Global NLO analysis of nPDFs: EPS09 + reanalysis for the LHeC

Kari J. Eskola & Hannu Paukkunen
University of Jyväskylä & Helsinki Institute of Physics

& **Carlos A. Salgado**
Universidade de Santiago de la Compostela

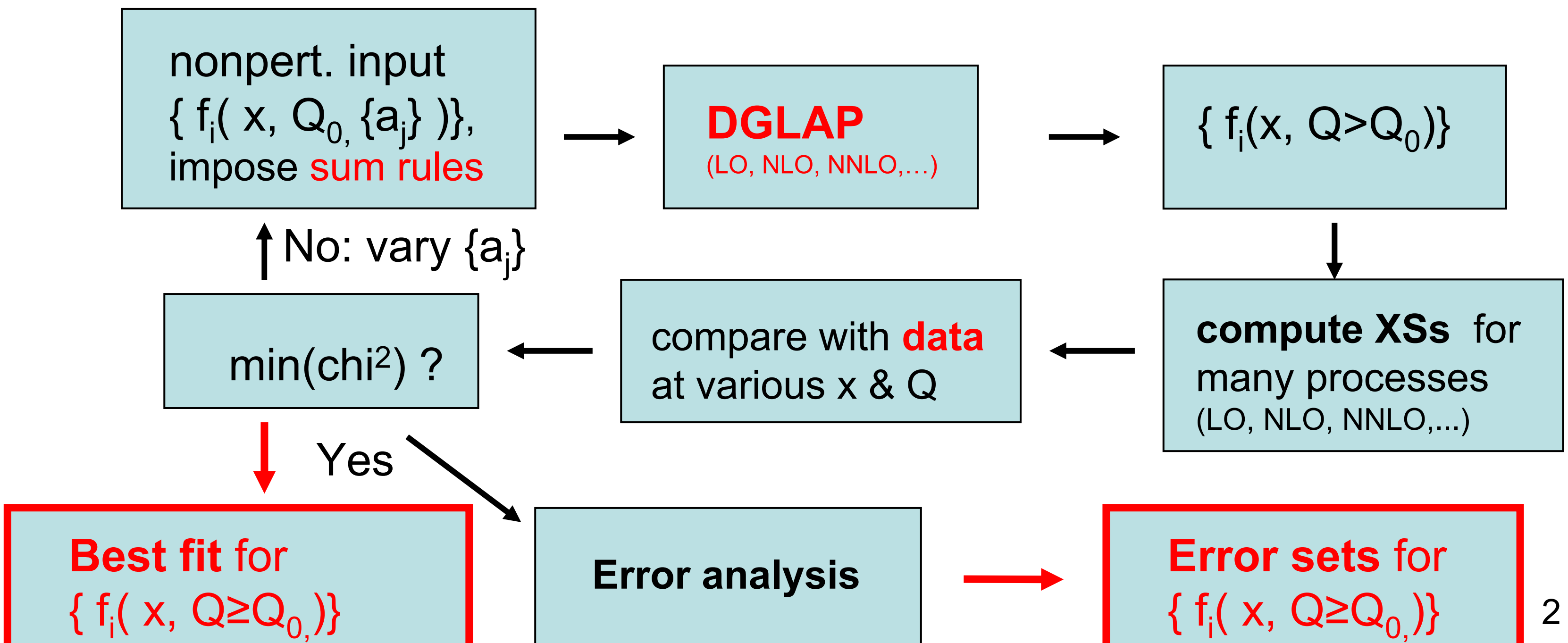
- **EPS09 – JHEP 0904:065,2009, arXiv:0902.4154 [hep-ph]**
  https://www.jyu.fi/fysiikka/en/research/highenergy/urhic/nPDFs
- **renanalysis with LHeC pseudodata included**

Global analysis of PDFs & nPDFs

- **test of pQCD & factorization** $\sigma_{AB\to h+X} = \sum_{i,j} f_i^A(Q^2) \otimes \hat{\sigma}_{ij\to h+X} \otimes f_j^B(Q^2) + \mathcal{O}\left(Q^2\right)^{-n}$

- **procedure to find a universal set $\{f_i(x,Q^2)\}$ at all x and $Q^2 >> \Lambda^2_{QCD}$**

1. Iterate until best fit, best set of initial parameters $\{a_j\}$ found
2. Perform error analysis for quantifying the PDF uncertainties



Challenges in extracting the nonperturbative input for PDFs and nPDFs

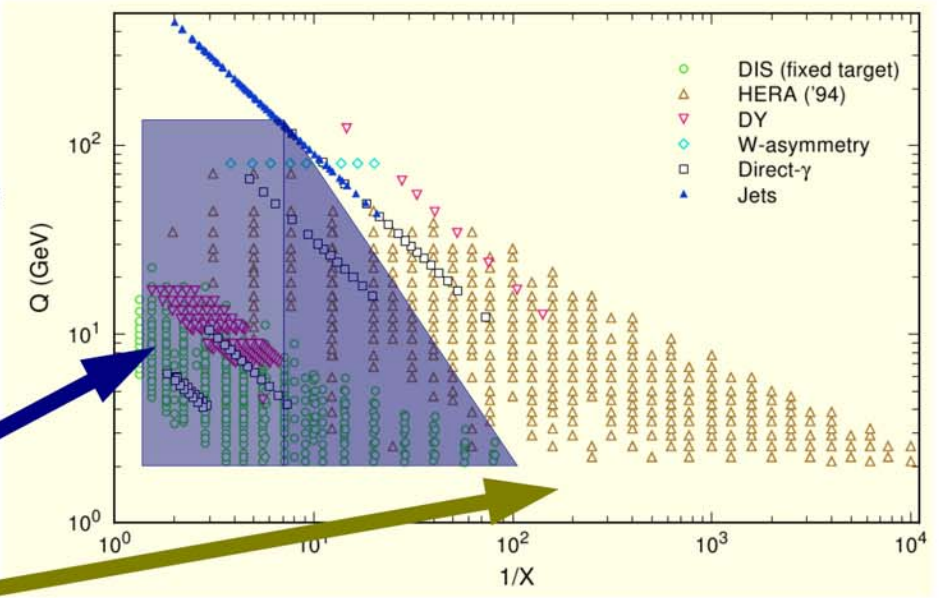
- data (=constraints) are not at one fixed Q^2 but correlated in x and Q^2 [fig.]
- how to account for the exp. statistical/systematic/normalization errors?
- parameter space 15-30 d & XSs require multi-d numerical integrations

→ need **very fast DGLAP solver and XS solvers**

– we (EPS09) have them now in NLO [see Hannus PhD thesis]

Further challenges in nPDF analysis vs. that of free-proton PDFs:

- **Smaller amount of data**
- **Not as many types of data**
- **Kinematical plane more restricted**
- **Need to parametrize the A-dependence**



Nuclear data

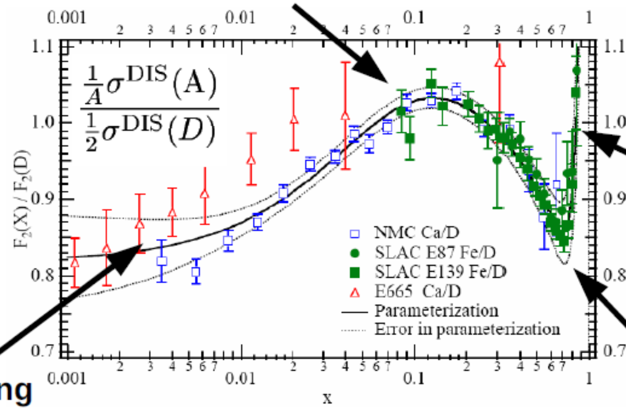
Free proton data

[from HP]

For hadron collisions, excellent global fits [CTEQ, MRST/W,…]
→ factorization theorem seems to hold well
→ PDF uncertainties have been quantified & error sets available

In nuclear hard scatterings:

$\sigma^{\text{bound nucleon}} \neq \sigma^{\text{free nucleon}}$



EPS09:

Q1: **How well does factorization work for nuclear collisions**
-- can we find process-independent NLO nPDFs such that

$$\sigma_{AB\to h+X} = \sum_{i,j} f_i^A(Q^2) \otimes \hat{\sigma}_{ij\to h+X} \otimes f_j^B(Q^2) + \mathcal{O}\left(Q^2\right)^{-n}$$

Q2: **How large are the nPDF uncertainties?**

Q3: **Release error sets for the nPDFs?**

Progress in the global nPDF analyses

1998:	EKS98	[Eskola, Kolhinen, Ruuskanen, Salgado]	LO
2001 :	HKM	[Hirai, Kumano, Miyama]	LO + error estimates
2004 :	HKN04	[Hirai, Kumano, Nagai]	LO + error est.
2004 :	nDS	[de Florian, Sassot]	NLO
2007 :	EKPS	[Eskola, Kolhinen, Paukkunen, Salgado]	LO + error est.
2007 :	HKN07	[Hirai, Kumano, Nagai]	NLO + error est.
2008 :	EPS08	[Eskola, Paukkunen, Salgado]	LO, w. RHIC data
2009 :	**EPS09**	[Eskola, Paukkunen, Salgado]	**NLO, w. RHIC data, + nPDF error sets!**

New elements in EPS09:

- with l+A DIS & DY in p+A, include also **RHIC pi0 data** from d+Au
- **error analysis** now at similar sophistication level as in CTEQ
- release the best fit + **30 error PDF sets** (for each A!) for general use

→ users can now study the **propagation of nPDF-uncertainties** into any hard XS!

EPS09 global analysis framework

- define the bound-proton PDFs vs. fixed free-p PDFs:

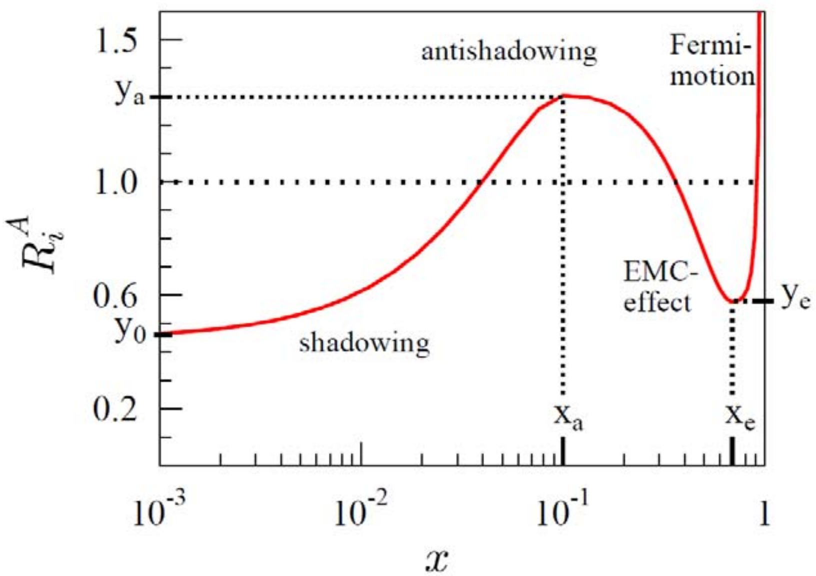
$$f_i^A(x,Q^2) \equiv \underline{R_i^A(x,Q^2)} f_i^{\mathrm{CTEQ6.1M}}(x,Q^2)$$

- MSbar & zero-mass variable flavor-number sceme
- isospin symmetry: $u_{p,n} = d_{n,p}$
- conserve momentum& baryon number
- initial parametrization at Q_0=1.3 GeV:

$R_G^A(x,Q_0^2)$ gluons

$R_V^A(x,Q_0^2)$ valence quarks

$R_S^A(x,Q_0^2)$ sea quarks



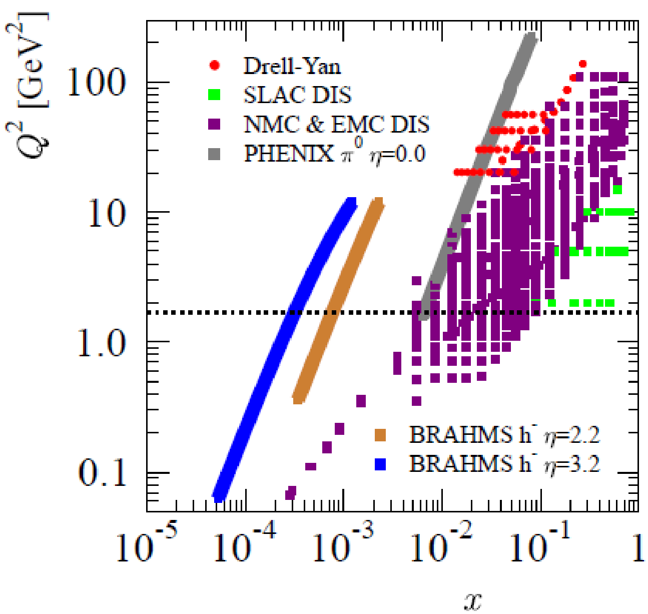
- **A-dependence** is in the parameters, e.g. $y_a = y_a(C)\,(A/12)^{p_a}$

