Status of proton PDFs and rôle of nuclear corrections

Graeme **W**att

CERN PH-TH

CERN Heavy Ion Forum
7th June 2010

Partly based on work done in collaboration with

A. D. **M**artin, W. J. **S**tirling and R. S. **T**horne

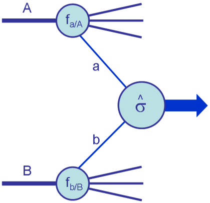
Introduction

- Protons are not elementary particles: made of **partons**.
  ⇒ **P**arton **D**istribution **F**unctions (**PDF**s) essential to relate theory to experiment at the LHC (and Tevatron, HERA, ...).
- $f_{a/A}(x, Q^2)$ gives *number density* of partons a in hadron A with momentum fraction x at a hard scale $Q^2 \gg \Lambda^2_{\rm QCD}$.

$$\sigma_{AB} = \sum_{a,b=q,g} \int_0^1 \mathrm{d}x_a \int_0^1 \mathrm{d}x_b \; f_{a/A}(x_a, Q^2)\; f_{b/B}(x_b, Q^2)\; \hat{\sigma}_{ab}$$

Outline of talk:

1. Introduction to global (proton) PDF analyses.
2. Recent developments and current status.
3. Benchmark cross sections for the LHC.
4. Rôle of nuclear corrections.



Fixed-order collinear factorisation at hadron colliders

- The "standard" pQCD framework: holds up to formally power-suppressed ("higher-twist") terms $\mathcal{O}(\Lambda^2_{\mathrm{QCD}}/Q^2)$.
- Expand $\hat{\sigma}_{ab}$, $P_{aa'}$ and β as perturbative series in α_S ($\mu_R = \mu_F = Q$).

$$\sigma_{AB} = \sum_{a,b=q,g} \left[\hat{\sigma}^{\mathrm{LO}}_{ab} + \alpha_S(Q^2)\hat{\sigma}^{\mathrm{NLO}}_{ab} + \ldots\right] \otimes f_{a/A}(x_a, Q^2) \otimes f_{b/B}(x_b, Q^2)$$

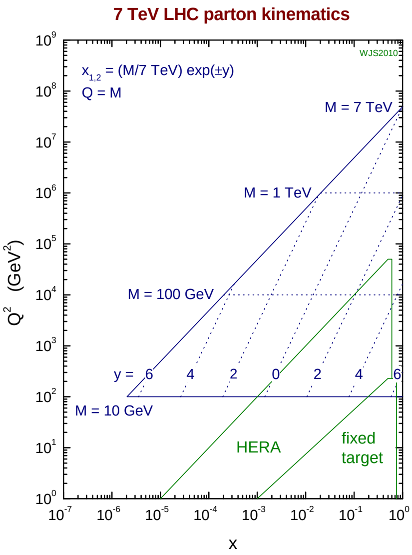
PDF evolution: $$\frac{\partial f_{a/A}}{\partial \ln Q^2} = \frac{\alpha_S}{2\pi} \sum_{a'=q,g} \left[P^{\mathrm{LO}}_{aa'} + \alpha_S P^{\mathrm{NLO}}_{aa'} + \ldots\right] \otimes f_{a'/A}$$

α_S evolution: $$\frac{\partial \alpha_S}{\partial \ln Q^2} = -\beta^{\mathrm{LO}}\alpha_S^2 - \beta^{\mathrm{NLO}}\alpha_S^3 - \ldots$$

- Need to extract input values $f_{a/A}(x, Q_0^2)$ and $\alpha_S(M_Z^2)$ from data.
- Structure functions in deep-inelastic scattering (DIS):

$$F_i\left(x_{\mathrm{Bj}}, Q^2\right) = \sum_{a=q,g} C_{i,a} \otimes f_{a/A}, \quad C_{i,a} = C^{\mathrm{LO}}_{i,a} + \alpha_S C^{\mathrm{NLO}}_{i,a} + \ldots$$

From HERA *et al.* to the LHC



- PDFs are **universal**.
- Fit existing data from HERA and fixed-target experiments, together with Tevatron data.
- HERA *ep* (H1, ZEUS).
- Fixed-target experiments: ℓp, ℓd (BCDMS, NMC, E665, SLAC), νN (CCFR, NuTeV, CHORUS), *pp*, *pd* (E866/NuSea).
- Tevatron $p\bar{p}$ (CDF, DØ).
- DGLAP evolution gives PDFs at higher Q^2 for LHC.

