\alpha_s^2)$$

Convolution

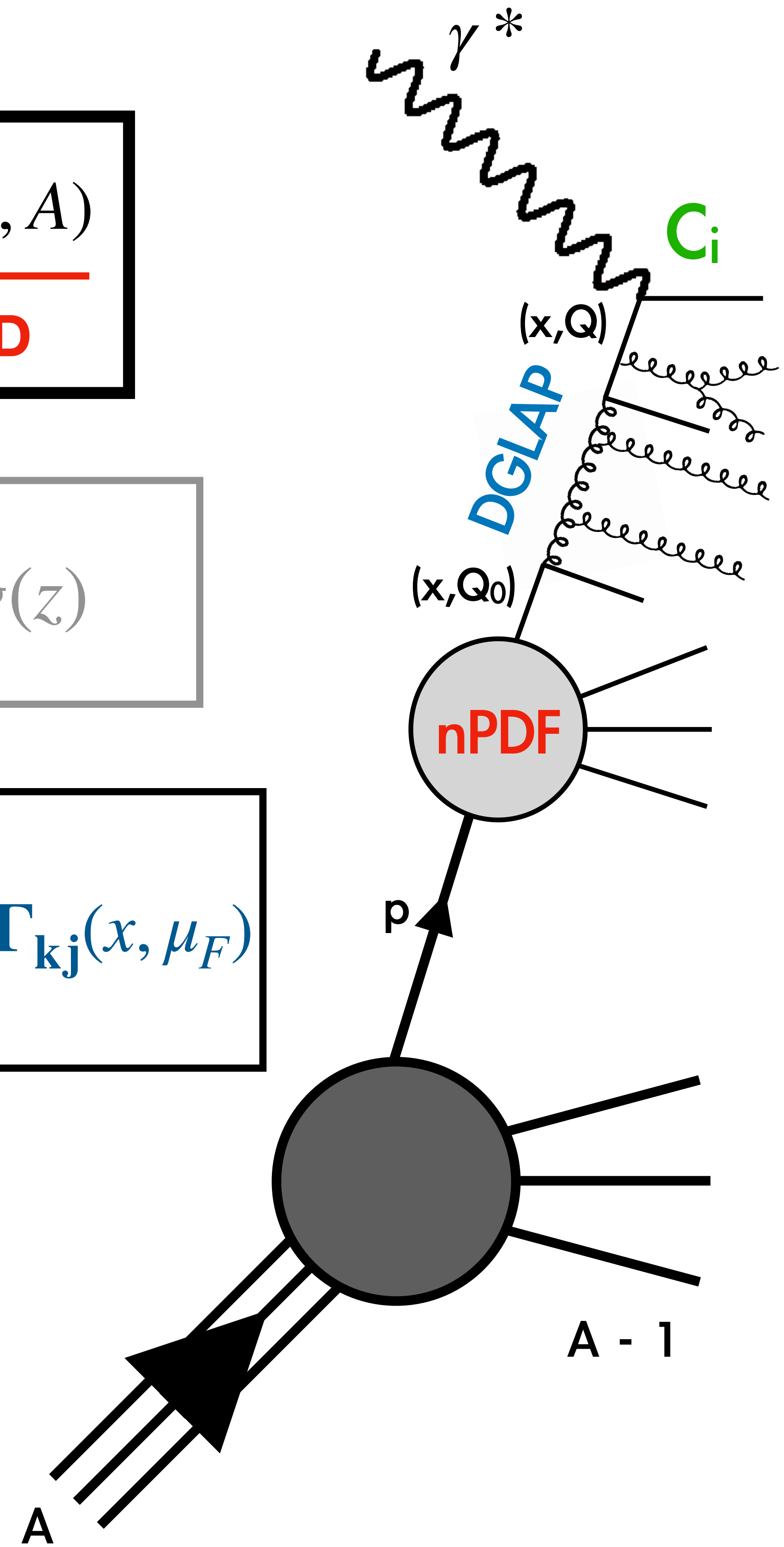
$$f \otimes g = \int_x^1 \frac{dz}{z} f\left(\frac{x}{z}\right) g(z)$$

DGLAP equation

$$\frac{\partial \mathbf{\Gamma_{ij}}}{\partial \ln \mu_F^2} = \sum_k \frac{\alpha_{\mathbf{s}}(\mu_{\mathbf{F}})}{4\pi} \left[P_{ik}^{(0)}(x) + \frac{\alpha_{\mathbf{s}}(\mu_{\mathbf{F}})}{4\pi} P_{ik}^{(1)}(x) + \ldots \right] \otimes \mathbf{\Gamma_{kj}}(x, \mu_F)$$

RG equation

$$\frac{\partial \alpha_{\mathbf{s}}}{4\pi \cdot \partial \ln \mu_R^2} = -\left(\frac{\alpha_s(\mu_R)}{4\pi} \right)^2 \left[\beta_0 + \frac{\alpha_s(\mu_R)}{4\pi} \beta_1 + \ldots \right]$$



Nuclear NC Inclusive DIS

EM F_2 ($Q < M_Z$)

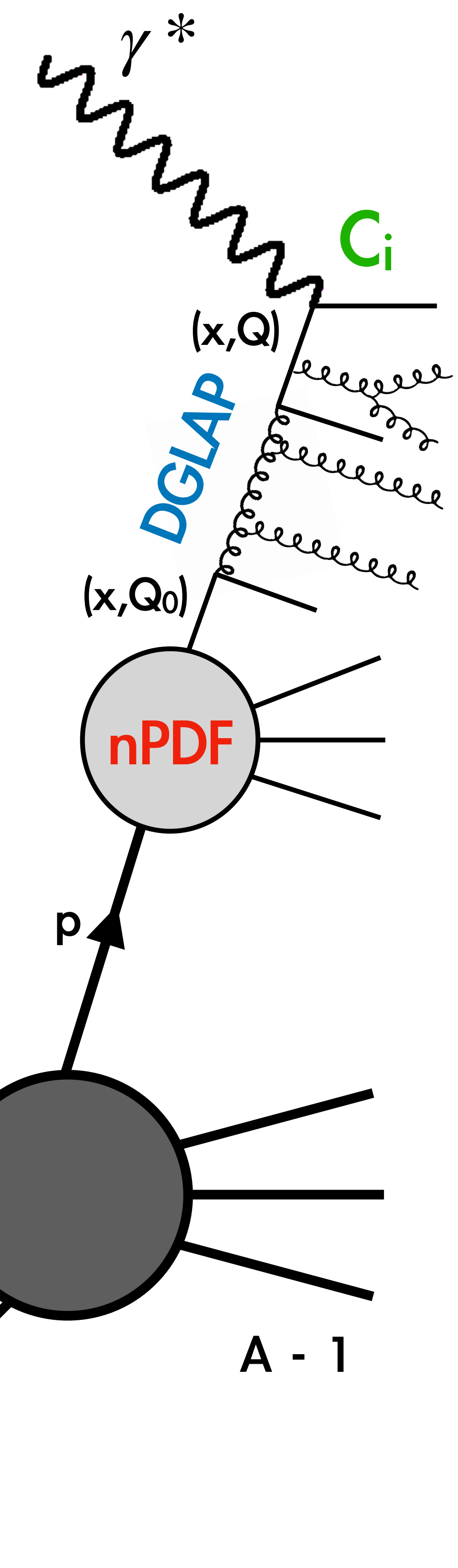
$$F_2(x, Q^2, A) = \sum_i^{n_f} \sum_j^{n_f} \underbrace{C_i(x, Q^2)}_{\text{pQCD}} \otimes \underbrace{\Gamma_{ij}(x, Q^2)}_{\text{DGLAP}} \otimes \underbrace{q_j(x, Q_0^2, A)}_{\text{npQCD}}$$

$$= C^S_{2,q}(x, \alpha_s(Q)) \otimes \Sigma(x, Q) \longrightarrow \textbf{Singlet}$$