The data in EPS09

- 3 types

DIS: l+A → l+X
DY: p+A → l^+l^-+X
pi0: d+Au → pi0+X at RHIC

Altogether
- 929 data points
- 32 data sets



Experiment	Process	Nuclei	Data points	χ^2 LO	χ^2 NLO	Weight
SLAC E-139	DIS	He(4)/D	21	6.5	7.3	1
NMC 95, re.	DIS	He/D	16	14.5	15.6	5
NMC 95	DIS	Li(6)/D	15	23.6	16.8	1
NMC 95, Q^2 dep.	DIS	Li(6)/D	153	162.2	157.0	1
SLAC E-139	DIS	Be(9)/D	20	9.6	12.2	1
NMC 96	DIS	Be(9)/C	15	3.8	3.8	1
SLAC E-139	DIS	C(12)/D	7	4.1	3.2	1
NMC 95	DIS	C/D	15	15.0	13.8	1
NMC 95, Q^2 dep.	DIS	C/D	165	141.8	142.0	1
NMC 95, re.	DIS	C/D	16	19.3	20.5	1
NMC 95, re.	DIS	C/Li	20	30.3	28.4	1
FNAL-E772	DY	C/D	9	7.5	8.3	1
SLAC E-139	DIS	Al(27)/D	20	10.9	12.5	1
NMC 96	DIS	Al/C	15	6.0	5.8	1
SLAC E-139	DIS	Ca(40)/D	7	5.0	4.1	1
FNAL-E772	DY	Ca/D	9	2.9	3.4	15
NMC 95, re.	DIS	Ca/D	15	25.4	24.7	1
NMC 95, re.	DIS	Ca/Li	20	23.9	19.6	1
NMC 96	DIS	Ca/C	15	6.0	6.0	1
SLAC E-139	DIS	Fe(56)/D	26	19.1	23.9	1
FNAL-E772	DY	Fe/D	9	2.1	2.2	15
NMC 96	DIS	Fe/C	15	11.0	10.8	1
FNAL-E866	DY	Fe/Be	28	20.9	21.7	1
CERN EMC	DIS	Cu(64)/D	19	13.4	14.8	1
SLAC E-139	DIS	Ag(108)/D	7	3.8	2.9	1
NMC 96	DIS	Sn(117)/C	15	9.6	9.1	1
NMC 96, Q^2 dep.	DIS	Sn/C	144	80.2	82.8	10 (x=0.0125 only)
FNAL-E772	DY	W(184)/D	9	7.0	6.7	10
FNAL-E866	DY	W/Be	28	27.3	24.2	1
SLAC E-139	DIS	Au(197)/D	21	11.6	13.8	1
RHIC-PHENIX	π^0 prod.	dAu/pp	20	7.3	6.3	20
NMC 96	DIS	Pb/C	15	6.90	7.2	1
Total			929	738.6	731.3	

Generalized Chi2 in EPS09

usual definition
$$\chi^2_N(\{z\}) \equiv \sum_{i\in N}\left[\frac{D_i - T_i(\{z\})}{\sigma_i}\right]^2$$

now define
$$\chi^2(\{z\}) \equiv \sum_N w_N\, \chi^2_N(\{z\})$$

$$\chi^2_N(\{z\}) \equiv \left(\frac{1-f_N}{\sigma_N^{\rm norm}}\right)^2 + \sum_{i\in N}\left[\frac{f_N D_i - T_i(\{z\})}{\sigma_i}\right]^2$$

1. **Weights w_N for data sets N**
   which provide important constraints but whose #data points is small

2. **Treatment of the additional overall normalization errors σ^{norm} given by the experiment** (also in e.g. CTEQ):
   - **f_N** is determined by minimizing the chi^2_N for each data set N**,** for each parameter-set candidate during the overall chi^2 minimization
   - **"penalty factor" $(\ldots)^2$** accounts for the fact that f_N=1 should be the optimal value ⇔ the normalization given by the experiment
   - **f_N is an output** of the analysis

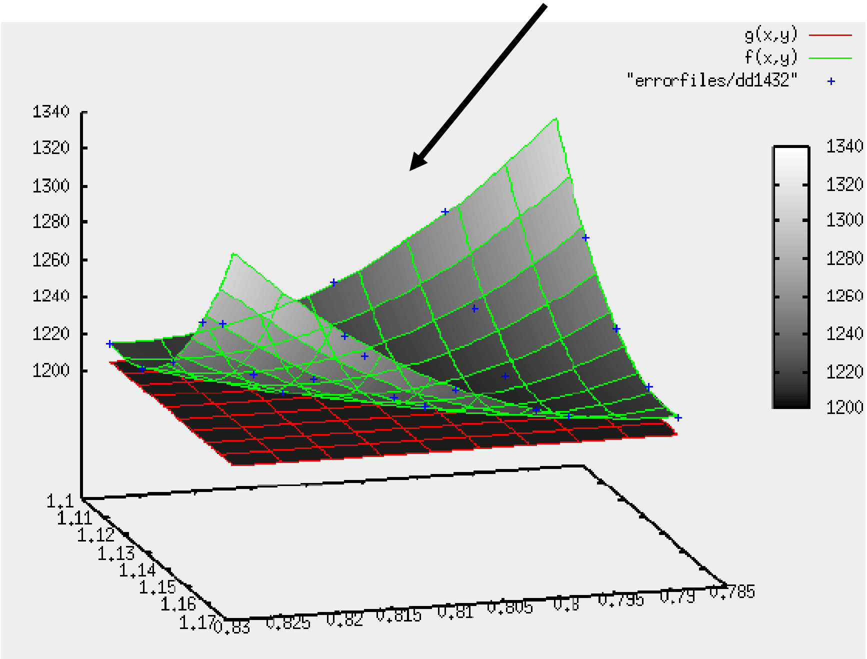
Q: How to estimate the uncertainties in the nPDFs ?
How to propagate these into hard XSs ?

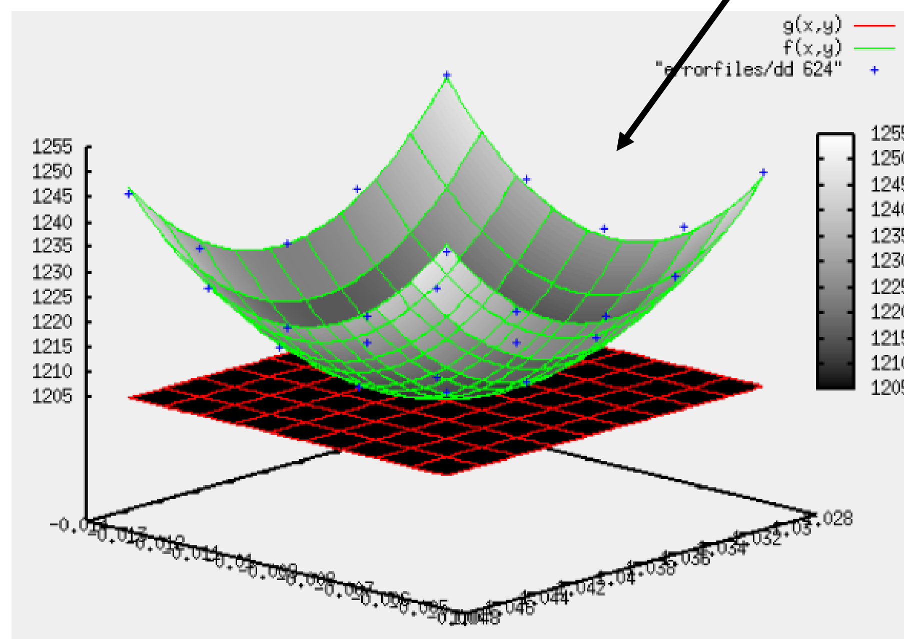
A: Error analysis in EPS09 via Hessian method

• expand around the minimum, this defines the Hessian matrix **H**

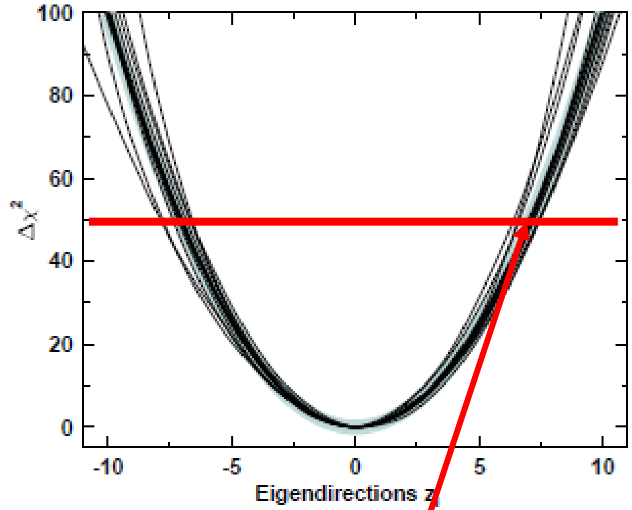
$$\chi^2 - \chi_0^2 \approx \sum_{ij} \frac{1}{2} \frac{\partial^2 \chi^2}{\partial a_i \partial a_j} (a_i - a_i^0)(a_j - a_j^0) \equiv \sum_{ij} H_{ij} (a_i - a_i^0)(a_j - a_j^0)$$

• parameters are correlated, **H** must be diagonalized and correlations removed





O(15) parameters, O(100) pairs



- eigenvectors of **H** provide a basis $\{z_i\}$ of uncorrelated parameter combinations, where

$$\chi^2 \approx \chi_0^2 + \sum_i z_i^2$$

- hence, the hard-cross-section uncertainties can be estimated as

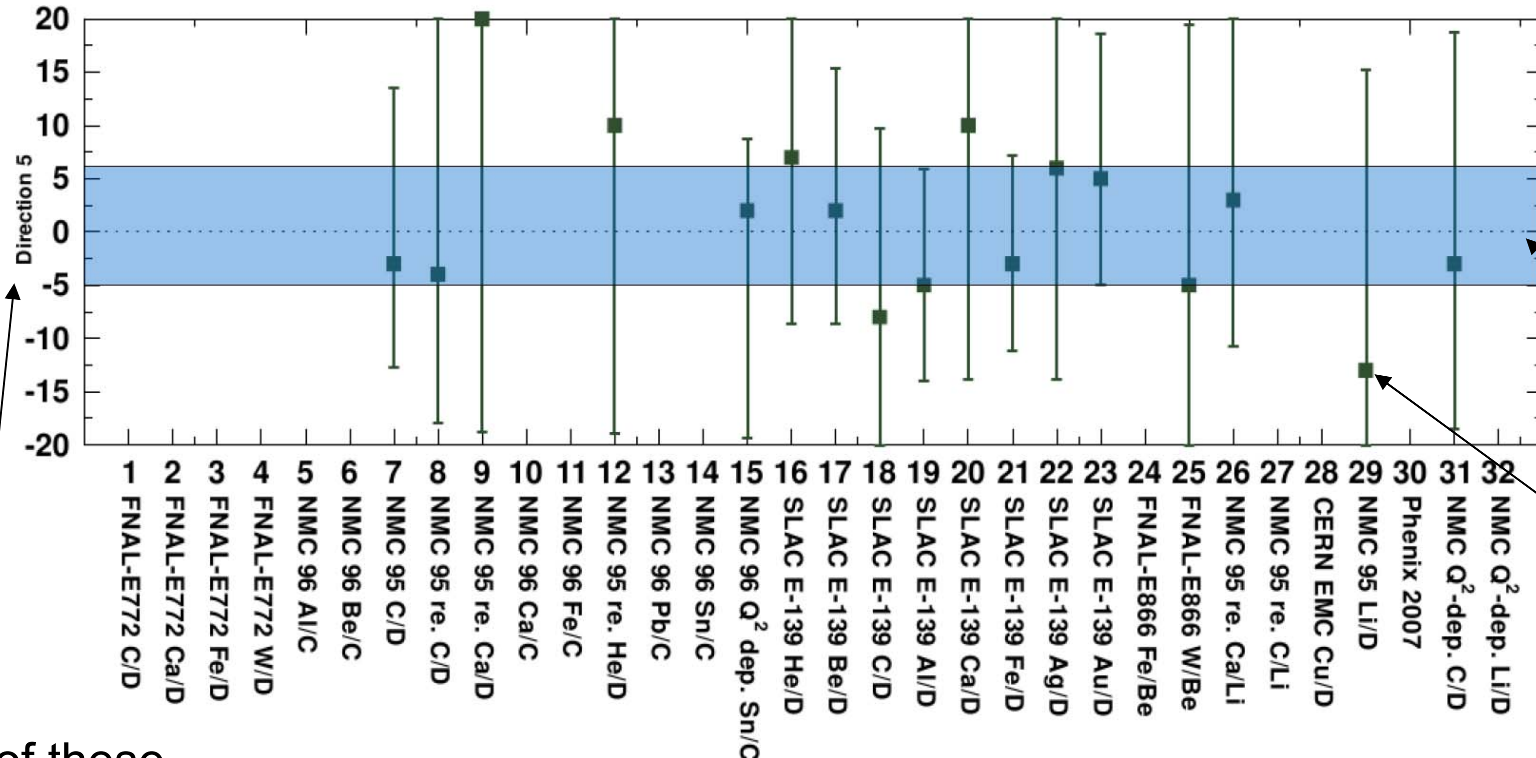
$$(\Delta X)^2 = \left(\frac{\partial X}{\partial z_1} \cdot \delta z_1\right)^2 + \left(\frac{\partial X}{\partial z_2} \cdot \delta z_2\right)^2 + \cdots$$

- but in the uncertainty estimates, how large should δz_k i.e. $\Delta\chi^2$ be, in order to correctly reflect the uncertainties in the data ?