Paradigm for PDF determination by "global analysis"

1. **Parameterise** the x dependence for each flavour $a = q, g$ at the input scale $Q_0^2 \sim 1\ \mathrm{GeV}^2$ in some flexible form, e.g.

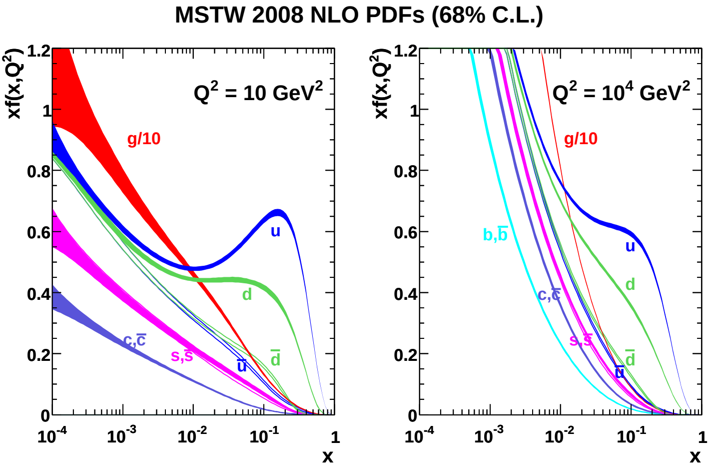
$$xf_{a/p}(x, Q_0^2) = A_a\, x^{\Delta_a}\, (1-x)^{\eta_a}\, (1 + \epsilon_a \sqrt{x} + \gamma_a\, x),$$

   subject to number- and momentum-sum rule constraints.
2. **Evolve** the PDFs to higher scales $Q^2 > Q_0^2$ using the DGLAP (Dokshitzer–Gribov–Lipatov–Altarelli–Parisi) evolution equations.
3. **Convolute** the evolved PDFs with $C_{i,a}$ and $\hat{\sigma}_{ab}$ to calculate theory predictions corresponding to a wide variety of data.
4. **Vary** the input parameters $\{A_a, \Delta_a, \eta_a, \epsilon_a, \gamma_a, \ldots\}$ to minimise

$$\chi^2 = \sum_{i=1}^{N_{\mathrm{pts.}}} \left(\frac{\mathrm{Data}_i - \mathrm{Theory}_i}{\mathrm{Error}_i} \right)^2$$

   or generalisations to account for *correlated* systematic errors.

Example of PDFs obtained from global analysis



- **Error bands** shown are obtained from propagation of experimental uncertainties on the fitted data points.

Criteria for choice of tolerance $T = \sqrt{\Delta\chi^2_{\rm global}}$

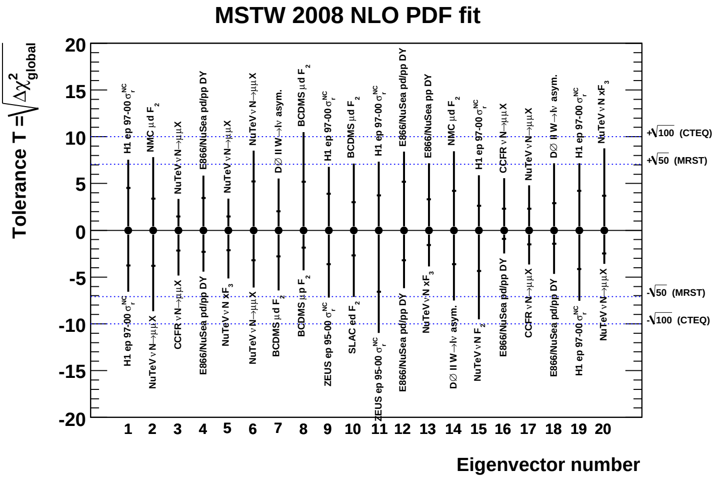
Parameter-fitting criterion

- $T^2 = 1$ for 68% (1-σ) C.L., $T^2 = 2.71$ for 90% C.L.
- **In practice:** minor inconsistencies between fitted data sets, and unknown experimental and theoretical uncertainties, so not appropriate for global PDF analysis.