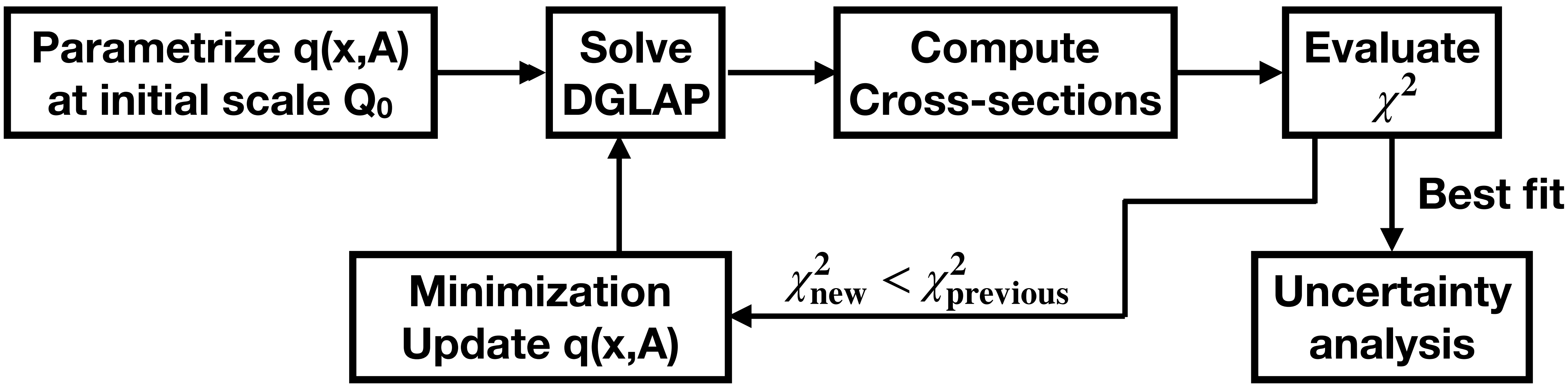
$$+ C^{NS}_{2,q}(x, \alpha_s(Q)) \otimes T(x, Q^2) \longrightarrow \textbf{Non-Singlet}$$

$$+ C^S_{2,g}(x, \alpha_s(Q)) \otimes g(x, Q^2) \longrightarrow \textbf{Gluon}$$

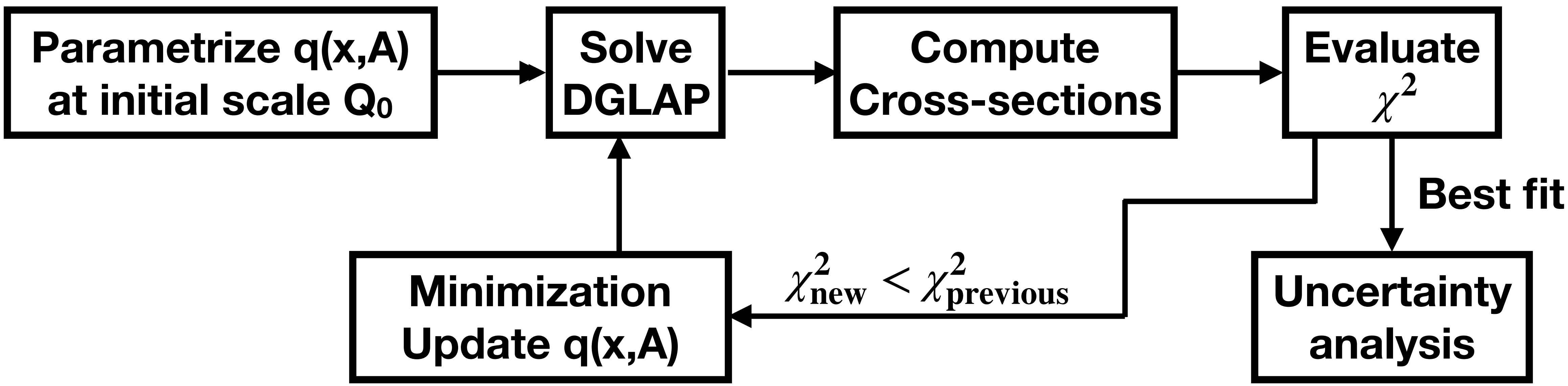
<u>Theoretically,</u> 3 Independent PDFs [backup]
<u>Practically,</u> depends on data kinematic range



Methodology (1)