→ No unique procedure exists for this

→ use a **"90 % confidence criterion"** similar to CTEQ

- **We take $\Delta\chi^2 = 50$ based on requiring the χ^2-contribution of each data set to stay within its 90% confidence range.**



total chi2 min

Min chi2 for data set 29

15 of these

- $$\Delta\chi^2 \equiv \sum_i \frac{\Delta\chi^2(z_i^+) + \Delta\chi^2(z_i^-)}{2N}$$ **= 50**

[from HP]

The EPS09 package

= the central set, **best fit S^0** + **30 nPDF error sets $S_k\pm$** where each $S_k\pm$ is obtained by displacing the fit parameters to the +/- direction along the eigendirections z_k such that chi2 grows by 50

the user can then

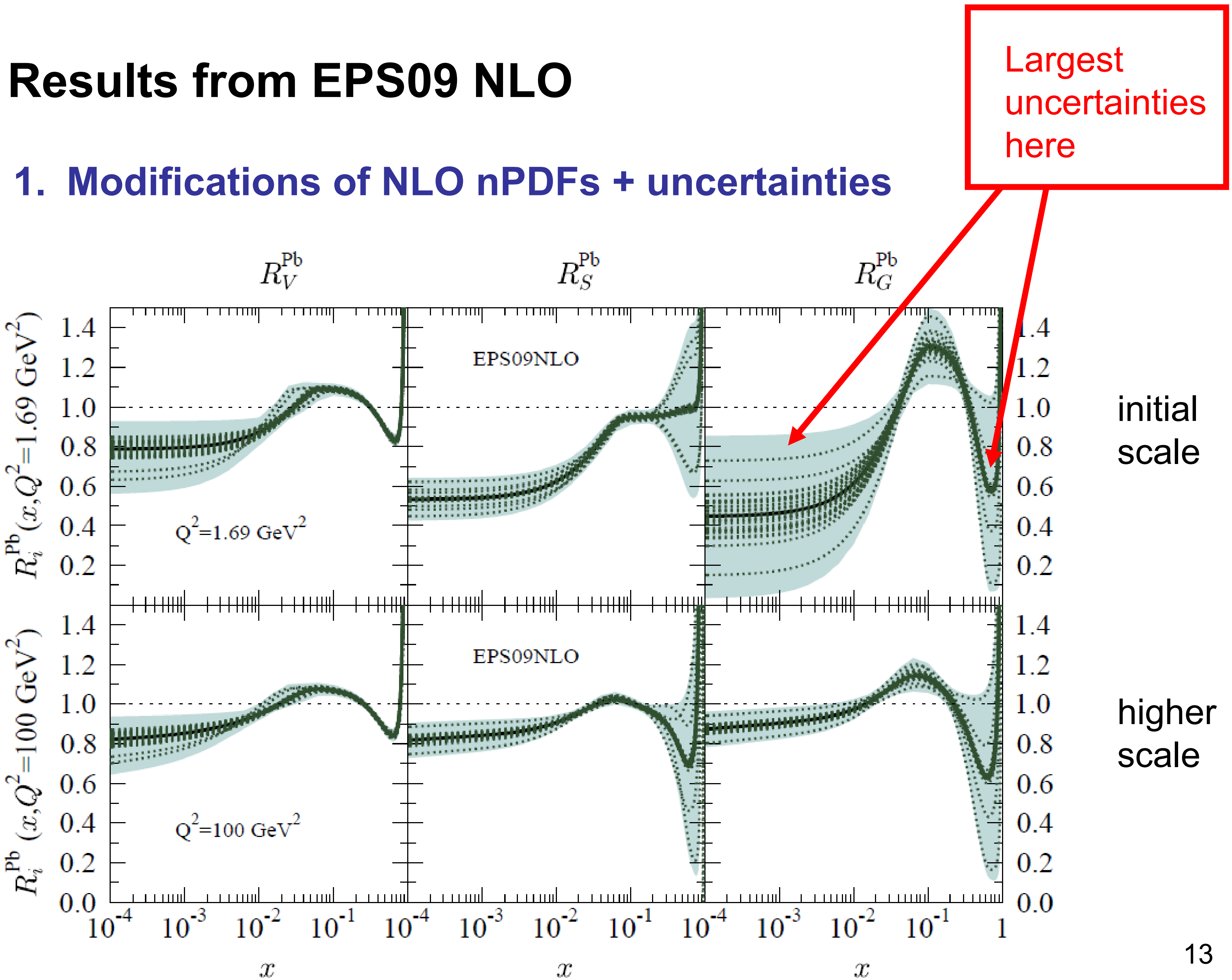
- compute the propagation of the nPDF-originating uncertainties into the cross section X of his/her interest as follows:

$$(\Delta X^+)^2 \approx \sum_k \left[\max\left\{X(S_k^+) - X(S^0), X(S_k^-) - X(S^0), 0\right\}\right]^2$$

$$(\Delta X^-)^2 \approx \sum_k \left[\max\left\{X(S^0) - X(S_k^+), X(S^0) - X(S_k^-), 0\right\}\right]^2$$

Results from EPS09 NLO

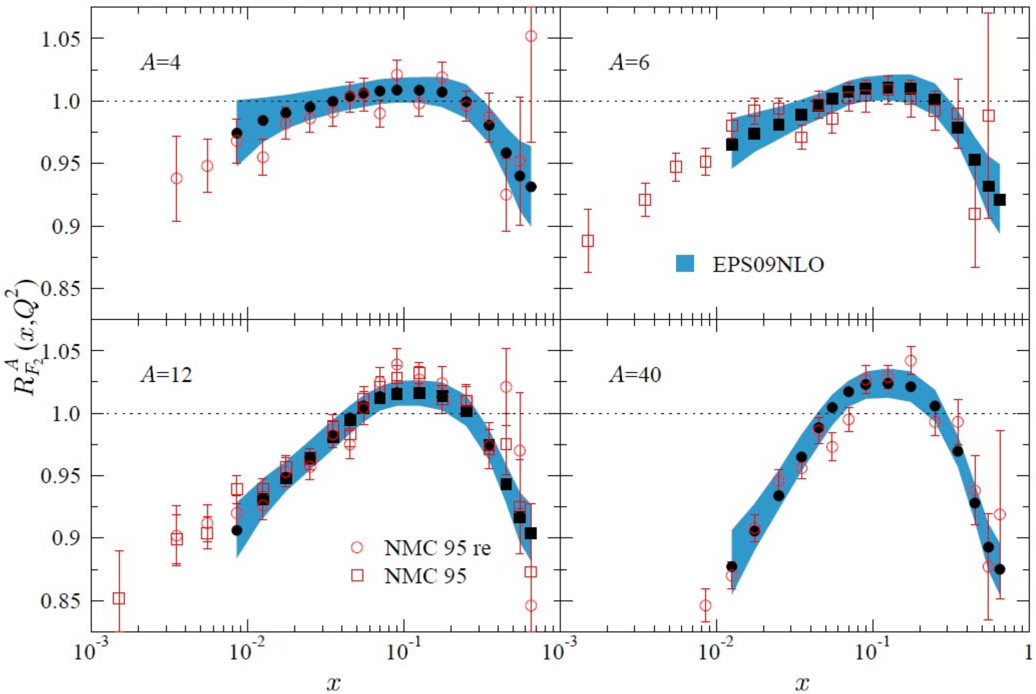
1. Modifications of NLO nPDFs + uncertainties

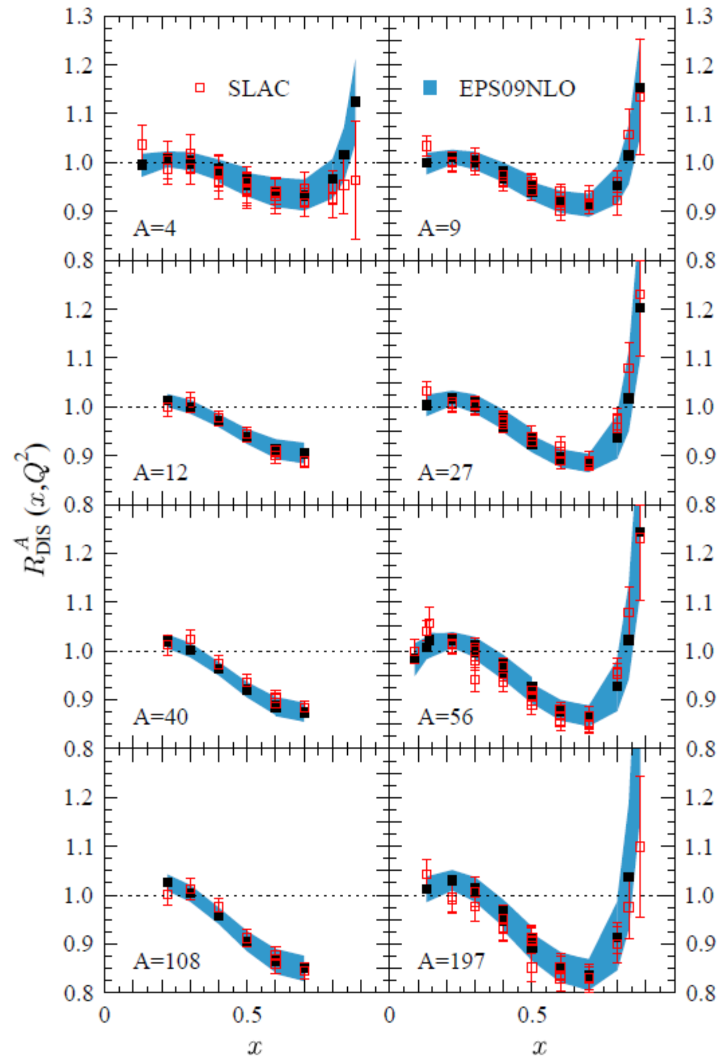


2. Comparison with data: NLO DIS + uncertainties

$$\sigma_{\rm DIS}^{\ell+A\to\ell+X} = \sum_{i=q,\overline{q},g} f_i^A(\mu^2) \otimes \hat{\sigma}_{\rm DIS}^{\ell+i\to\ell+X}(\mu^2)$$

$$R_{\rm DIS}^{\rm A}(x,Q^2) \equiv \frac{\frac{1}{A}d\sigma_{\rm DIS}^{l\rm A}/dQ^2dx}{\frac{1}{2}d\sigma_{\rm DIS}^{l\rm d}/dQ^2dx}$$



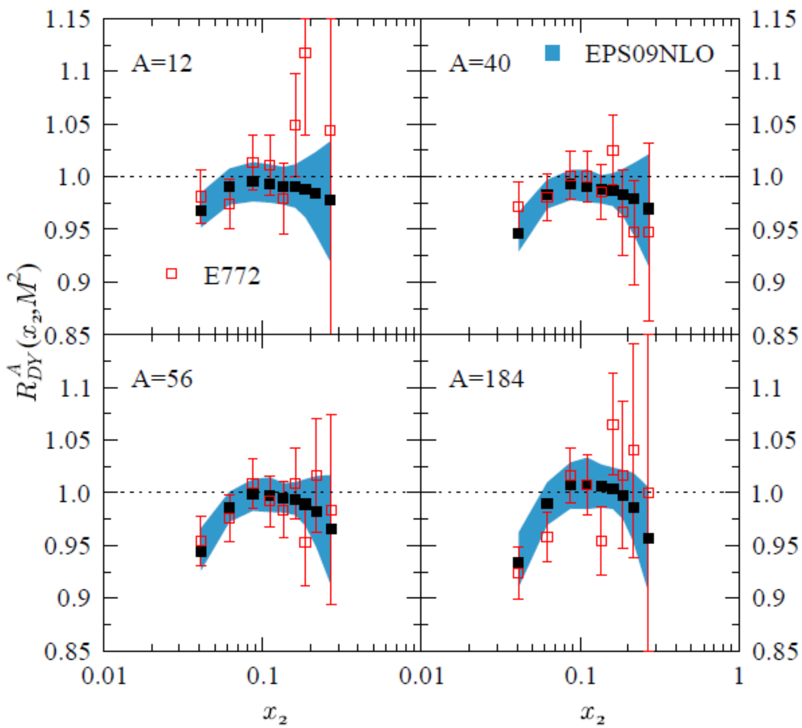


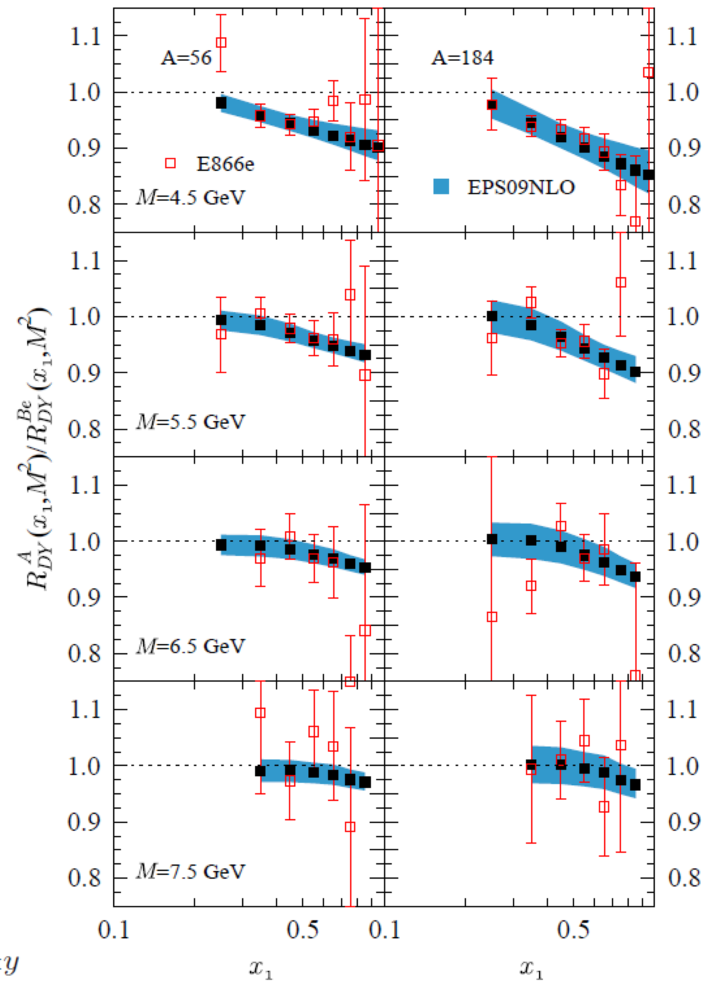
Good fits & Propagation of the exp. uncertainties into the nPDFs is OK !