Hypothesis-testing criterion (proposed by CTEQ)

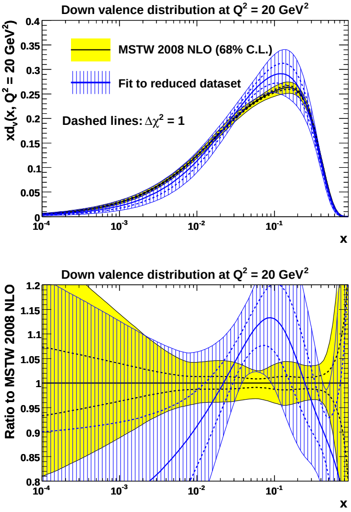
- Much weaker: treat PDF sets obtained from eigenvectors of covariance matrix as alternative hypotheses.
- Determine T^2 from the criterion that each data set should be described within its 90% C.L. limit. Very roughly, a "good" fit has $\chi^2 \simeq N_{\rm pts.} \pm \sqrt{2N_{\rm pts.}}$ for each data set.
- **CTEQ:** $T^2 = 100$ for 90% C.L. limit, **MRST:** $T^2 = 50$.
- Large T^2 also for *nuclear* PDFs, e.g. $T^2 = 50$ for **EPS09**.

Dynamic tolerance: different for each eigenvector



- Outer (inner) error bars give tolerance for 90% (68%) C.L.

Test of dynamic tolerance: fit to reduced dataset



- Fit to reduced dataset comprising 589 DIS data points, cf. 2699 data points in global fit.
- Errors given by $T^2 = 1$ don't overlap $\Rightarrow$ inconsistent data sets included in global fit.
- **Dynamic tolerance** $T^2 > 1$ accommodates mildly inconsistent data sets.
- **Issues:**
  1. $T^2 > 1$ not rigorous?
  2. Dependence on input parameterisation?

Alternative approach: NNPDF Collaboration

MSTW approach [arXiv:0901.0002] (CTEQ similar)

Parameterisation	$xf_{a/p} \sim A_a\, x^{\Delta_a}\, (1-x)^{\eta_a}\, (1+\epsilon_a\, \sqrt{x} + \gamma_a\, x)$
Minimisation	Non-linear least-squares (Marquardt method)
Error propagation	Hessian method with dynamical tolerance
Application	Use best-fit and 40 eigenvector PDF sets

NNPDF approach [arXiv:1002.4407, and references therein]

Parameterisation	Neural network (37 free parameters per PDF)
Minimisation	Genetic algorithm (stop before overlearning)
Error propagation	Generate $N_{\rm rep} \sim \mathcal{O}(1000)$ MC data replicas
Application	Calculate average and s.d. over $N_{\rm rep}$ PDF sets

- It would be interesting to apply the NNPDF approach to the determination of *nuclear* PDFs, where parameterisation bias is more relevant than for *proton* PDFs due to less available data.

Alternative approach: NNPDF Collaboration

MSTW approach [arXiv:0901.0002] (CTEQ similar)

Parameterisation	$xf_{a/p} \sim A_a\, x^{\Delta_a}\, (1-x)^{\eta_a}\, (1+\epsilon_a\,\sqrt{x}+\gamma_a\, x)$
Minimisation	Non-linear least-squares (Marquardt method)
Error propagation	Hessian method with dynamical tolerance
Application	Use best-fit and 40 eigenvector PDF sets

NNPDF approach [arXiv:1002.4407, and references therein]

Parameterisation	Neural network (37 free parameters per PDF)
Minimisation	Genetic algorithm (stop before overlearning)
Error propagation	Generate $N_{\rm rep} \sim \mathcal{O}(1000)$ MC data replicas
Application	Calculate average and s.d. over $N_{\rm rep}$ PDF sets

- It would be interesting to apply the NNPDF approach to the determination of *nuclear* PDFs, where parameterisation bias is more relevant than for *proton* PDFs due to less available data.

Alternative approach: NNPDF Collaboration

MSTW approach [arXiv:0901.0002] (CTEQ similar)

Parameterisation	$xf_{a/p} \sim A_a \, x^{\Delta_a} \, (1-x)^{\eta_a} \, (1 + \epsilon_a \, \sqrt{x} + \gamma_a \, x)$
Minimisation	Non-linear least-squares (Marquardt method)
Error propagation	Hessian method with dynamical tolerance
Application	Use best-fit and 40 eigenvector PDF sets

NNPDF approach [arXiv:1002.4407, and references therein]

Parameterisation	Neural network (37 free parameters per PDF)
Minimisation	Genetic algorithm (stop before overlearning)
Error propagation	Generate $N_{\rm rep} \sim \mathcal{O}(1000)$ MC data replicas
Application	Calculate average and s.d. over $N_{\rm rep}$ PDF sets

- It would be interesting to apply the NNPDF approach to the determination of *nuclear* PDFs, where parameterisation bias is more relevant than for *proton* PDFs due to less available data.