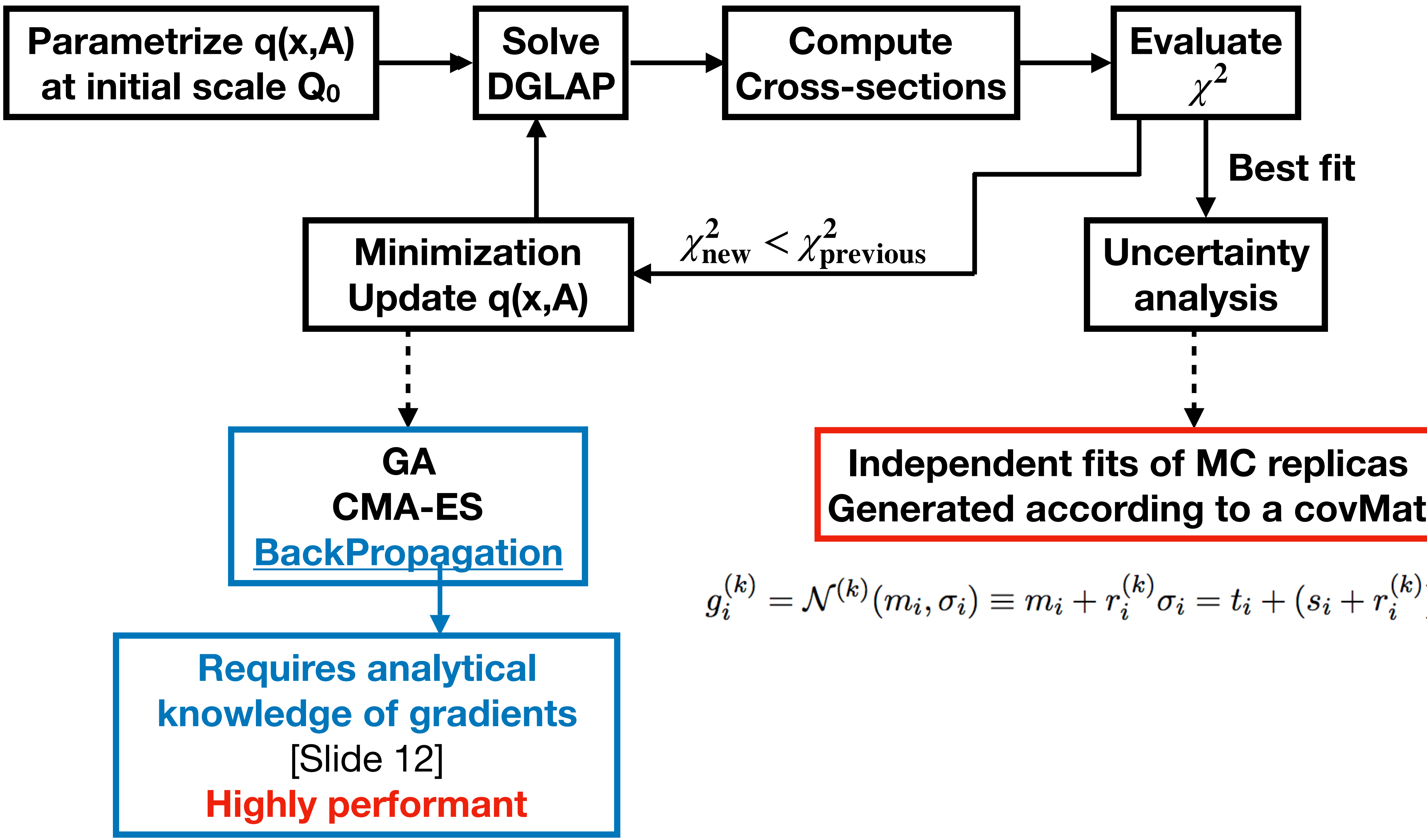


Methodology (1)



$$\chi^2 = \frac{1}{N_{data}} \sum_{n,m}^{N_{data}} (\hat{F}_2^{(n)} - F_2^{(n)})\, Cov_{nm}^{-1}\, (\hat{F}_2^{(m)} - F_2^{(m)})$$

Methodology (1)



$$g_i^{(k)} = \mathcal{N}^{(k)}(m_i, \sigma_i) \equiv m_i + r_i^{(k)}\sigma_i = t_i + (s_i + r_i^{(k)})\sigma_i$$

Parametrizations

EM F_2 ($Q < M_Z$)

$$F_2(x, Q^2, A) = \sum_i^{n_f} \sum_j^{n_f} C_i(x, Q^2) \otimes \Gamma_{ij}(x, Q^2) \otimes q_j(x, Q_0^2, A)$$

$$= \sum_i^{n_f} \widetilde{C}_i(x, Q^2) \otimes \mathbf{q_i(x, Q_0^2, A)}$$

Pre-computed FK tables [NNPDF methodology]

Parametrizations

EM F_2 ($Q < M_Z$)

$$F_2(x, Q^2, A) = \sum_i^{n_f} \sum_j^{n_f} C_i(x, Q^2) \otimes \Gamma_{ij}(x, Q^2) \otimes q_j(x, Q_0^2, A)$$

$$= \sum_i^{n_f} \widetilde{C}_i(x, Q^2) \otimes \mathbf{q_i}(\mathbf{x}, \mathbf{Q_0^2}, \mathbf{A})$$

Pre-computed FK tables [NNPDF methodology]

PDF - Proton case

$$q_i^p(x, Q_0^2) = x f_i(x, Q_0^2) \propto \mathbf{x}^{\alpha}(\mathbf{1} - \mathbf{x})^{\beta}\mathbf{NN}(\mathbf{x})$$

Parametrizations

EM F2 (Q < Mz)

$$F_2(x,Q^2,A) = \sum_i^{n_f}\sum_j^{n_f} C_i(x,Q^2)\otimes\Gamma_{ij}(x,Q^2)\otimes q_j(x,Q_0^2,A)$$

$$= \sum_i^{n_f}\widetilde{C}_i(x,Q^2)\otimes \mathbf{q_i(x,Q_0^2,A)}$$

Pre-computed FK tables [NNPDF methodology]

PDF - Proton case

$$q_i^p(x,Q_0^2) = xf_i(x,Q_0^2) \propto \mathbf{x^{\alpha}(1-x)^{\beta}NN(x)}$$

nPDF - Direct Fit

$$q_i(x,A,Q_0^2) = \mathbf{NN_i(x,A,Q_0^2)}$$

Constraint

$$\mathbf{-NN_i(x,A=1,Q_0^2)+q_i^p(x,Q_0^2)}$$

nPDF - Nuclear Modification Fit

$$q_i(x,A,Q_0^2) = \mathbf{NN_i(x,A)q^p(x,Q_0)}$$

Status of recent nPDFs

EPPS16 [1612.05741]

$$f_i^{\mathrm{p}/A}(x,Q^2) = R_i^A(x,Q^2) f_i^{\mathrm{p}}(x,Q^2)$$

$$R_i^A(x,Q_0^2) = \begin{cases} a_0 + a_1(x-x_a)^2 & x \le x_a \\ b_0 + b_1 x^\alpha + b_2 x^{2\alpha} + b_3 x^{3\alpha} & x_a \le x \le x_e \\ c_0 + (c_1 - c_2 x)(1-x)^{-\beta} & x_e \le x \le 1, \end{cases}$$

nCTEQ15 [1509.00792]

$$x f_i^{p/A}(x,Q_0) = c_0 x^{c_1} (1-x)^{c_2} e^{c_3 x} (1+e^{c_4} x)^{c_5},$$
$$\text{for} \quad i = u_v, d_v, g, \bar{u}+\bar{d}, s+\bar{s}, s-\bar{s},$$

$$\frac{\bar{d}(x,Q_0)}{\bar{u}(x,Q_0)} = c_0 x^{c_1} (1-x)^{c_2} + (1+c_3 x)(1-x)^{c_4}.$$

$$c_k \to c_k(A) \equiv c_{k,0} + c_{k,1}\left(1 - A^{-c_{k,2}}\right),$$
$$k = \{1, \ldots, 5\}.$$