3. Comparison with data: NLO DY + uncertainties

$$\sigma_{\rm DY}^{p+A\to l^+l^-+X} = \sum_{i,j=q,\overline{q},g} f_i^p(\mu^2)\otimes f_j^A(\mu^2)\otimes \hat{\sigma}^{ij\to l^+l^-+X}(\mu^2)$$

$$R_{\rm DY}^{\rm A}(x_{1,2},M^2) \equiv \frac{\frac{1}{A}d\sigma_{\rm DY}^{\rm pA}/dM^2dx_{1,2}}{\frac{1}{2}d\sigma_{\rm DY}^{\rm pd}/dM^2dx_{1,2}}$$





$$x_{1,2} \equiv \sqrt{M^2/s}\, e^{\pm y}$$

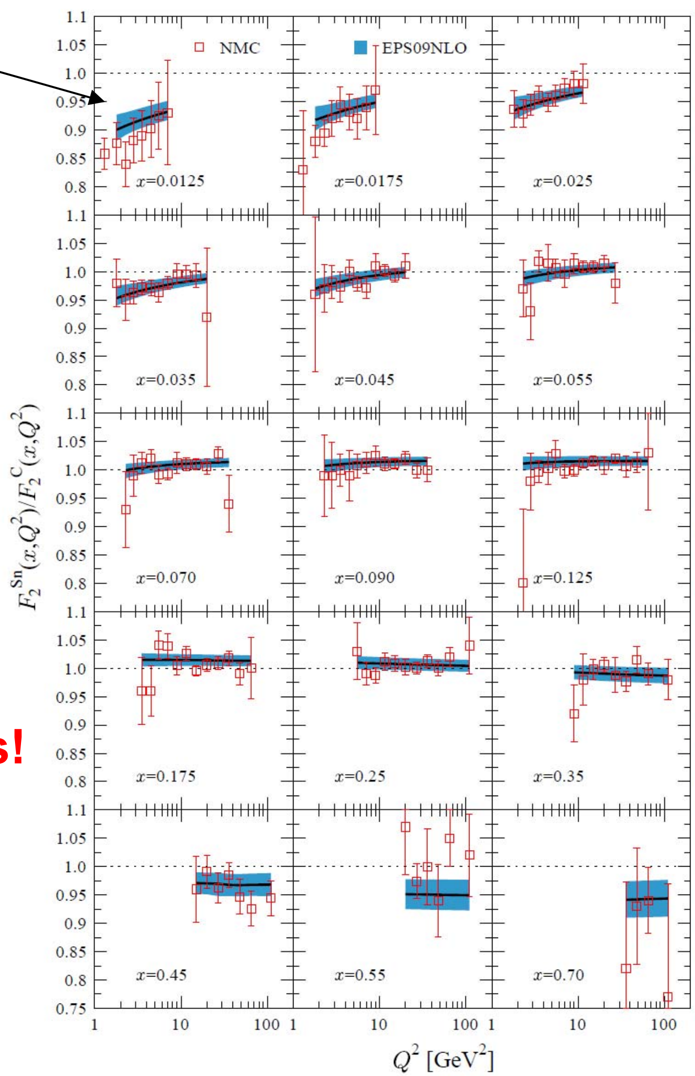
4. Comparison with data: DIS vs. NLO DGLAP evolution

Due to this panel, gluons at x~0.03 are ~well constrained

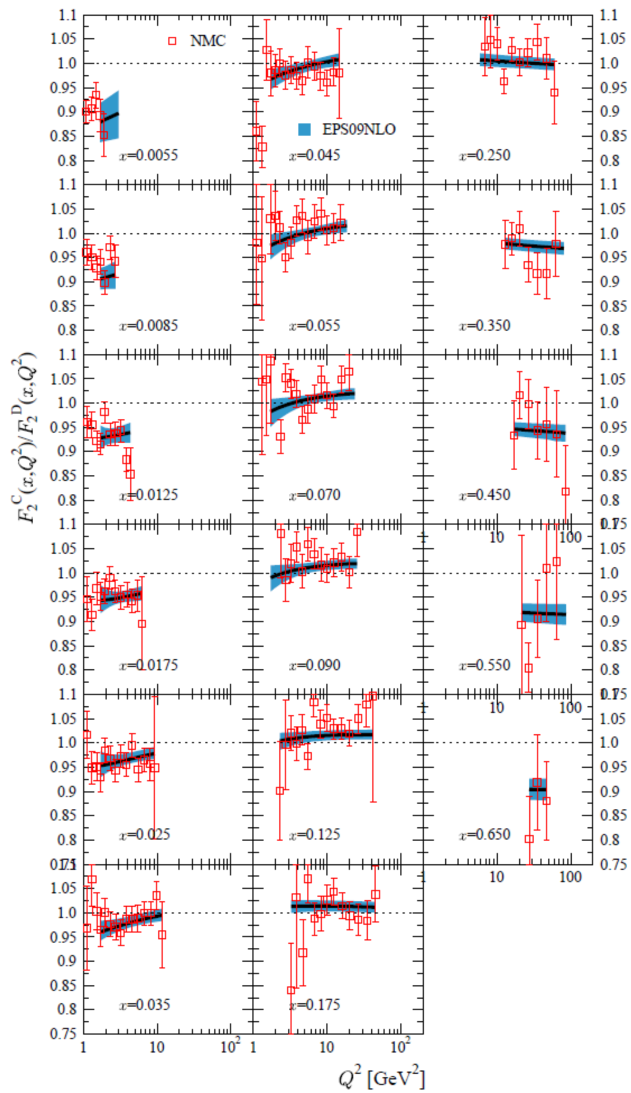
but…

smaller-x data would be needed to reduce the large uncertainties in small-x gluons!

F_2^{Sn}/F_2^{C}



F_2^{C}/F_2^{D}

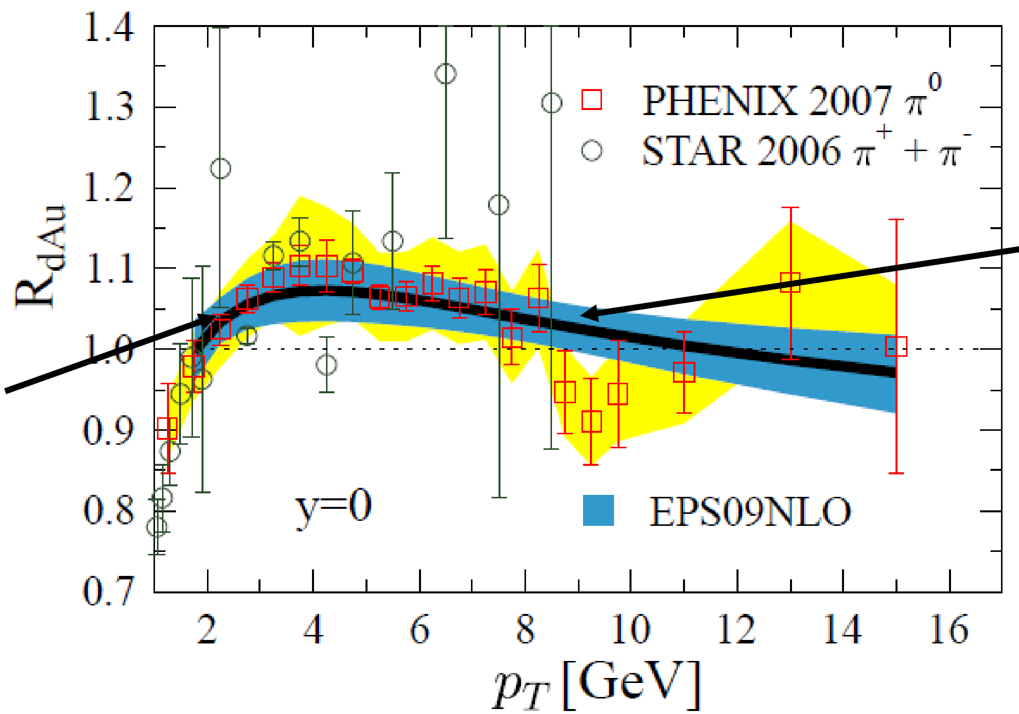


5. Comparison with data: pi0 (y=0, pT>1.7 GeV) in d+Au at RHIC

$$\sigma^{A+B\to\pi+X} = \sum_{i,j,k=q,\overline{q},g} f_i^A(\mu^2)\otimes f_j^B(\mu^2)\otimes\hat{\sigma}^{ij\to k+X}(\mu^2)\otimes D_{k\to\pi}(\mu^2)$$

$$R^{\pi}_{\mathrm{dAu}} \equiv \frac{1}{\langle N_{\mathrm{coll}}\rangle}\frac{d^2N^{\mathrm{dAu}}_{\pi}/dp_T dy}{d^2N^{\mathrm{pp}}_{\pi}/dp_T dy} \overset{\mathrm{min.bias}}{=} \frac{\frac{1}{2A}d^2\sigma^{\mathrm{dAu}}_{\pi}/dp_T dy}{d^2\sigma^{\mathrm{pp}}_{\pi}/dp_T dy}$$

Surprisingly good, fairly tensionless fit



From shadowing

Suggests an EMC-depletion for gluons – the treatment of overall normalization errors important!

EPS09 – summary

• Excellent agreement between NLO pQCD and the hard-process nuclear data for DIS, DY, and π^0 production in the kinematical range $0.005 < x < 1$, $1.69\ \mathrm{GeV}^2 < Q^2 < 150\ \mathrm{GeV}^2$ --- chi2/N = 0.79

• No significant tension between the data sets from different processes

→ Factorization seems to work well in describing high-E inclusive nuclear hard processes

• The **EPS09** nPDF package

central set (best fit) + 30 error sets both in NLO and LO,
https://www.jyu.fi/fysiikka/en/research/highenergy/urhic/nPDFs

→ Estimation of the nPDF-uncertainty propagation into hard nuclear XSs is now, finally, possible for anybody!

- **further tests of factorization and nPDFs provided by**
  RHIC data for
    direct photons in d+Au (also Au+Au)
    heavy quarks in d+Au
    pion production at fwd-y

  p+A at the LHC – but are such runs foreseen?

  future facilities eRHIC, **LHeC?**

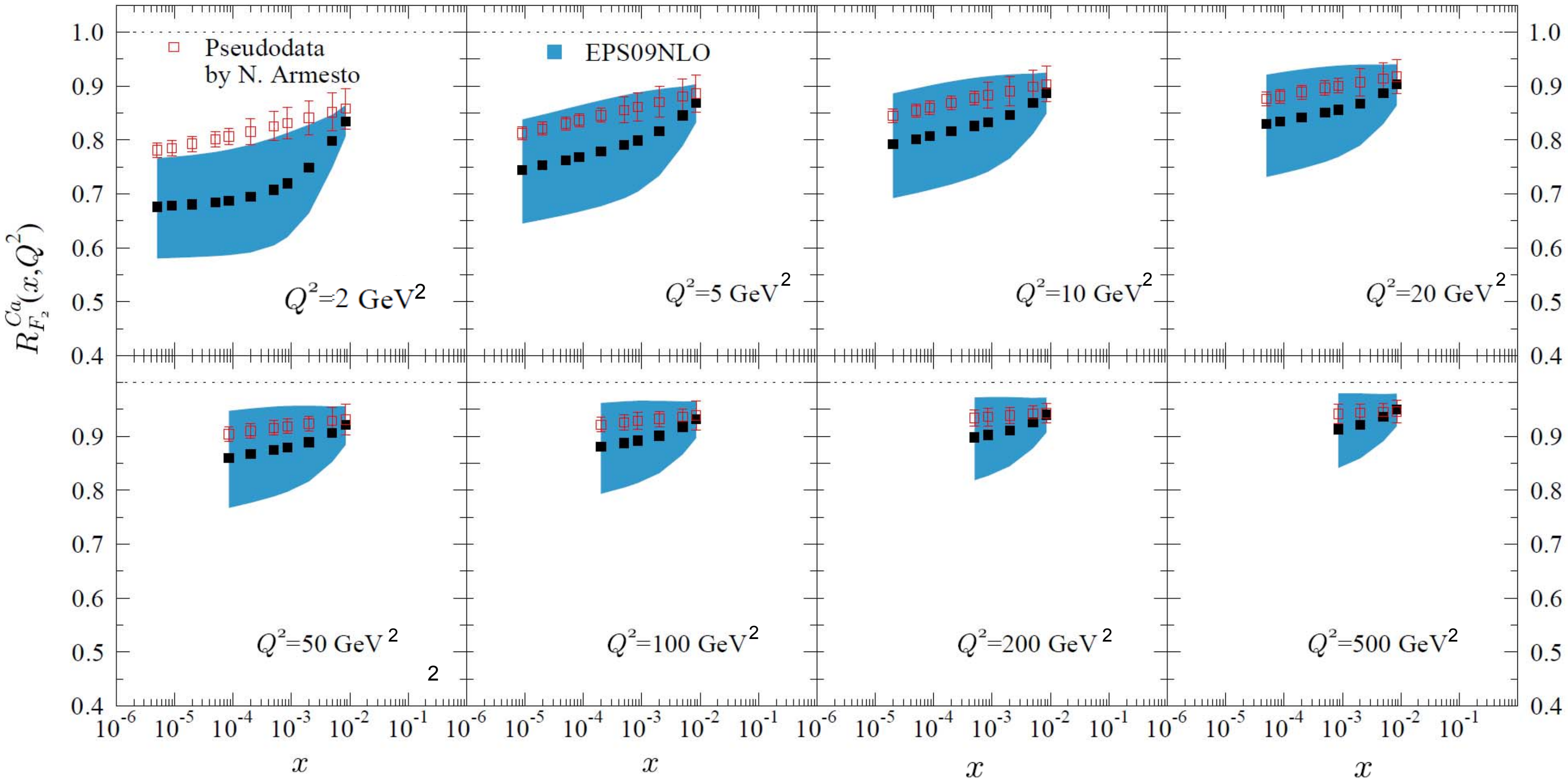
- **on our near-future work list**
  + include also neutrino DIS data (NuTeV, …)
    → further constraints for quarks?
  + extend the analysis to general-mass framework

- **a longer-term goal**
  a master global analysis which combines PDFs and nPDFs

Impact of the LHeC for the global analysis of nPDFs?