Alternative approach: NNPDF Collaboration

MSTW approach [arXiv:0901.0002] (CTEQ similar)

Parameterisation	$xf_{a/p} \sim A_a\, x^{\Delta_a}\, (1-x)^{\eta_a}\, (1 + \epsilon_a \sqrt{x} + \gamma_a\, x)$
Minimisation	Non-linear least-squares (Marquardt method)
Error propagation	Hessian method with dynamical tolerance
Application	Use best-fit and 40 eigenvector PDF sets

NNPDF approach [arXiv:1002.4407, and references therein]

Parameterisation	Neural network (37 free parameters per PDF)
Minimisation	Genetic algorithm (stop before overlearning)
Error propagation	Generate $N_{\rm rep} \sim \mathcal{O}(1000)$ MC data replicas
Application	Calculate average and s.d. over $N_{\rm rep}$ PDF sets

- It would be interesting to apply the NNPDF approach to the determination of *nuclear* PDFs, where parameterisation bias is more relevant than for *proton* PDFs due to less available data.

Alternative approach: NNPDF Collaboration

MSTW approach [arXiv:0901.0002] (CTEQ similar)

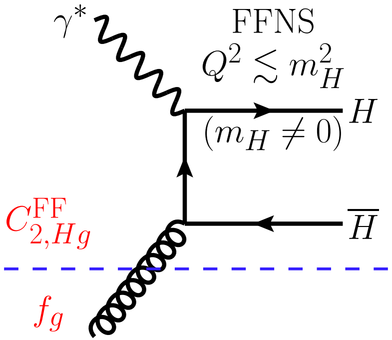
Parameterisation	$xf_{a/p} \sim A_a\, x^{\Delta_a}\, (1-x)^{\eta_a}\, (1+\epsilon_a\sqrt{x}+\gamma_a\, x)$
Minimisation	Non-linear least-squares (Marquardt method)
Error propagation	Hessian method with dynamical tolerance
Application	Use best-fit and 40 eigenvector PDF sets

NNPDF approach [arXiv:1002.4407, and references therein]

Parameterisation	Neural network (37 free parameters per PDF)
Minimisation	Genetic algorithm (stop before overlearning)
Error propagation	Generate $N_{\rm rep} \sim \mathcal{O}(1000)$ MC data replicas
Application	Calculate average and s.d. over $N_{\rm rep}$ PDF sets

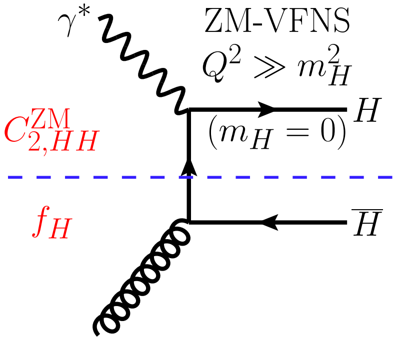
- It would be interesting to apply the NNPDF approach to the determination of *nuclear* PDFs, where parameterisation bias is more relevant than for *proton* PDFs due to less available data.

Heavy quark contribution to DIS structure function F_2



Fixed flavour number scheme

- No heavy quark PDF.
- Includes $\mathcal{O}(m_H^2/Q^2)$ terms.
- No resummation of $\alpha_S \ln(Q^2/m_H^2)$ terms.



Zero-mass variable flavour number scheme

- Use heavy quark PDF.
- Mass dependence neglected.
- Resums $\alpha_S \ln(Q^2/m_H^2)$ terms similar to light quarks.

General-mass variable flavour number scheme (GM-VFNS)

- Interpolate between two well-defined regions.
- FFNS for $Q^2 \leq m_H^2$, ZM-VFNS for $Q^2 \gg m_H^2$.
- Details of interpolation are ambiguous ($\Rightarrow$ **uncertainty**).

CTEQ6.1 (ZM-VFNS) → CTEQ6.5 (GM-VFNS)

- 8% increase in W and Z cross sections at LHC (14 TeV).