X

Status of recent nPDFs

EPPS16 [1612.05741]

$$f_i^{\mathrm{p}/A}(x,Q^2) = \underline{R_i^A}(x,Q^2) f_i^{\mathrm{p}}(x,Q^2)$$

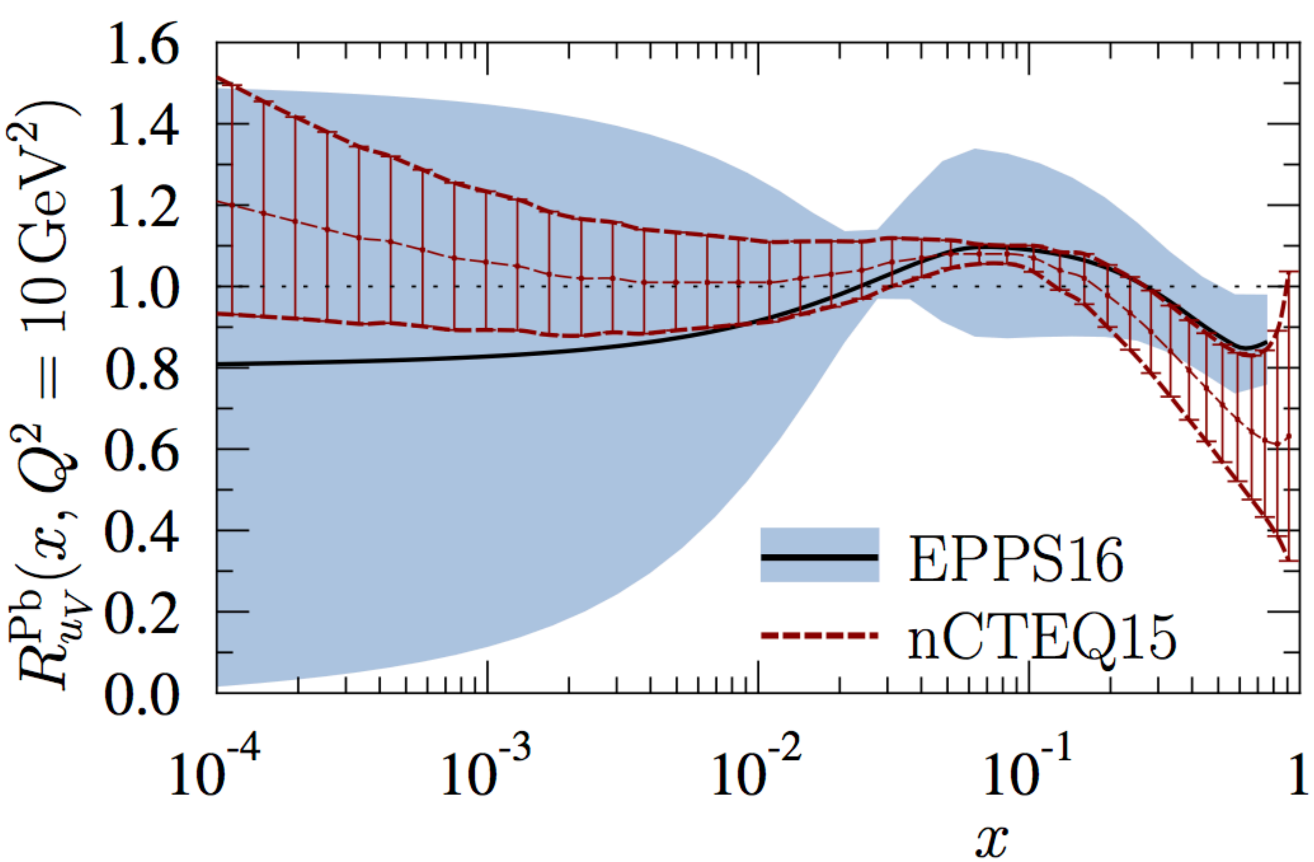
$$R_i^A(x,Q_0^2) = \begin{cases} a_0 + a_1(x-x_a)^2 & x \le x_a \\ b_0 + b_1 x^\alpha + b_2 x^{2\alpha} + b_3 x^{3\alpha} & x_a \le x \le x_e \\ c_0 + (c_1 - c_2 x)(1-x)^{-\beta} & x_e \le x \le 1, \end{cases}$$

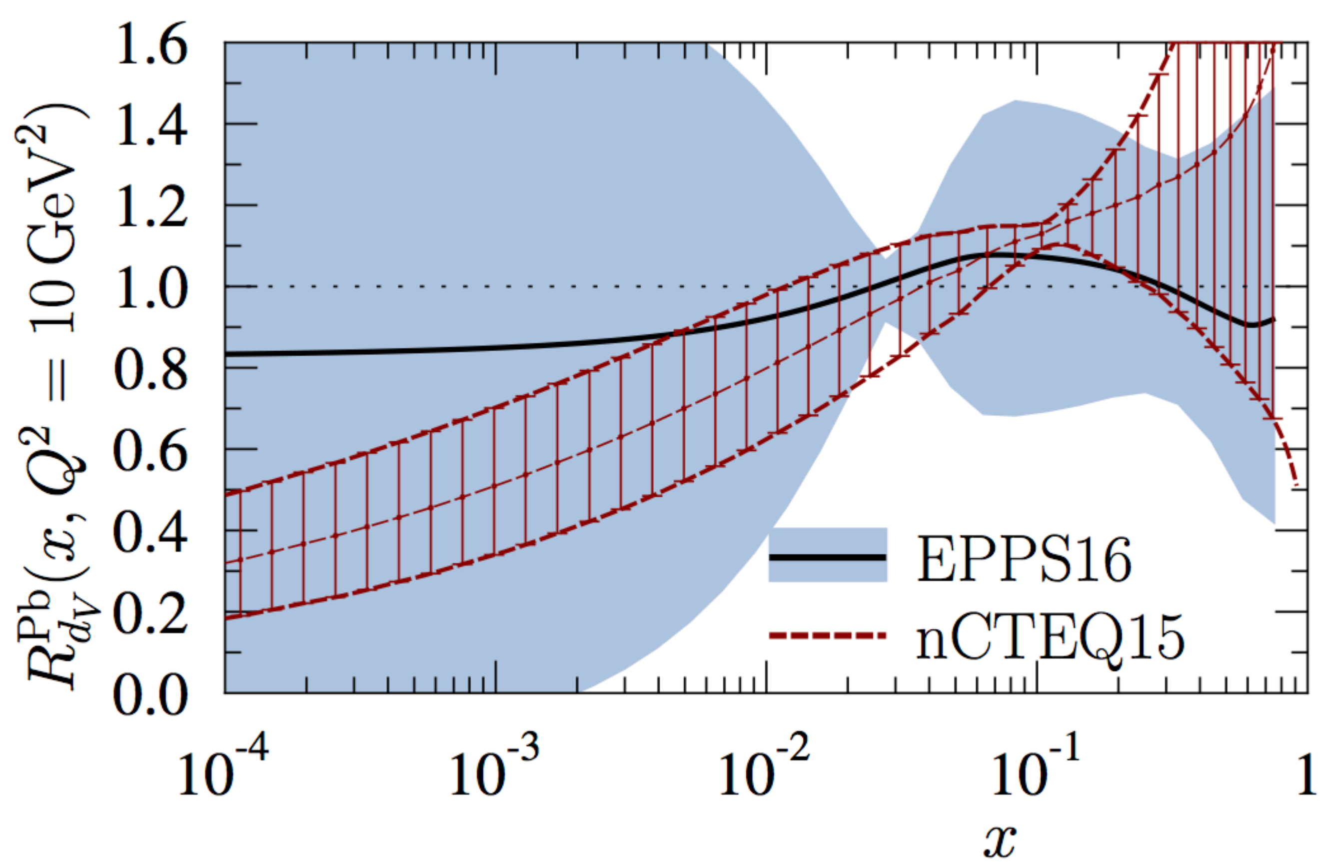
nCTEQ15 [1509.00792]

$$x f_i^{p/A}(x,Q_0) = c_0 x^{c_1}(1-x)^{c_2} e^{c_3 x}(1+e^{c_4}x)^{c_5},$$
$$\text{for} \quad i = u_v, d_v, g, \bar{u}+\bar{d}, s+\bar{s}, s-\bar{s},$$

$$\frac{\bar{d}(x,Q_0)}{\bar{u}(x,Q_0)} = c_0 x^{c_1}(1-x)^{c_2} + (1+c_3 x)(1-x)^{c_4}.$$

$$c_k \to c_k(A) \equiv c_{k,0} + c_{k,1}\left(1 - A^{-c_{k,2}}\right), \quad k = \{1,\ldots,5\}.$$