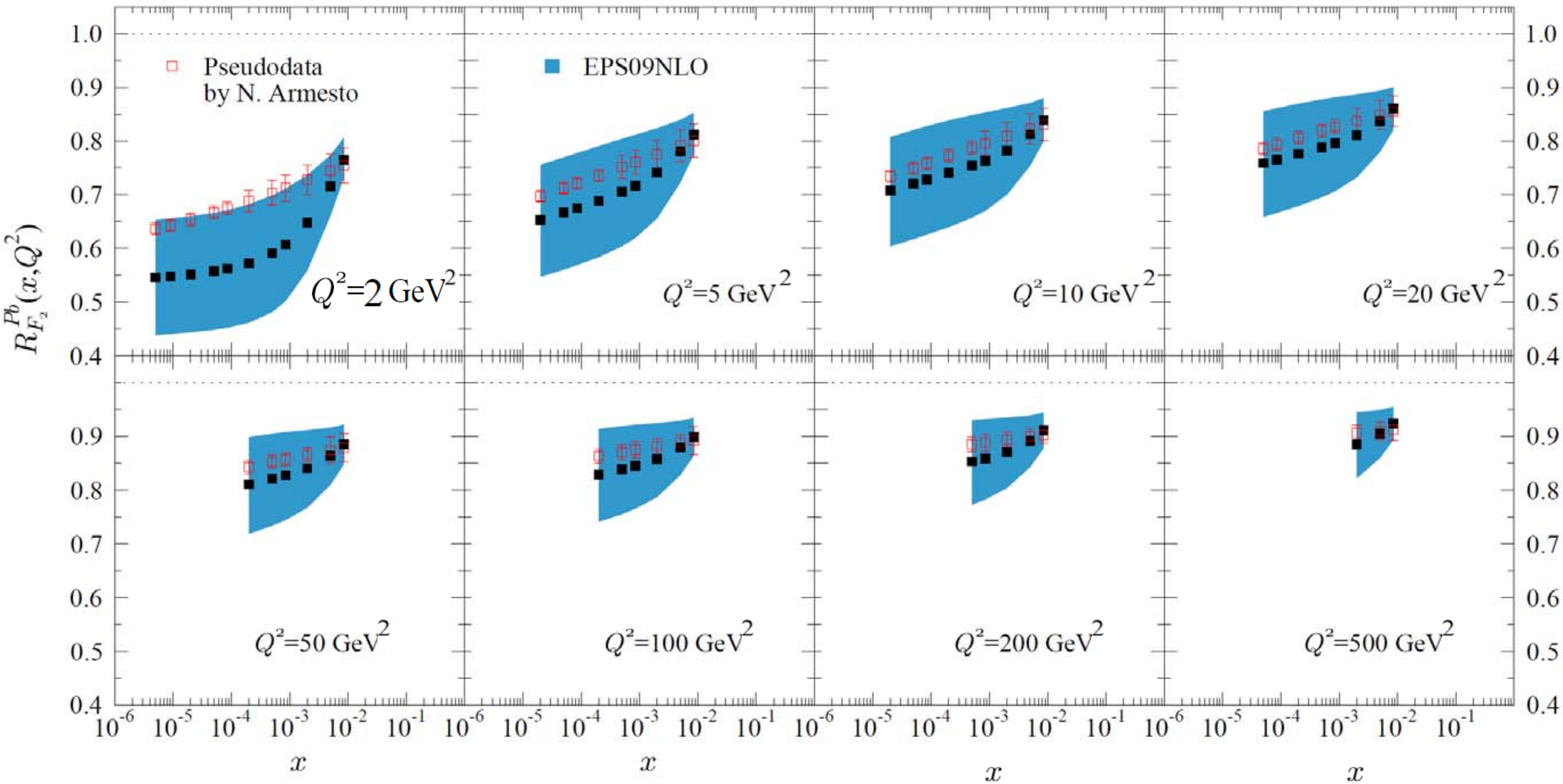
1. Plot the LHeC pseudodata [generated by N. Armesto] against EPS09

Calsium, A=40



[figure from HP]

Lead, A=208



Inclusion of such LHeC data in the global analysis should reduce the nPDF uncertainties at small-x significantly

[figure from HP]

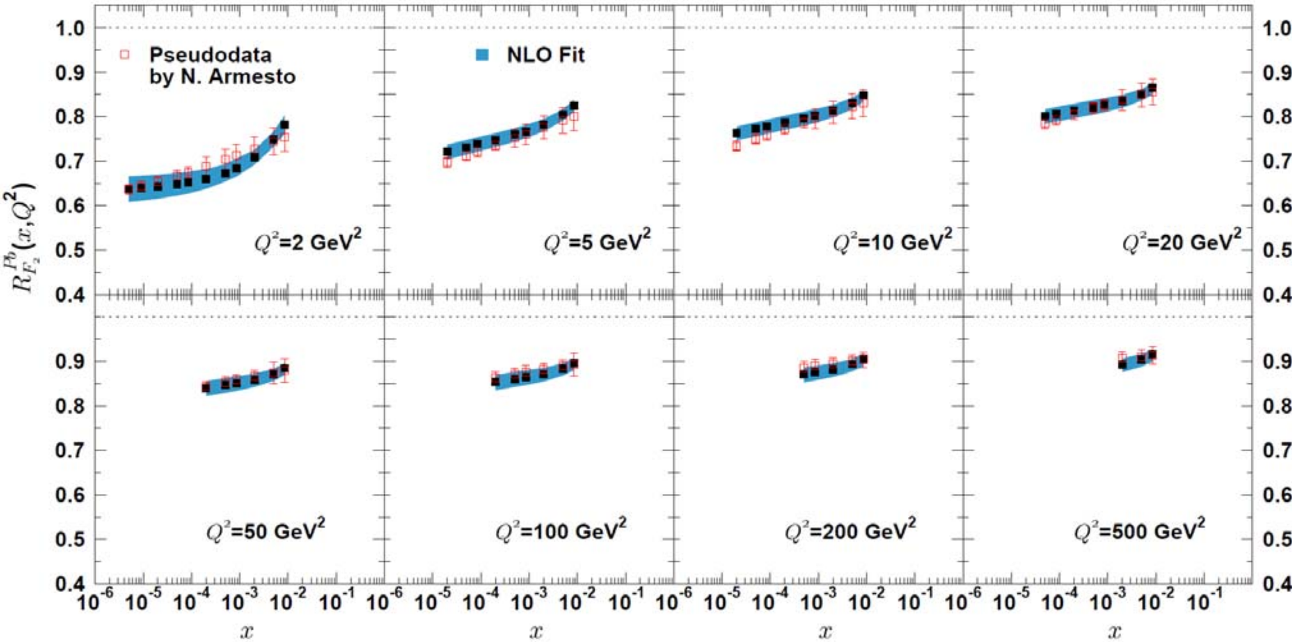
2. Global NLO fit with LHeC pseudodata [from N. Armesto] included

[performed by Hannu Paukkunen for this workshop]

- keep the EPS09 set-up (fit functions, khi2 definition, weighting, error analysis, etc)
- keep the same $\Delta\chi^2$ = 50 in the error analysis
- decrease the PHENIX data weight (now more data on gluons, can now do this)
- with the LHeC data, we can release one more gluon parameter (x_a)

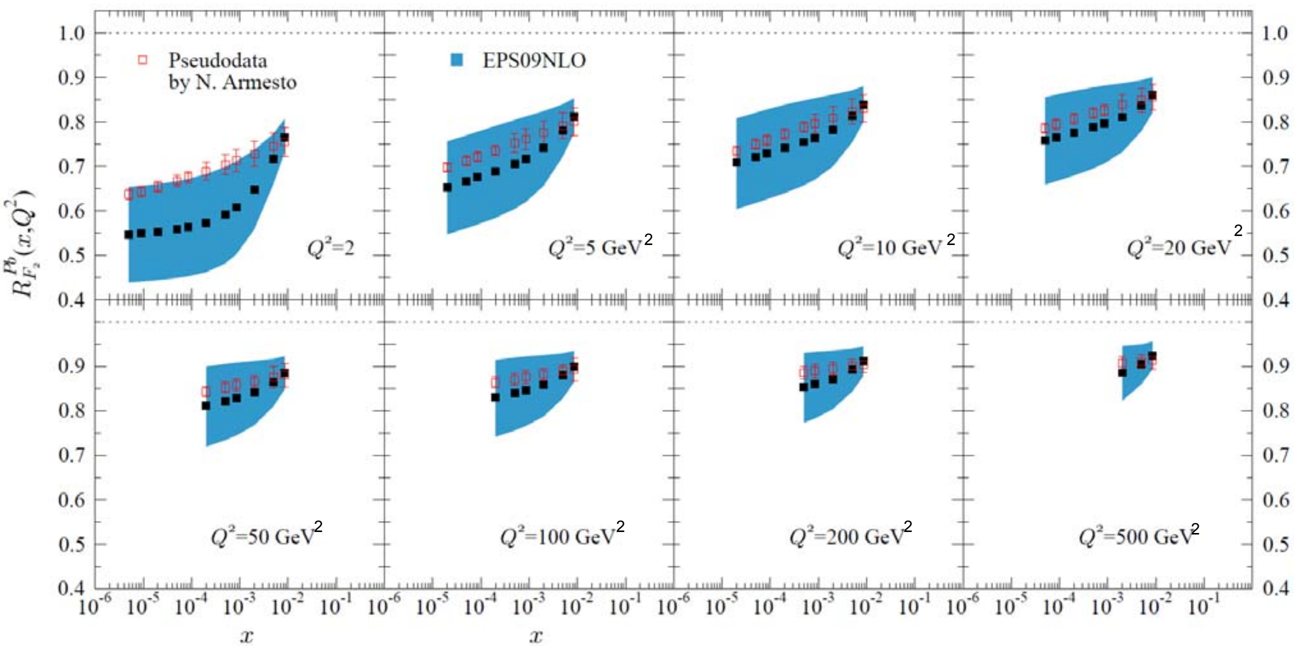
→ Quantify how much the nPDF uncertainties can be expected to reduce

A=208



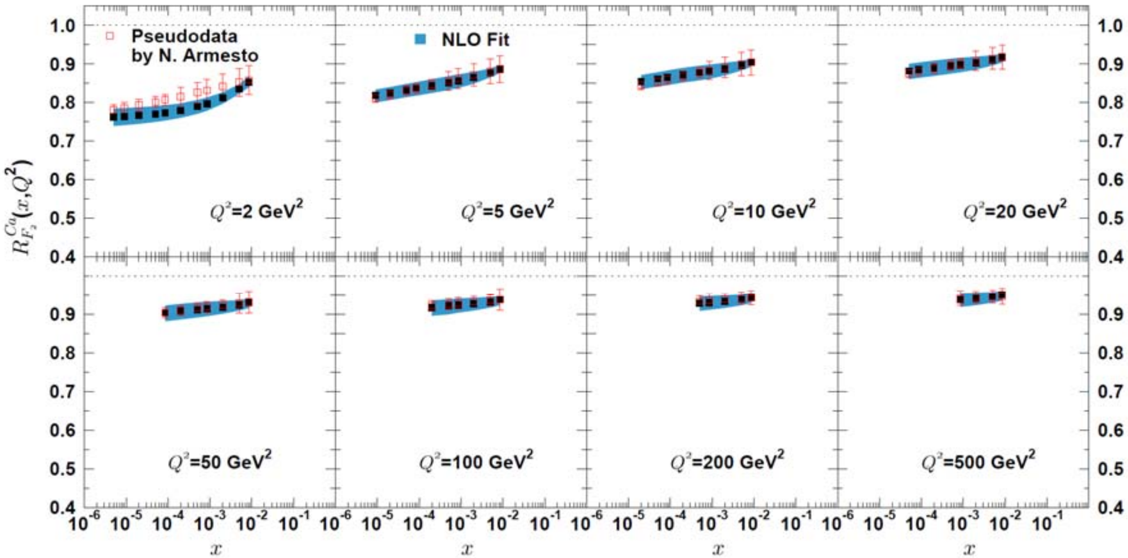
LHeC pseudodata included in the fit; uncertainties much smaller