MRST 1998–2006 and MSTW 2008

- MRST/MSTW group have used a GM-VFNS since 1998.

NNPDF 1.0–2.0

- Fits still use ZM-VFNS ($\to$ GM-VFNS [arXiv:1002.4407]).

Recent comparison between GM-VFNS used by different groups:

J. Rojo, S. Forte, J. Huston, P. Nadolsky, P. Nason, F. Olness, **R. Thorne and G. W.**

"*The Les Houches benchmarks for GM-VFN heavy quark schemes in DIS*" [arXiv:1003.1241]

Most recent public NLO PDF sets from fitting groups

MSTW08 [arXiv:0901.0002]
GM-VFNS **global** fit to DIS, DY and jet data.

CTEQ6.6 [arXiv:0802.0007]
GM-VFNS **global** fit to DIS, DY and jet data.

NNPDF2.0 [arXiv:1002.4407]
ZM-VFNS **global** fit to DIS, DY and jet data.

HERAPDF1.0 [arXiv:0911.0884]
GM-VFNS fit only to inclusive HERA data.

ABKM09 [arXiv:0908.2766, with NLO update in January 2010]
FFNS fit to DIS and DY data.

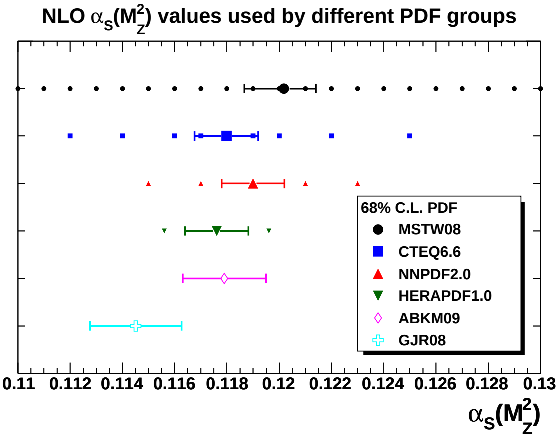
GJR08 [arXiv:0709.0614]
FFNS 'dynamical' fit to DIS, DY and jet data.

- **Only** MSTW08, ABKM09 and JR09 have **NNLO** PDFs.

Benchmark cross sections: where do we stand?

- **PDF4LHC** exercise: results presented by G.W. on 26th March.
- Use most recent public NLO PDFs from **all** fitting groups to calculate LHC benchmark processes: $W^{\pm}$, Z^0, $t\bar{t}$, $gg \to H$.
- Comparisons at **NLO** since only three groups have NNLO PDFs.
- Complication: all groups use different values of $\alpha_S(M_Z^2)$. However:
  - All groups now provide a prescription for "PDF+α_S" uncertainties.
  - Most provide alternative PDF sets for different $\alpha_S(M_Z^2)$ values.
- **Aims of exercise:**
  1. Establish degree of compatibility and identify outliers.
  2. Compare cross sections at **same** α_S values.
  3. To what extent are differences in predictions due to different α_S values used by each group, rather than differences in PDFs?
- Plots:
  `http://projects.hepforge.org/mstwpdf/pdf4lhc/`

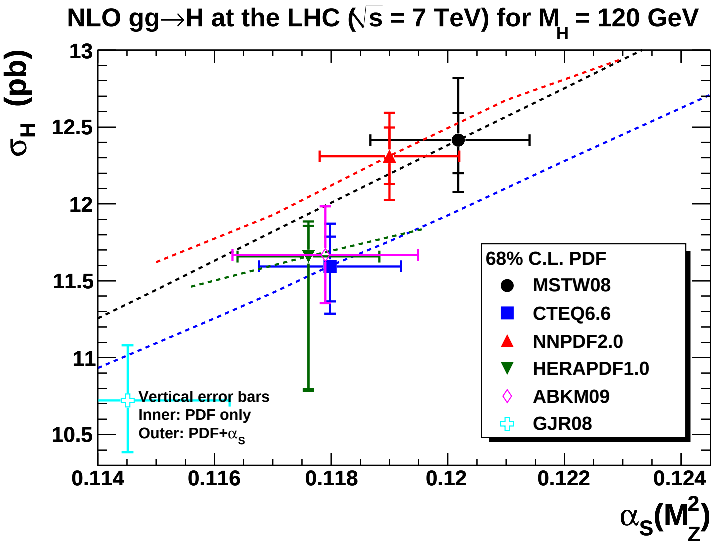
Values of $\alpha_S(M_Z^2)$ used by different fitting groups