Neural Network

Direct Fit

$$\hat{F}_2(x, Q^2, A) = \sum_{i}^{n_f} \widetilde{C}_i(x, Q^2) \otimes \mathbf{NN_i(x, Q_0^2, A \,|\, \{w, b\})}$$

{w, b} are the weights and thresholds of the NN

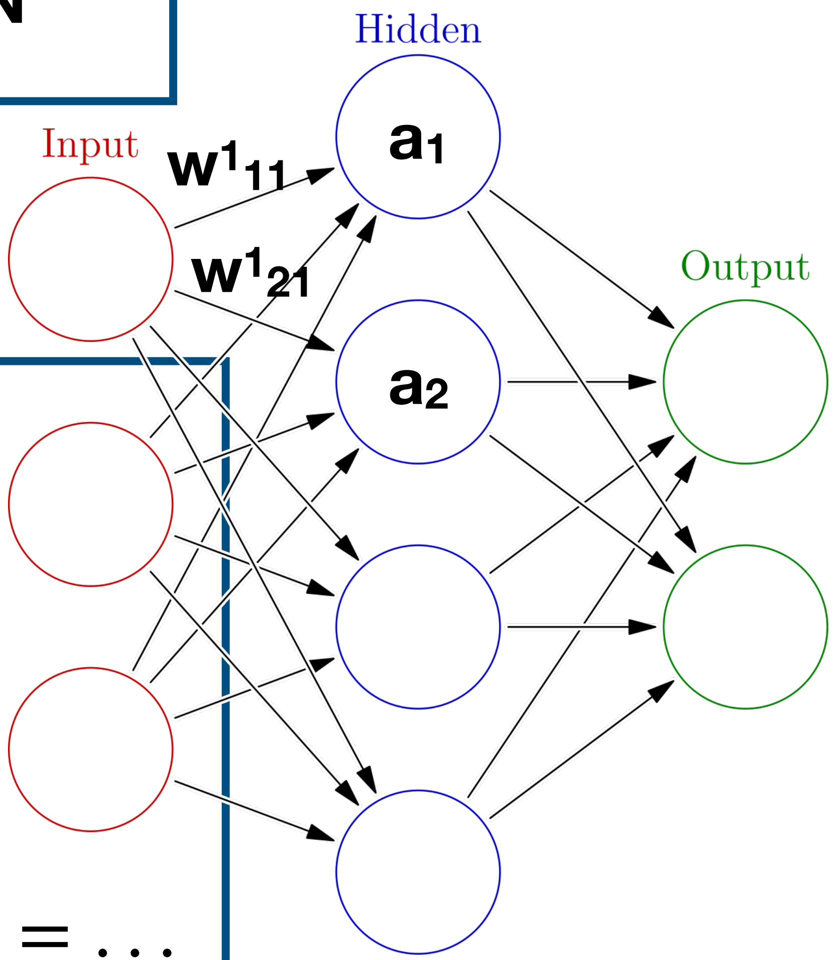
X

Neural Network

Direct Fit

$$\hat{F}_2(x, Q^2, A) = \sum_i^{n_f} \widetilde{C}_i(x, Q^2) \otimes \mathbf{NN_i(x, Q_0^2, A \,|\, \{w, b\})}$$

{w, b} are the weights and thresholds of the NN



Neural Network

$$NN_i(x; \{w_{ij}^{(\ell)}, b_i^{(\ell)}\}) = \sigma_L \left(\sum_j^{N_{L-1}} w_{ij}^{(L)} a_j^{(L-1)} + b_i^{(L)} \right)$$

$$= \sigma_L \left(\sum_j^{N_{L-1}} w_{ij}^{(L)} \sigma_{L-1} \left(\sum_k^{N_{L-2}} w_{jk}^{(L-1)} a_k^{(L-2)} + b_j^{(L-1)} \right) + b_i^{(L)} \right) = \ldots$$

Backpropagation

Gradient Descent

$$w_{ij} = w_{ij} - \frac{\eta}{N_{data}} \sum_{n}^{N_{data}} \frac{\partial \chi^{2(n)}}{\partial w_{ij}}$$

$$b_i = b_i - \frac{\eta}{N_{data}} \sum_{n}^{N_{data}} \frac{\partial \chi^{2(n)}}{\partial b_i}$$

χ^2 Derivatives

$$\frac{\partial \chi^2}{\partial w_{mn}^{(\ell)}} = 2\left(\frac{\widetilde{\mathbf{C}} \otimes \mathbf{NN} - F}{\sigma^2}\right) \widetilde{\mathbf{C}} \otimes \frac{\partial \mathbf{NN}}{\partial w_{mn}^{(\ell)}}$$

$$\frac{\partial \chi^2}{\partial b_{m}^{(\ell)}} = 2\left(\frac{\widetilde{\mathbf{C}} \otimes \mathbf{NN} - F}{\sigma^2}\right) \widetilde{\mathbf{C}} \otimes \frac{\partial \mathbf{NN}}{\partial b_{m}^{(\ell)}}$$

NN Derivatives

$$\frac{\partial \mathbf{NN}}{\partial w_{ij}^{(\ell)}} = \left[\prod_{\alpha=L}^{\ell+1} \mathbf{W}^{(\alpha)}\right] \cdot \frac{\partial \mathbf{a}^{(\ell)}}{\partial \mathbf{w}_{\mathbf{ij}}^{(\ell)}}$$

$$\frac{\partial \mathbf{NN}}{\partial b_{i}^{(\ell)}} = \left[\prod_{\alpha=L}^{\ell+1} \mathbf{W}^{(\alpha)}\right] \cdot \frac{\partial \mathbf{a}^{(\ell)}}{\partial \mathbf{b}_{\mathbf{i}}^{(\ell)}}$$

Neurones Derivatives

$$\frac{\partial \mathbf{a}_{\mathbf{k}}^{(\ell)}}{\partial \mathbf{w}_{\mathbf{ij}}^{(\ell)}} = \sigma'_{\ell}\left(z_k^{(\ell)}\right) \delta_{ik} a_j^{(\ell-1)}$$