$\Delta\chi^2$ = 50 seems OK here, too

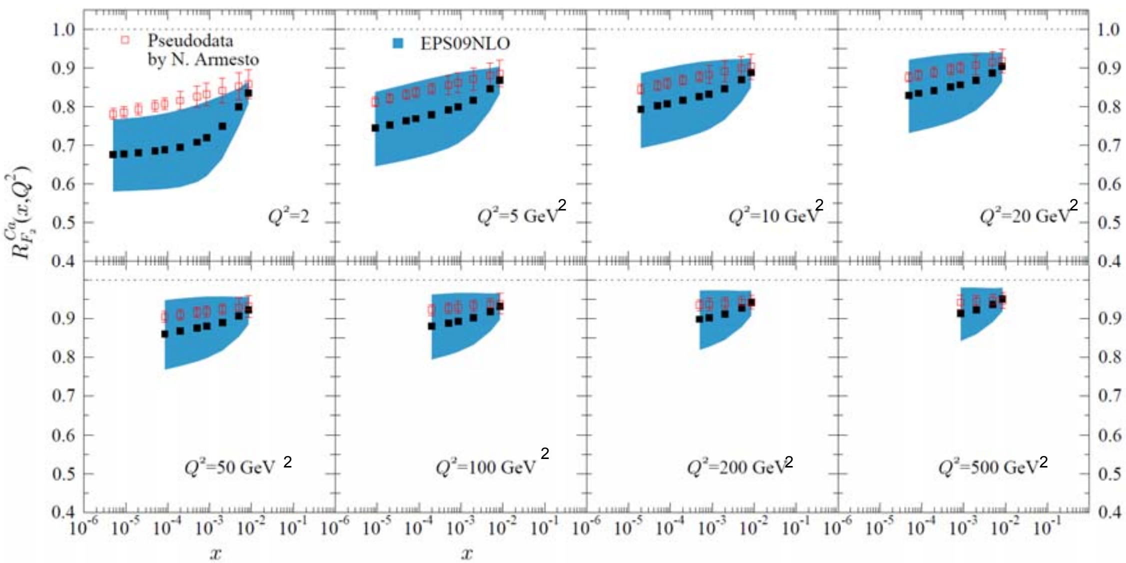


EPS09 – LHeC pseudodata not included

A=40



LHeC pseudodata included in the fit; uncertainties much smaller

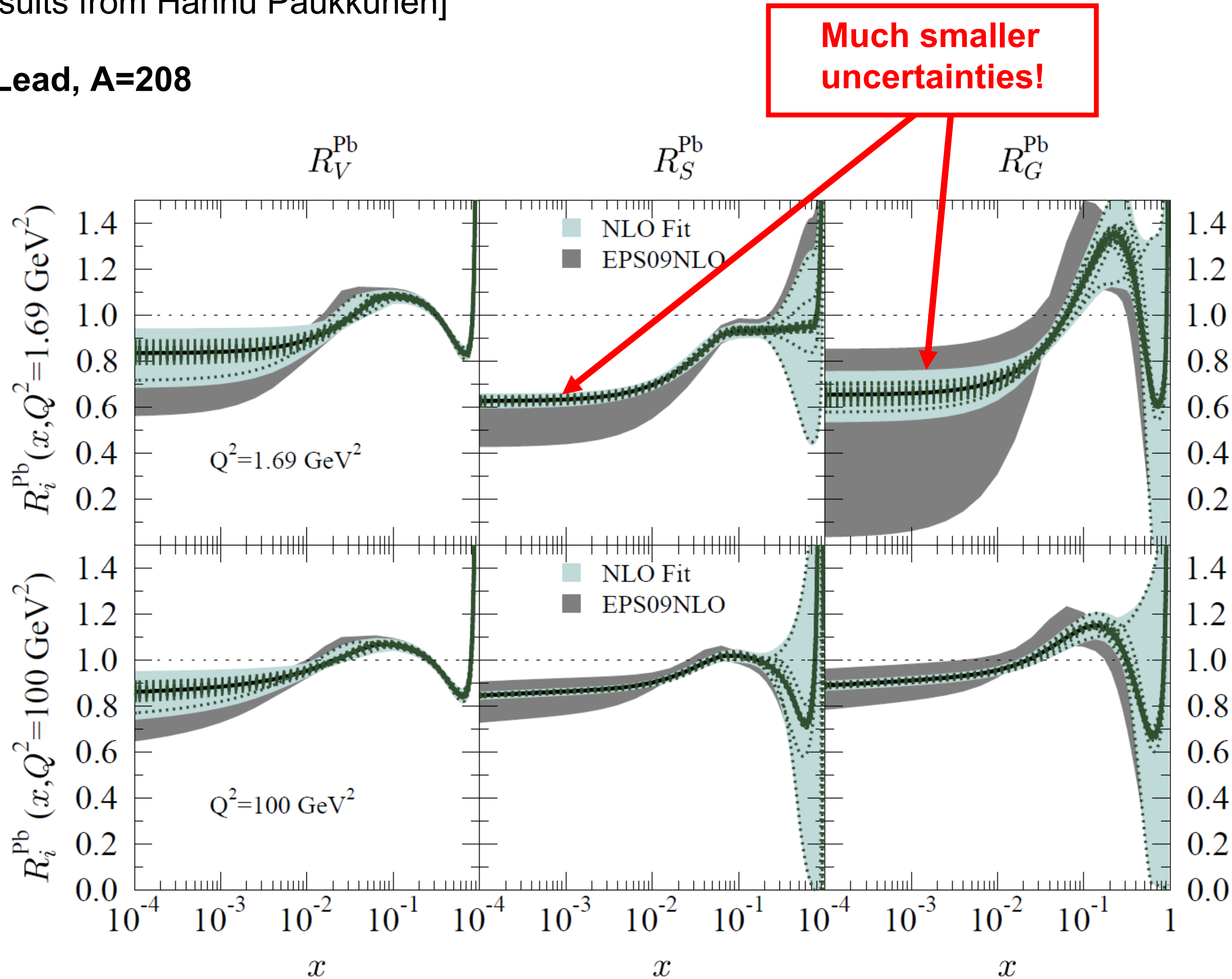


EPS09;
LHeC pseudodata
not included

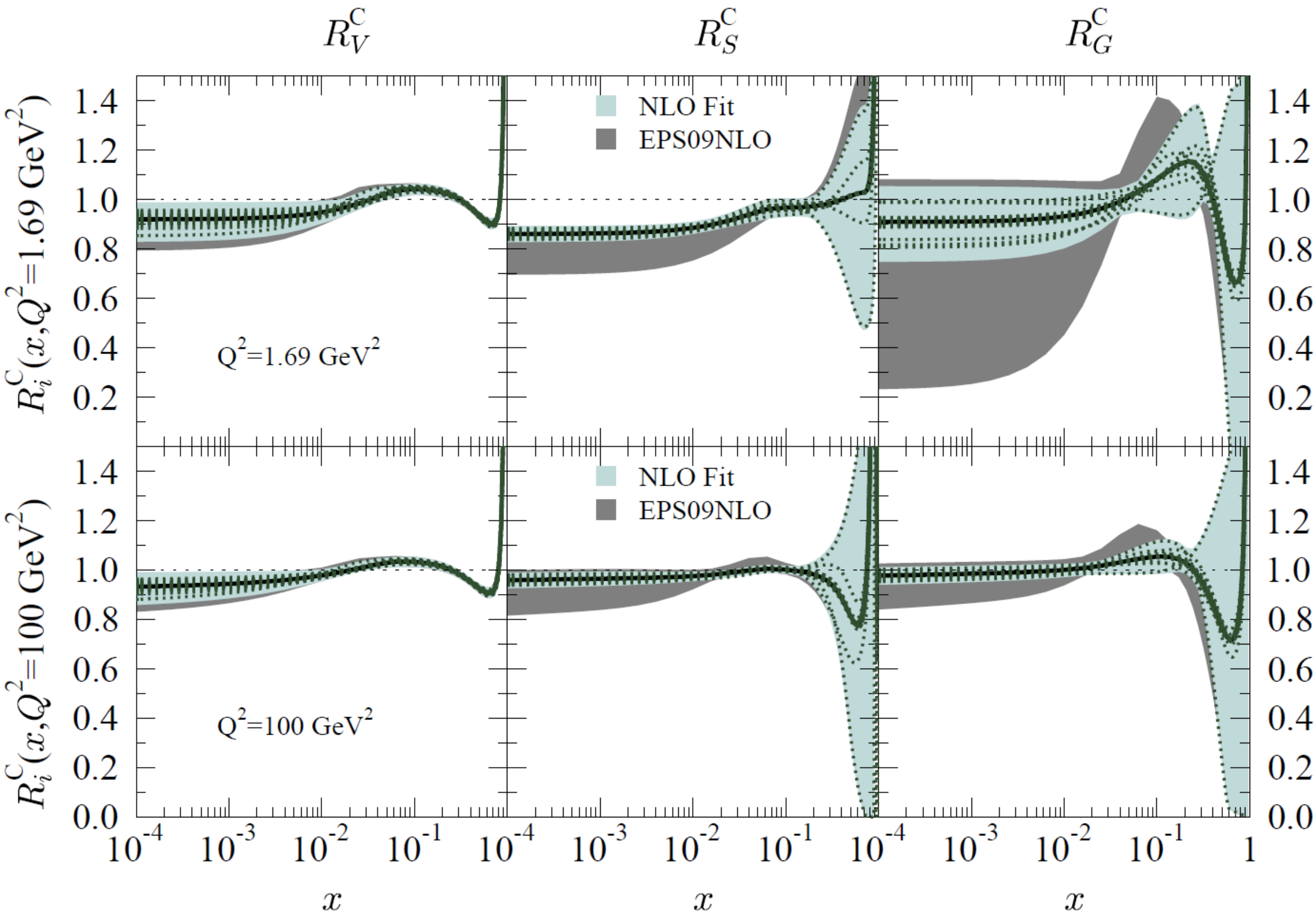
Global NLO fit with LHeC pseudodata [from N. Armesto] included

[results from Hannu Paukkunen]

Lead, A=208

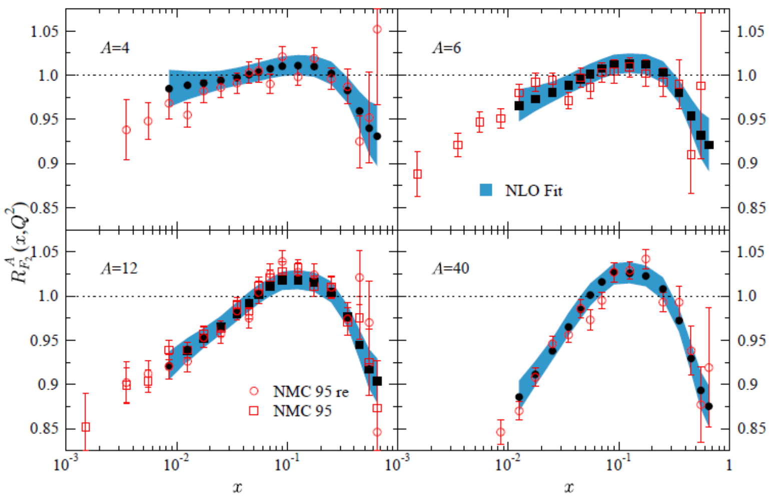


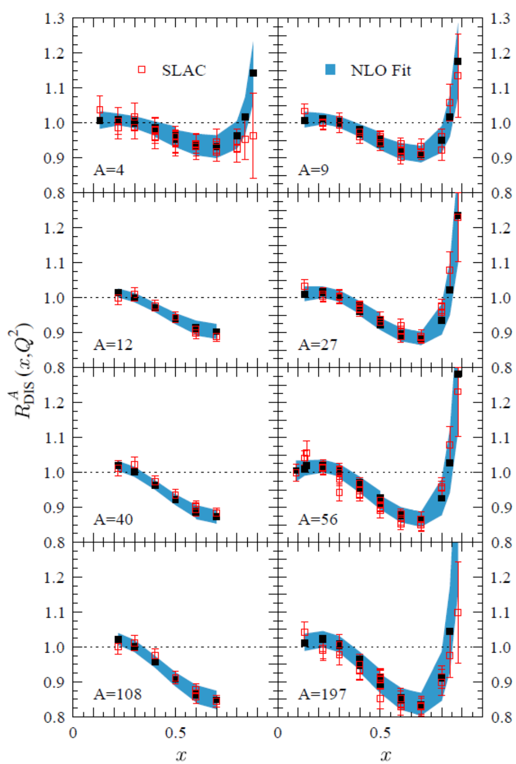
Carbon, A=12



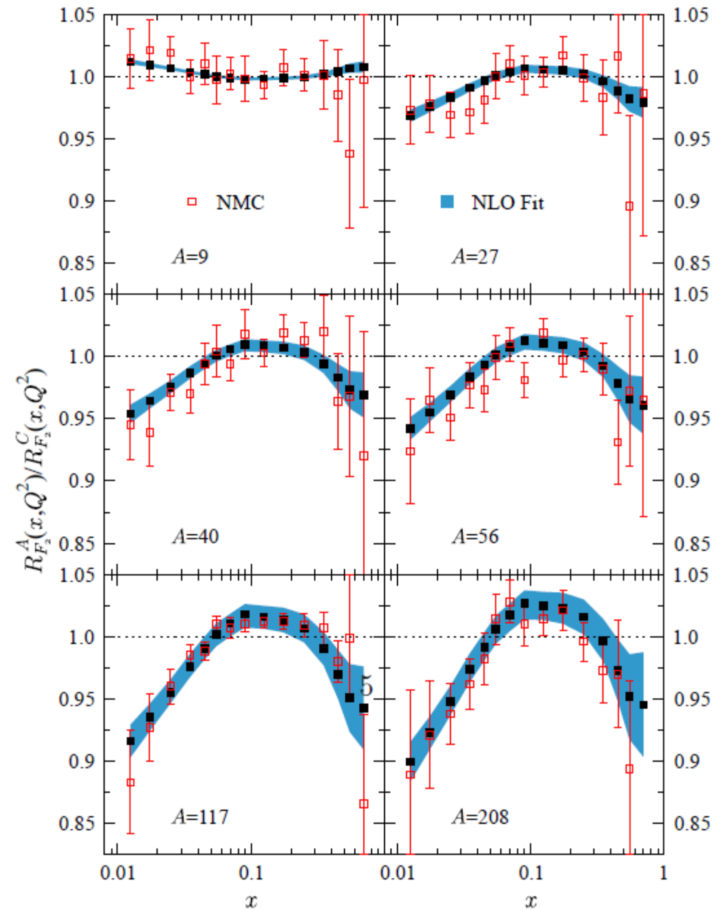
R_V^C
R_S^C
R_G^C
NLO Fit
EPS09NLO
NLO Fit
EPS09NLO
$R_i^C(x,Q^2=1.69\ \mathrm{GeV}^2)$
$R_i^C(x,Q^2=100\ \mathrm{GeV}^2)$
Q^2=1.69 GeV^2
Q^2=100 GeV^2
x
x
x