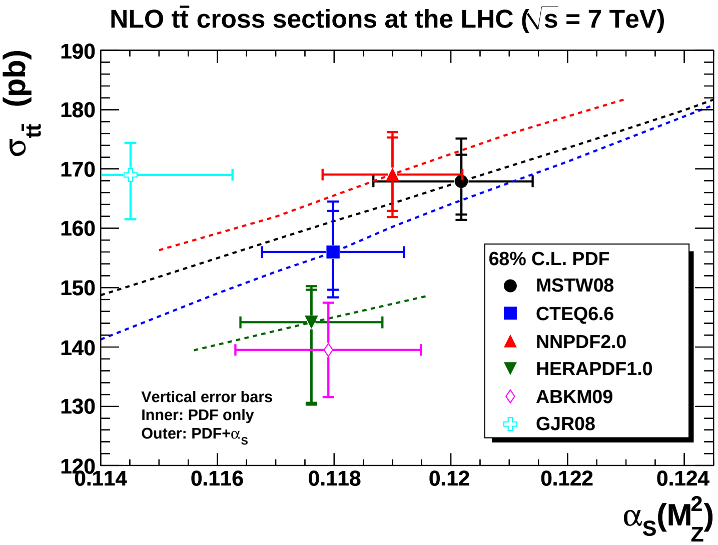
- $\alpha_S(M_Z^2)$ for MSTW08, ABKM09 and GJR08 obtained from fit.
- $\alpha_S(M_Z^2)$ for other groups applied as an external constraint.
- Smaller symbols indicate alternative $\alpha_S(M_Z^2)$ values provided.

(NNPDF and HERAPDF now provide more values in the latest LHAPDF V5.8.3 released 18th May.)

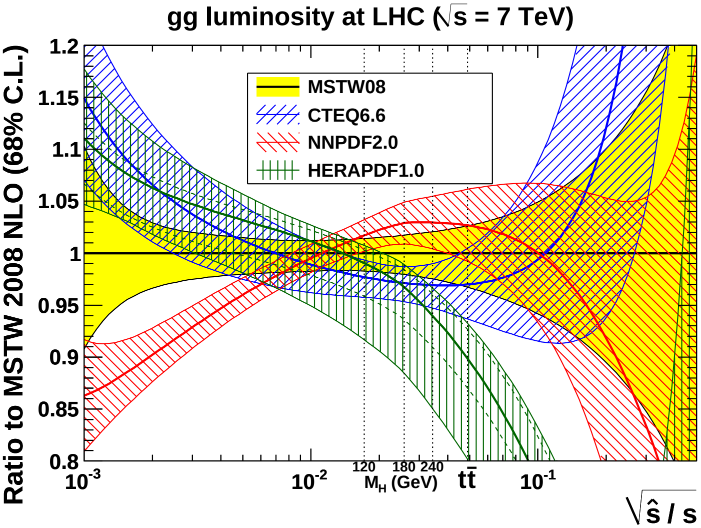
Higgs ($M_H = 120$ GeV) total cross section vs. $\alpha_S(M_Z^2)$



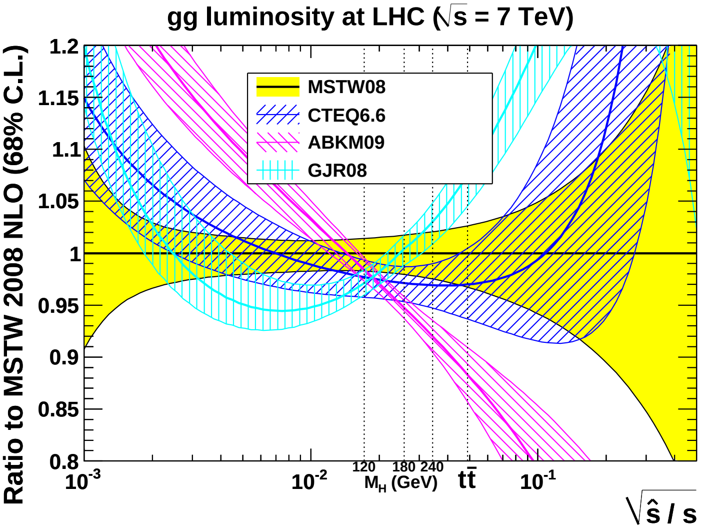
$t\bar{t}$ total cross section vs. $\alpha_S(M_Z^2)$