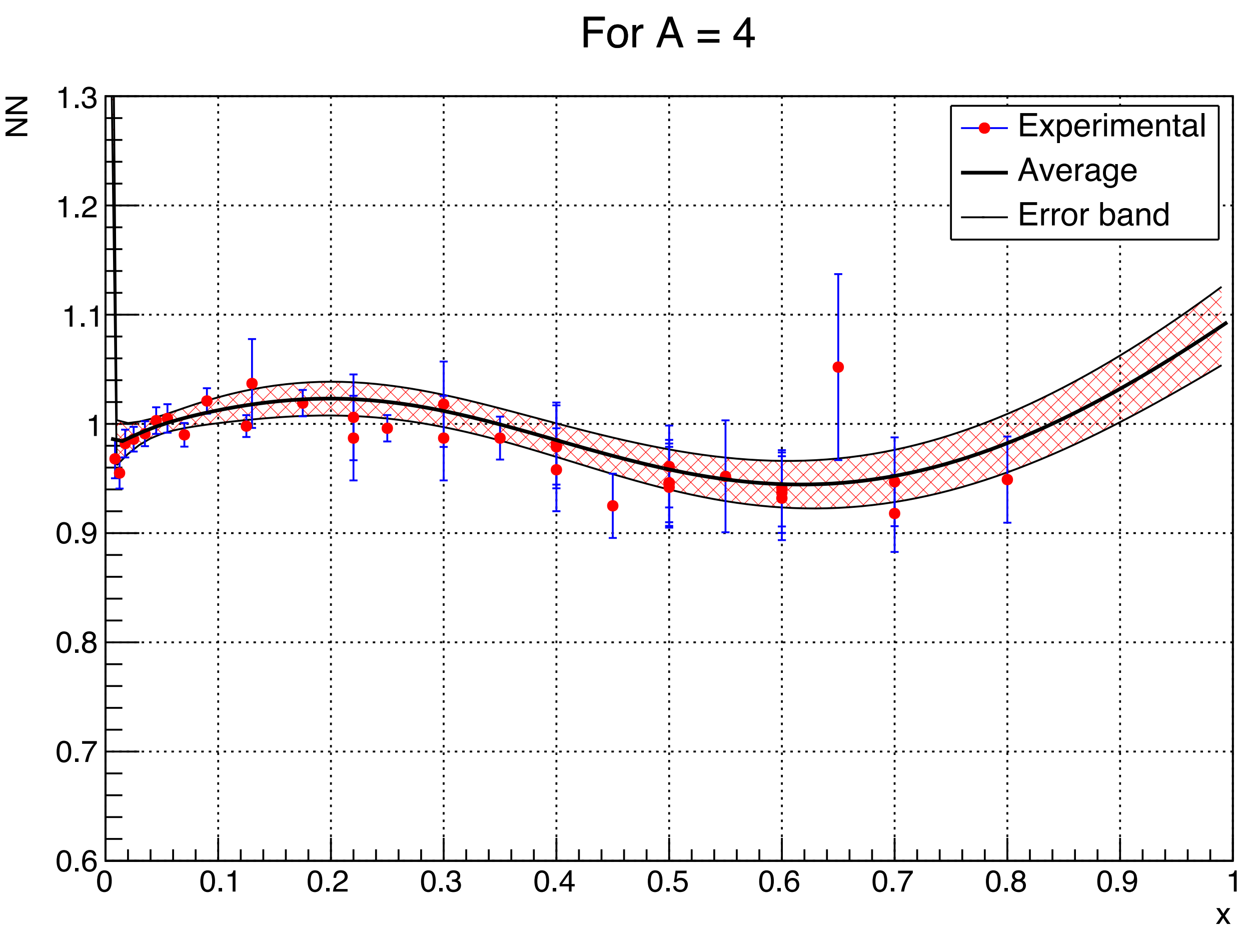
$$\frac{\partial \mathbf{a}_{\mathbf{k}}^{(\ell)}}{\partial \mathbf{b}_{\mathbf{i}}^{(\ell)}} = \sigma'_{\ell}\left(z_k^{(\ell)}\right) \delta_{ik}$$