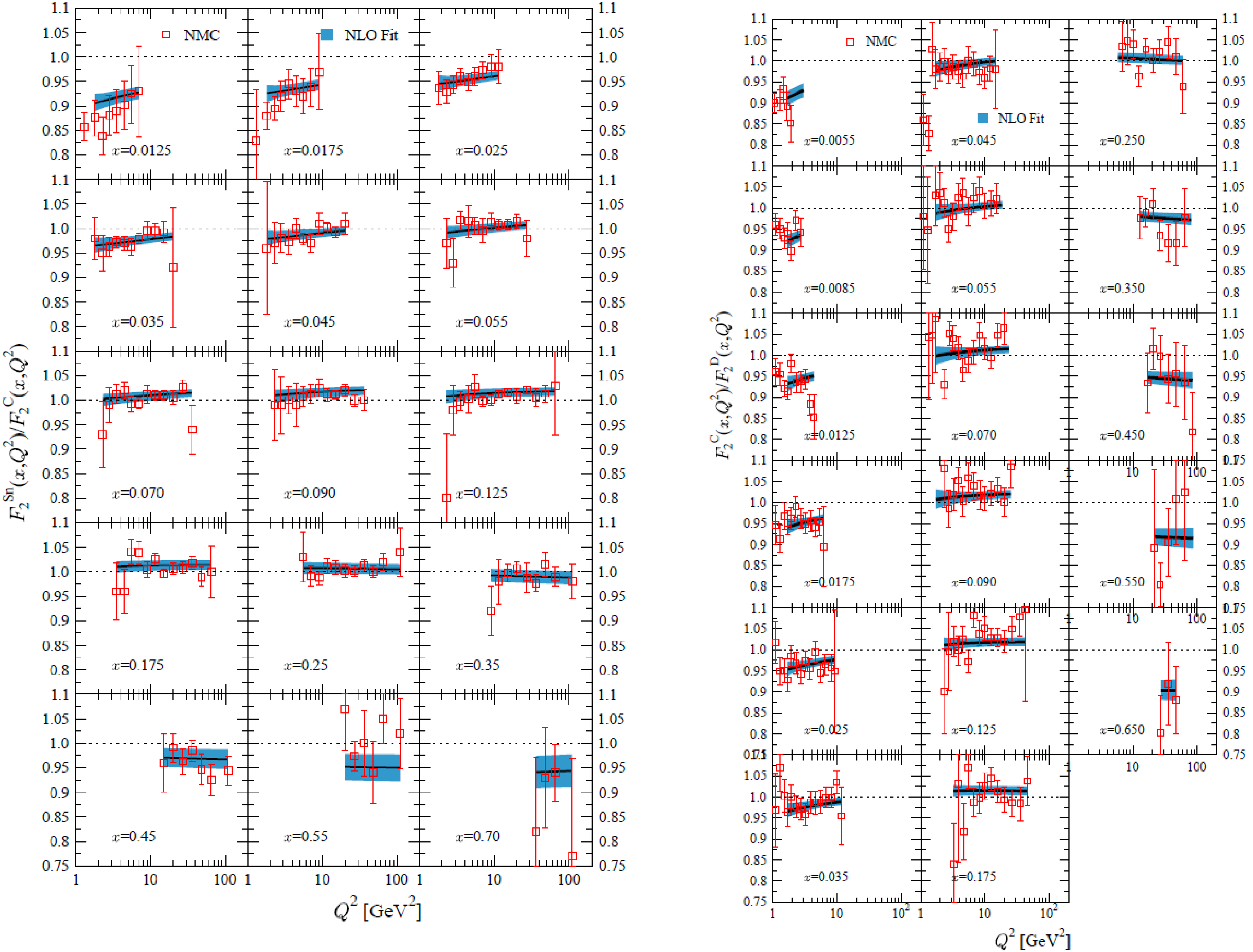
Also with the LHeC pseudodata included, the global NLO fit remains very good…

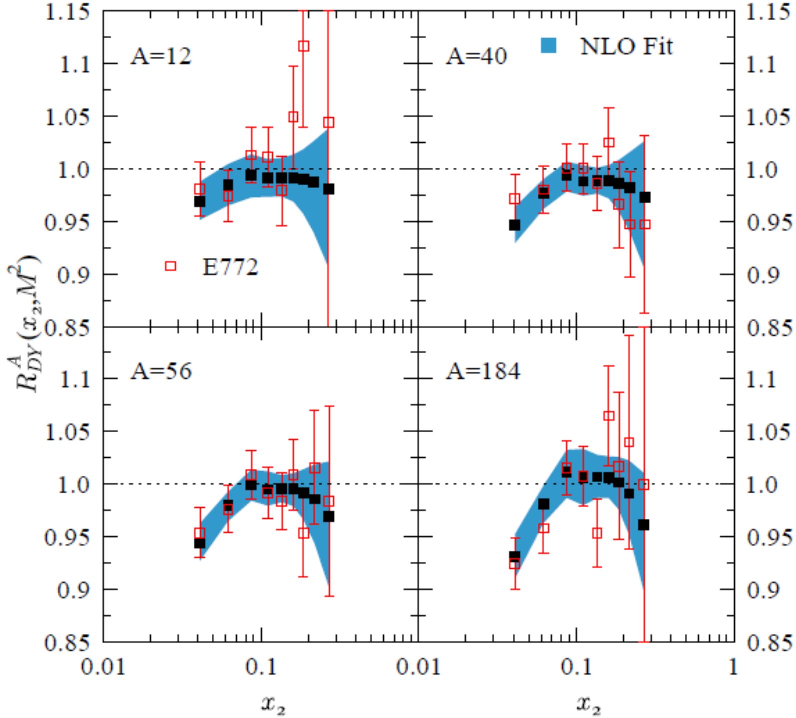




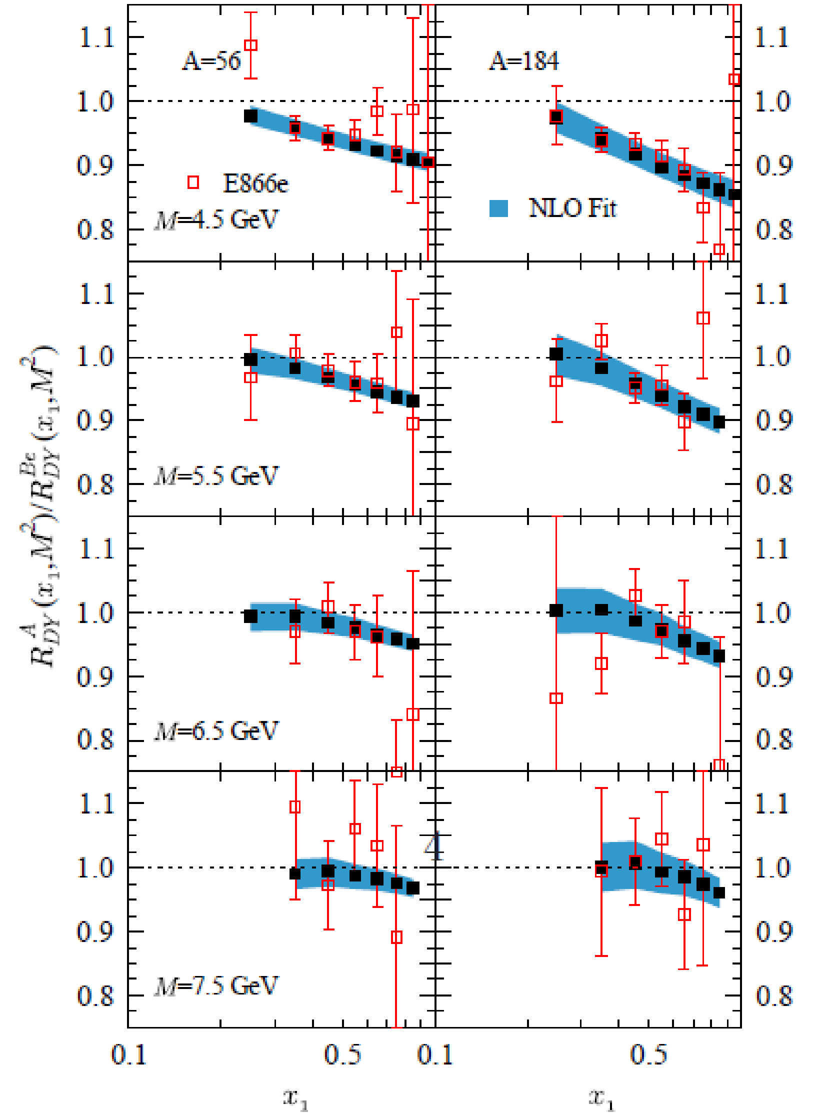
lepton+A DIS



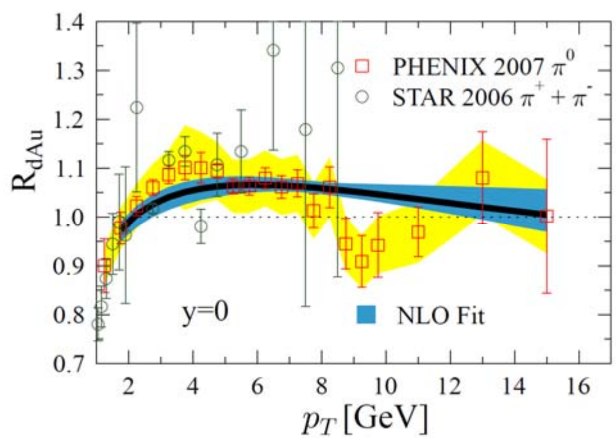

NMC
NLO Fit
$F_2^{Sn}(x,Q^2)/F_2^{C}(x,Q^2)$
x=0.0125
x=0.0175
x=0.025
x=0.035
x=0.045
x=0.055
x=0.070
x=0.090
x=0.125
x=0.175
x=0.25
x=0.35
x=0.45
x=0.55
x=0.70
Q^2 [GeV^2]
$F_2^{C}(x,Q^2)/F_2^{D}(x,Q^2)$
x=0.0055
x=0.0085
x=0.0125
x=0.0175
x=0.025
x=0.035
x=0.045
x=0.055
x=0.070
x=0.090
x=0.125
x=0.175
x=0.250
x=0.350
x=0.450
x=0.550
x=0.650
Q^2 [GeV^2]



DY in p+A



pi0 in d+A at RHIC



Conclusion

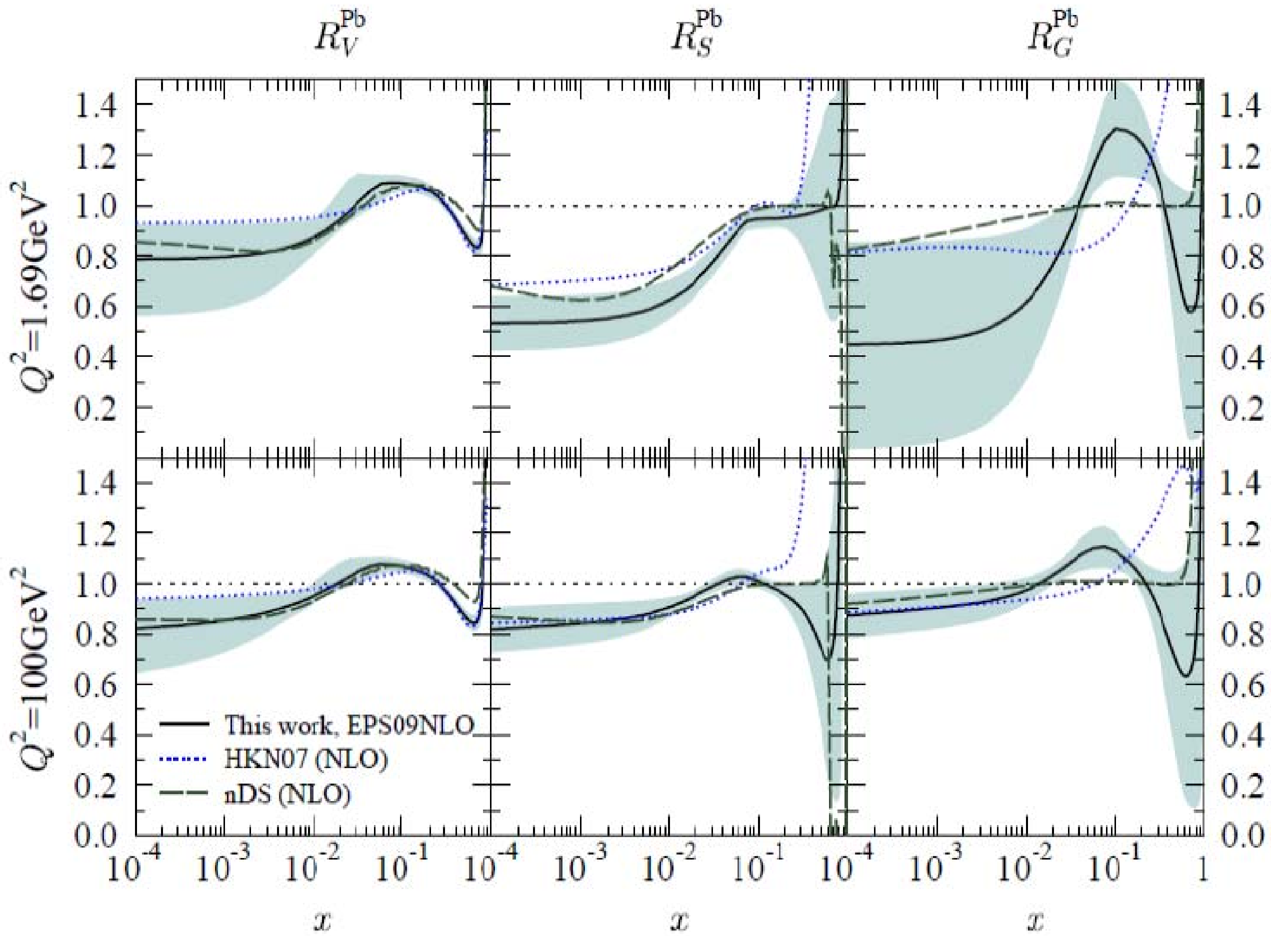
Performing a global analysis of nPDFs with a set of simulated LHeC pseudodata, we have demonstrated the effect of the expected LHeC data on the nPDF uncertainties.