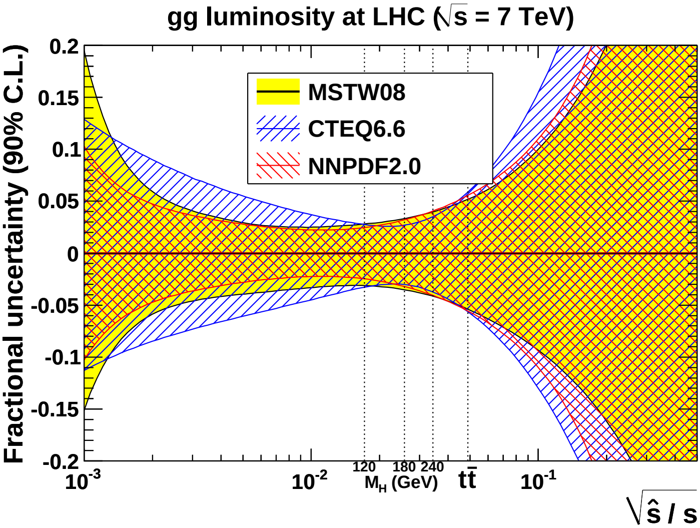
Ratio of gluon–gluon luminosity functions (1)



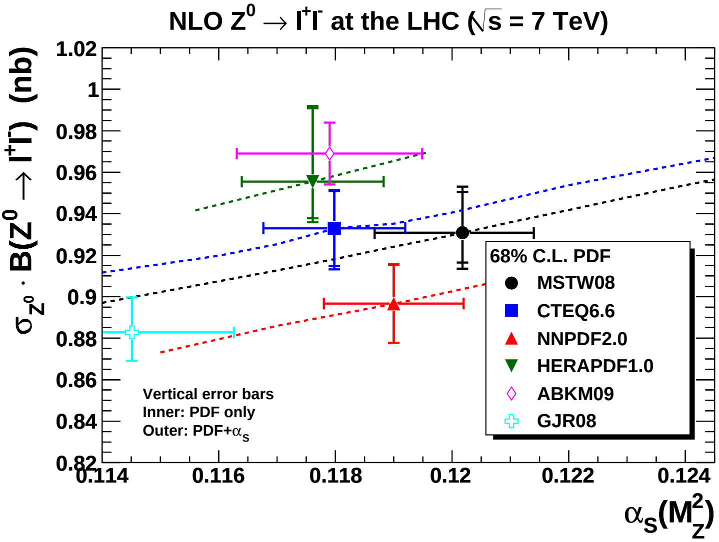
Ratio of gluon–gluon luminosity functions (2)



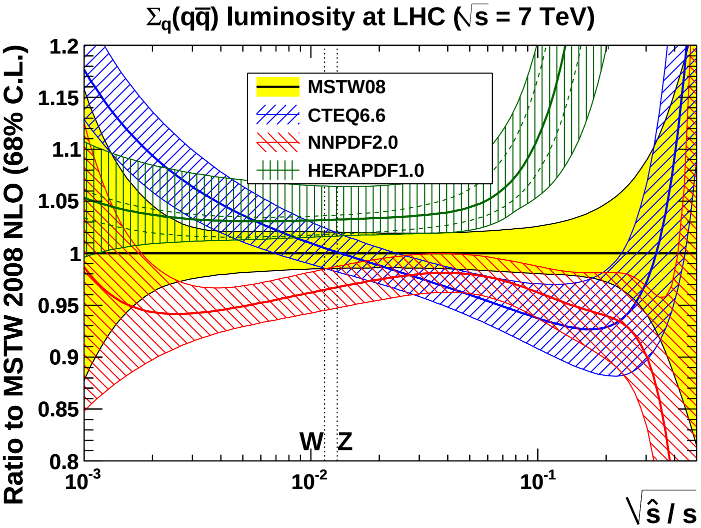
Fractional uncertainty in gluon–gluon luminosity