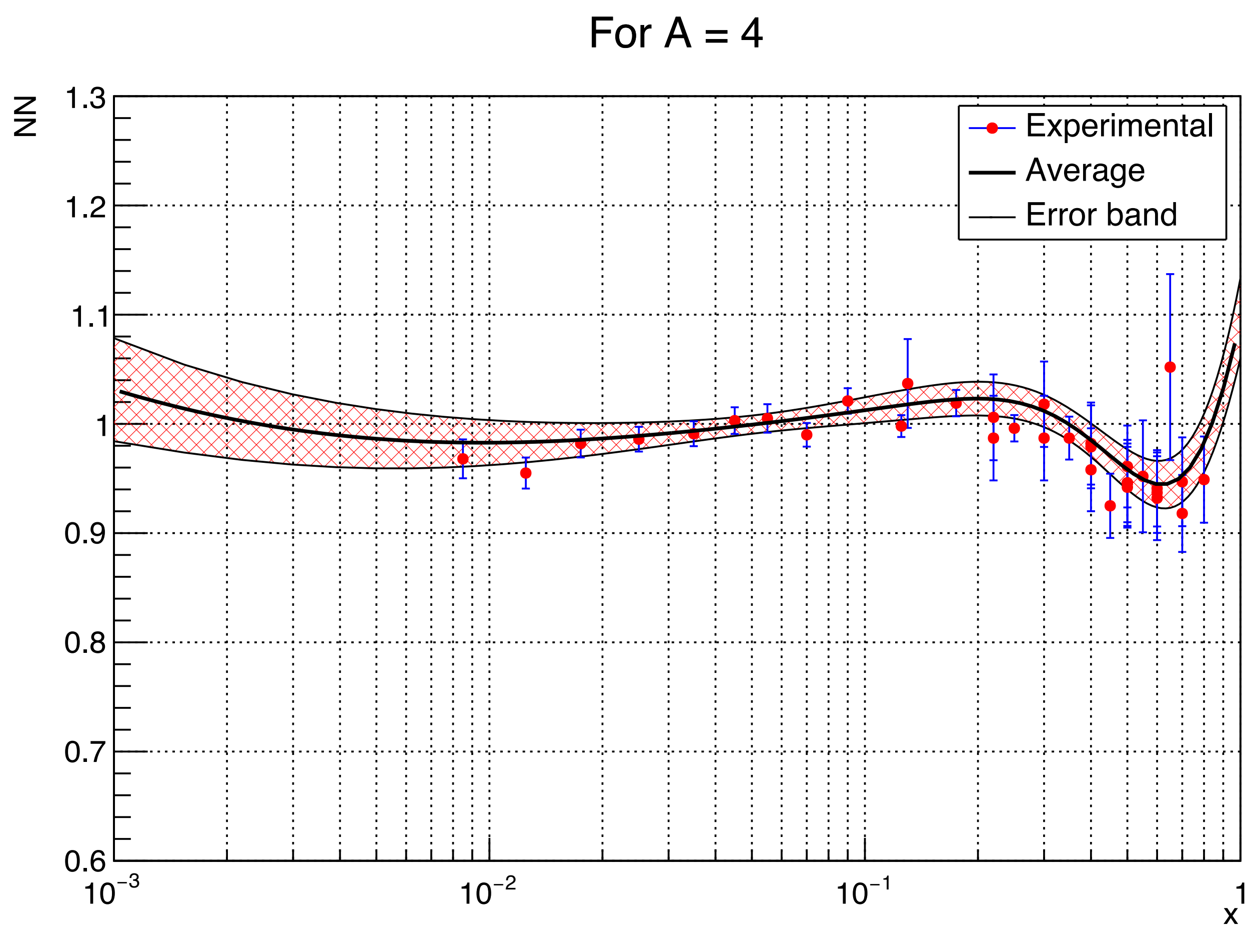
Results (1) - WarmUp

Fitting F_2^A/F_2^D Ratio

$$NN(x, A) = \frac{F_2^A(x)}{F_2^D(x)}$$

2D fit in (A, x) - Uncorrelated uncertainty

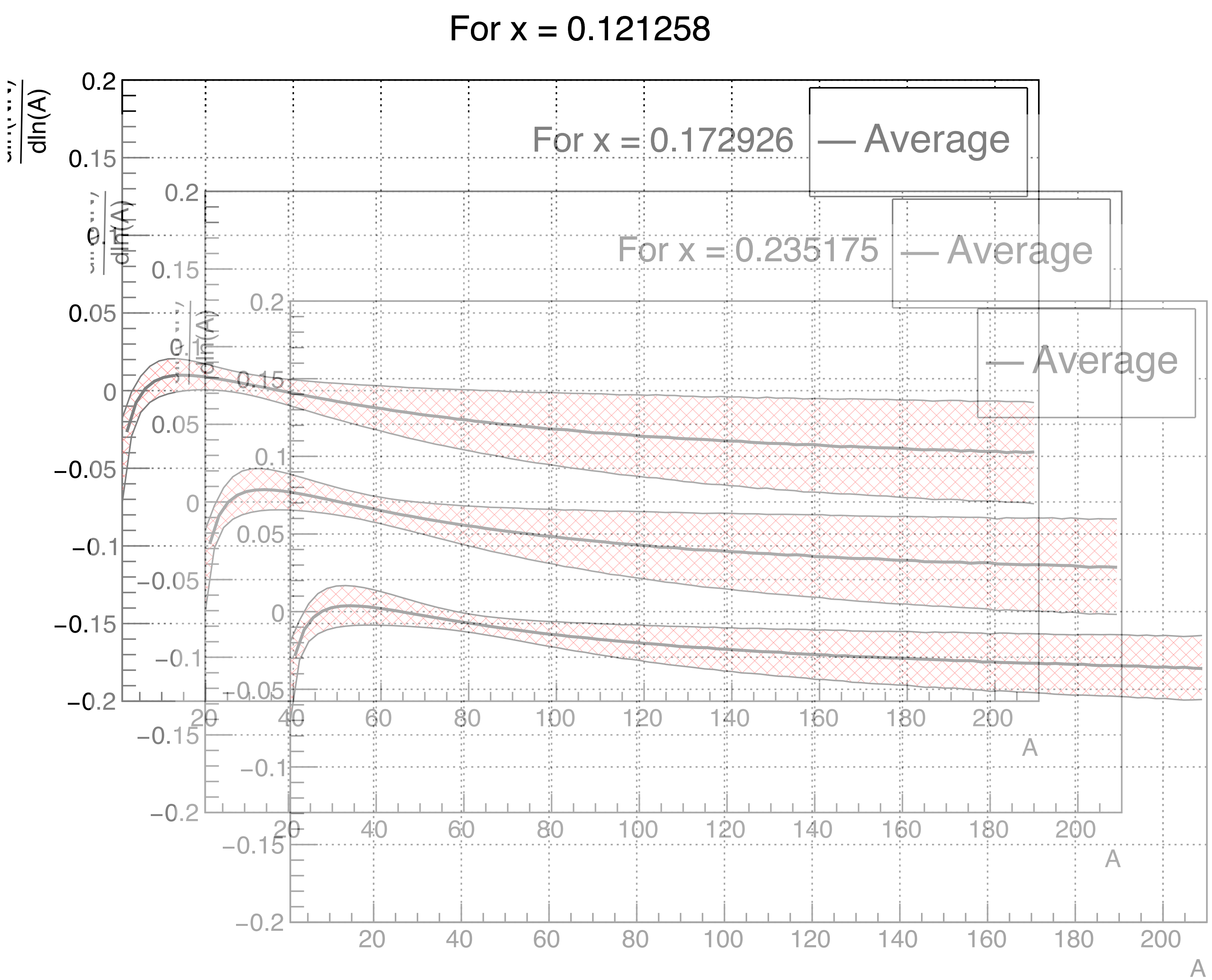




Results (2) - WarmUp

Constraints on A-dependance

Testing the assumption $\mathrm{NN(x,A)} = A^n f(x)$ **via** $\dfrac{d\,ln(NN(x,A))}{d\,ln(A)} = n(x)$



Outlook

- Reproducing EPPS16 and nCTEQ15 results to **verify the theory implementation** (FK tables).

EMC (He/D)	APFEL	EPPS16
χ^2_{total}	**18.16**	**18.0**

16 Datapoints

- Running **DIS-only fits** using the NNPDF methodology.

- Trying different NN **architectures** and different **parametrizations** + **constraints**, e.g at A=1, nPDF = PDF

- DIS-only Fits should be available in the **next few weeks.** HardProbes conference in Oct (1-5)

- Next, correlated uncertainties, different processes.

Backup

$$F_2(x,Q^2,A) = \sum_i^{n_f} C_i(x,Q^2) \otimes q_i(x,Q^2,A) \tag{2}$$

where $\otimes$ denotes the Mellin convolution defined as:

$$f(x) \otimes g(x) = \int_0^1 dy \int_0^1 dz f(y) g(z) \delta(x - yz) = \int_x^1 \frac{dy}{y} f(y) g(\frac{x}{y}) = \int_x^1 \frac{dz}{z} f(\frac{x}{z}) g(z) \tag{3}$$

We can factorize the nPDF dependance out of the convolution via expansion over a set of interpolating functions, spanning Q^2 and x such as:

$$q_i(x,Q^2,A) = \sum_\beta \sum_\tau q_{i,\beta\tau} I_\beta(x) I_\tau(Q^2) \tag{4}$$

where the nPDF $q_{i,\beta\tau}$ may be expressed as a product of nPDF at some initial fitting scale Q_0 and an evolution operator obtained by the solution of the DGLAP equation via the interpolation procedure as:

$$q_{i,\beta\tau} \equiv q_i(x_\beta, Q^2_\tau, A) = \sum_j \sum_\alpha \Gamma^\tau_{ij,\alpha\beta}\, q_j(x_\alpha, Q_0^2, A) \tag{5}$$