The high-precision LHeC data would play a crucial role in pinning down the nPDFs in the small-x region at perturbative scales.

Extra slides

NLO	Param.	Valence R_V^A	Sea R_S^A	Gluon R_G^A
1	y_0	Baryon sum rule	**0.785**	Momentum sum rule
2	p_{y_0}	—	**-0.136**	—
3	x_a	**6.56 $\times 10^{-2}$**	**8.74 $\times 10^{-2}$**	0.1 fixed
4	p_{x_a}	0, fixed	0, fixed	0, fixed
5	x_e	**0.688**	as valence	as valence
6	p_{x_e}	0, fixed	0, fixed	0, fixed
7	y_a	**1.05**	**0.970**	**1.207**
8	p_{y_a}	**1.47 $\times 10^{-2}$**	**-8.12 $\times 10^{-3}$**	**2.72 $\times 10^{-2}$**
9	y_e	**0.901**	**1.076**	**0.625**
10	p_{y_e}	**-2.81 $\times 10^{-2}$**	as valence	as valence
11	β	1.31	1.3, fixed	1.3, fixed
12	p_β	4.63 $\times 10^{-2}$	0, fixed	0, fixed

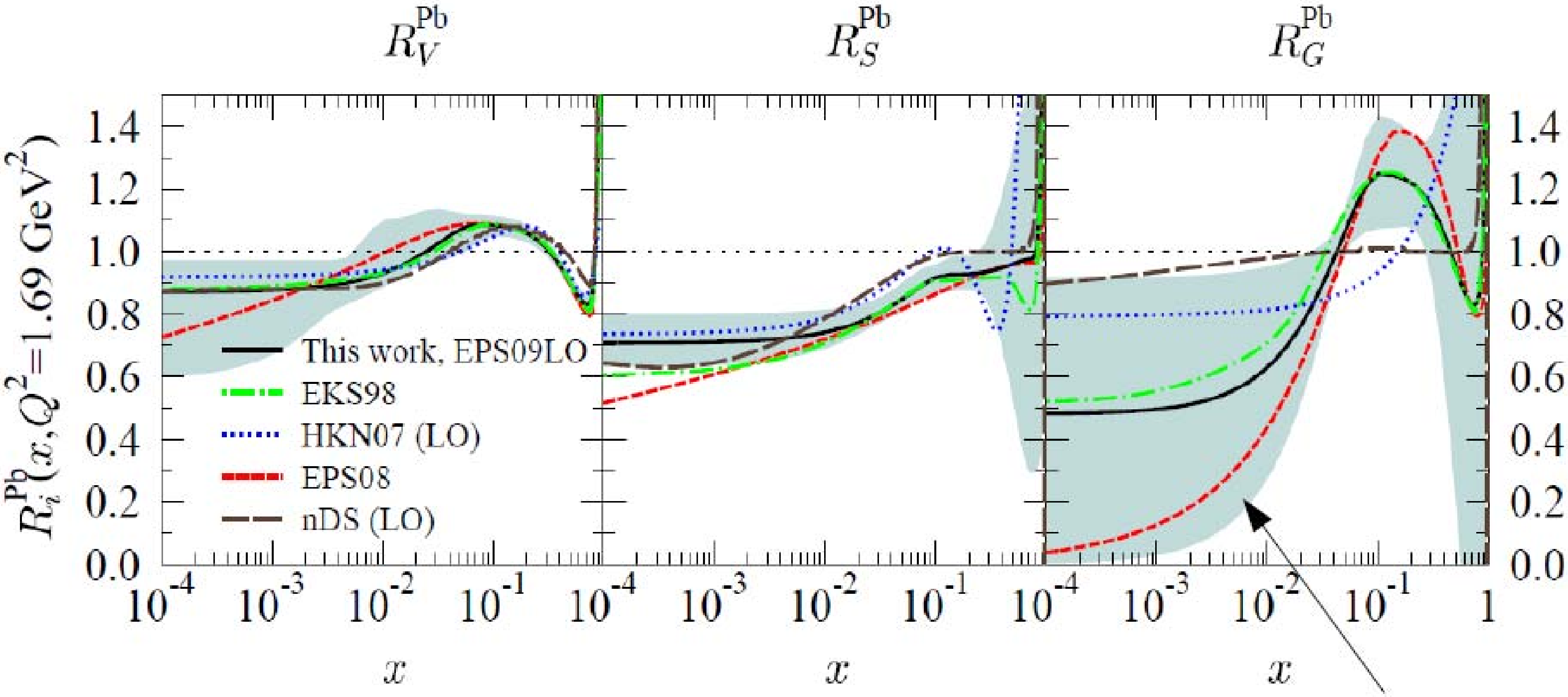
Comparison with earlier works: NLO



- **Difference with EPS09NLO & HKN07 gluons results mainly from the more restricted form of HKN07 fit function and inclusion of pion data**

[from HP]

Comparison with earlier works: LO

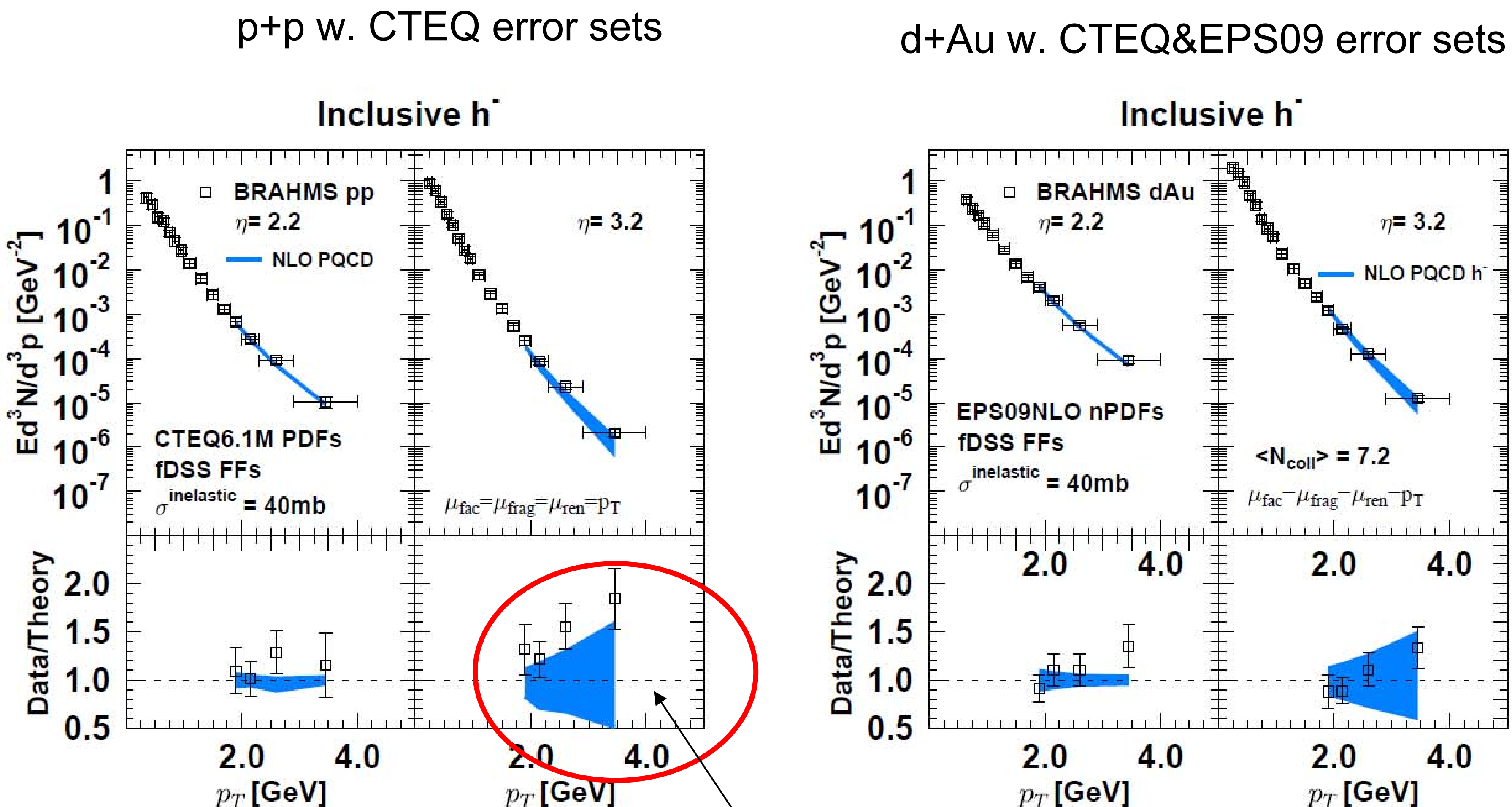


Recall: EPS08 was a set with extreme shadowing

- **The EPS09LO errorbands are clearly larger than in NLO**

[from HP]

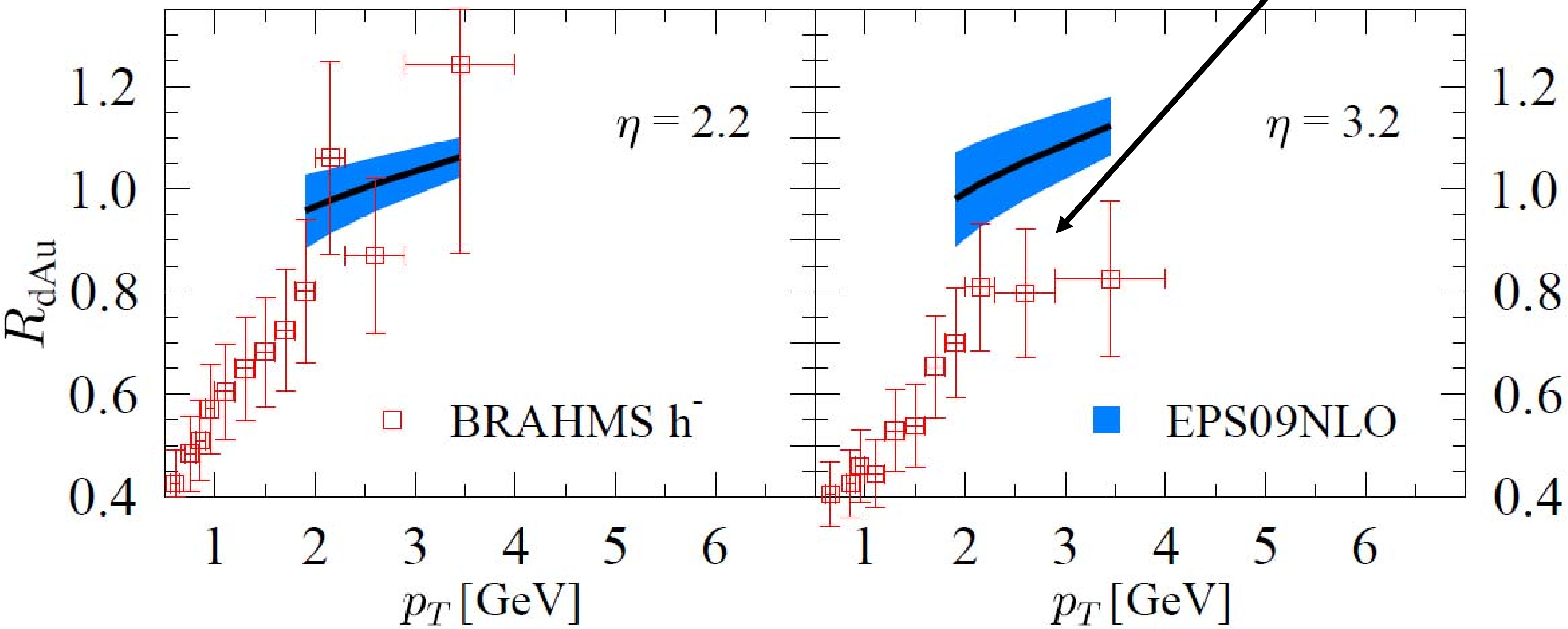
4. An Application: fwd-y h⁻ production at RHIC vs. BRAHMS data



Get 3 panels out of 4 right
– is something wrong with the data here? (difficult measurement)

Origin of the much-debated suppression in the BRAHMS R_{dAu}??

EPS09 → the suppression/saturation observed in the ratio R_{dAu} at pT>2 GeV, y=3.2, is caused by the excessive yield measured in the p+p case, NOT by a suppression in d+Au!



N.B: This data set was not included in EPS09