Finally:

$$F_2(x,Q^2,A) = \sum_i^{n_f} C_i(x,Q^2) \otimes [\sum_\alpha \sum_\tau \sum_j \sum_\beta \Gamma^\tau_{ij,\alpha\beta}\, q_j(x_\alpha, Q_0^2) I_\beta(x) I_\tau(Q^2)] \tag{6}$$

contracting the sum over τ, i and β then replacing j with i :

$$F_2(x,Q^2) = \sum_i^{n_f} \sum_\alpha^{n_x} FK_{i,\alpha}(x, x_\alpha, Q^2, Q_0^2) q_i(x_\alpha, Q_0^2) \tag{7}$$

where:

$$FK_{j,\alpha} = \sum_i C_i(x,Q^2) \otimes [\sum_\tau \sum_\beta \Gamma^\tau_{ij,\alpha\beta}\, I_\beta(x) I_\tau(Q^2)] \tag{8}$$

Backup

At LO, DGLAP basis

$$F_2^\gamma(x, Q^2, A) = \frac{1}{A}\left(Z\,\mathbf{F_2^p} + (A-Z)\,\mathbf{F_2^n}\right)$$

$$= \frac{5}{18}\Sigma - \left(\frac{\mathbf{Z}}{\mathbf{3A}} - \frac{1}{6}\right)T_3 + \frac{1}{18}(T_8 - T_{15}) + \frac{1}{30}(T_{24} - T_{35})$$

$$= \frac{2}{9}\Sigma - \left(\frac{\mathbf{Z}}{\mathbf{3A}} - \frac{1}{9}\right)T_3$$

Evolution ($Q_0 < M_c$)

$$\Sigma(Q^2) = \Gamma_{qq}\,\Sigma(Q_0^2) + \Gamma_{qg}\,g(Q_0^2)$$

$$g(Q^2) = \Gamma_{gq}\,\Sigma(Q_0^2) + \Gamma_{gg}\,g(Q_0^2)$$

$$\Gamma^+ = \Gamma_{T_3} = \Gamma_{T_8}$$

$$T_{15}(Q_0^2) = T_{24}(Q_0^2) = T_{35}(Q_0^2) = \Sigma(Q_0^2)$$

3 independent distributions

$\Sigma,\ g,\ T_3$

Evolution basis

$$\Sigma = \sum_{i=1}^{n_f} q_i^+$$ **where:** $q^\pm = q \pm \bar{q}$

$$T_3 = u^+ - d^+$$

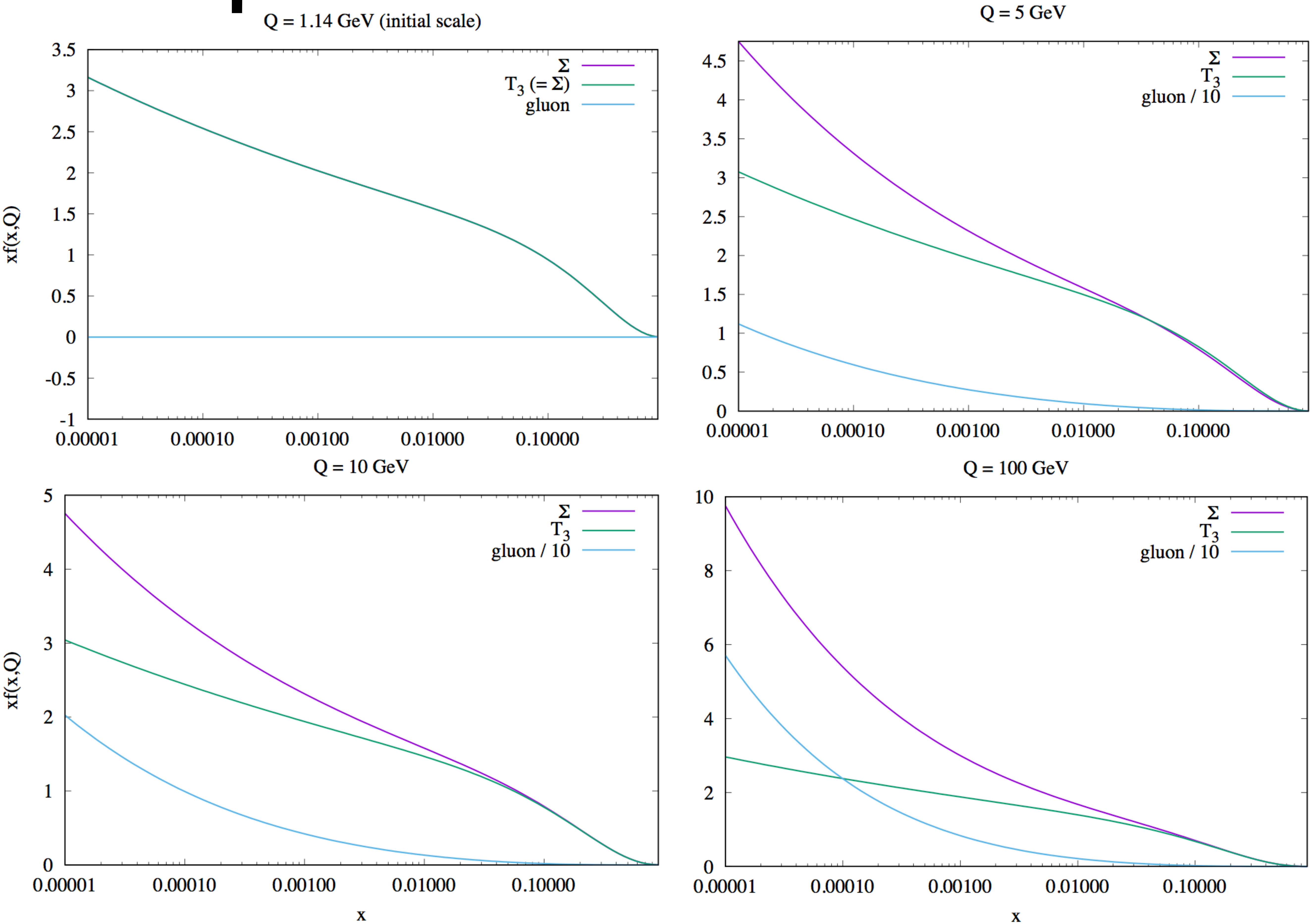
$$T_8 = u^+ + d^+ - 2s^+$$

$$T_{15} = u^+ + d^+ + s^+ - 3c^+$$

$$T_{24} = u^+ + d^+ + s^+ + c^+ - 4b^+$$

$$T_{35} = u^+ + d^+ + s^+ + c^+ + b^+ - 5t^+$$

Backup


Q = 1.14 GeV (initial scale)
Σ
T3 (= Σ)
gluon
xf(x,Q)
Q = 5 GeV
Σ
T3
gluon / 10
Q = 10 GeV
Σ
T3
gluon / 10
xf(x,Q)
x
Q = 100 GeV
Σ
T3
gluon / 10
x
0.00001
0.00010
0.00100
0.01000
0.10000