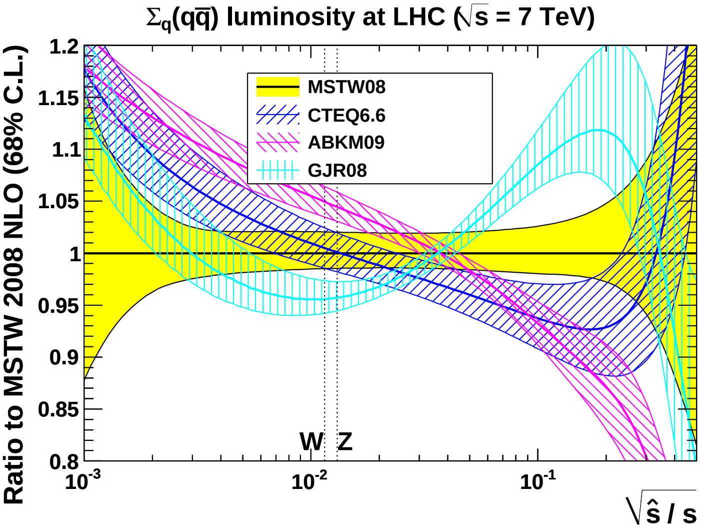
Z^0 total cross section vs. $\alpha_S(M_Z^2)$



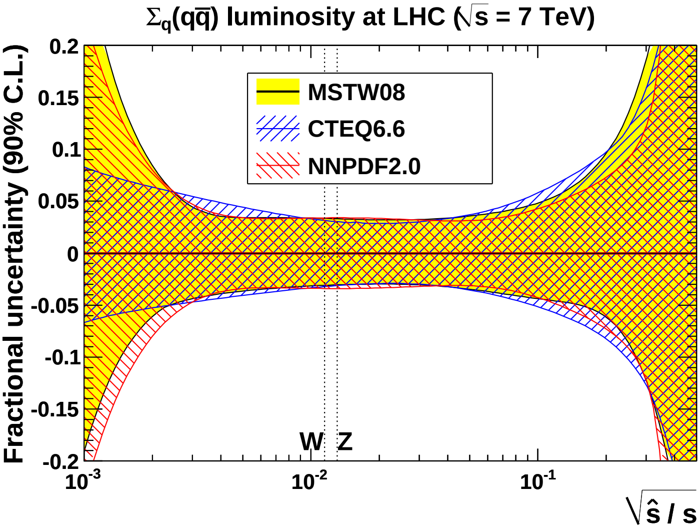
Ratio of quark–antiquark luminosity functions (1)



Ratio of quark–antiquark luminosity functions (2)



Fractional uncertainty in quark–antiquark luminosity



Summary: current status of proton PDFs

- Highlight major differences in data and theory between groups.

	MSTW08	**CTEQ6.6**	**NNPDF2.0**	HERAPDF1.0	ABKM09	GJR08
HERA DIS	✓	✓	✓	✓	✓	✓
Fixed-target DIS	✓	✓	✓	✗	✓	✓
Fixed-target DY	✓	✓	✓	✗	✓	✓
Tevatron W,Z	✓	✓	✓	✗	✗	✗
Tevatron jets	✓	✓	✓	✗	✗	✓
GM-VFNS	✓	✓	✗	✓	✗	✗
NNLO	✓	✗	✗	✗	✓	✓

- *Three* **NLO** global fits, but only *one* **NNLO** global fit.
- CTEQ6.6 only uses Tevatron Run I data, not Run II.
- Only NNPDF2.0 and HERAPDF1.0 use *combined* HERA data.

Which processes constrain different PDFs?

- Processes included in MSTW 2008 analysis [arXiv:0901.0002]:

Process	Subprocess	Partons	x range
$\ell^{\pm}\{p,n\} \to \ell^{\pm} X$	$\gamma^* q \to q$	$q, \bar{q}, g$	$x \gtrsim 0.01$
$\ell^{\pm} n/p \to \ell^{\pm} X$	$\gamma^* d/u \to d/u$	d/u	$x \gtrsim 0.01$
$pp \to \mu^+\mu^- X$	$u\bar{u}, d\bar{d} \to \gamma^*$	$\bar{q}$	$0.015 \lesssim x \lesssim 0.35$
$pn/pp \to \mu^+\mu^- X$	$(u\bar{d})/(u\bar{u}) \to \gamma^*$	$\bar{d}/\bar{u}$	$0.015 \lesssim x \lesssim 0.35$
$\nu(\bar{\nu})\, N \to \mu^-(\mu^+)\, X$	$W^* q \to q'$	$q, \bar{q}$	$0.01 \lesssim x \lesssim 0.5$
$\nu\, N \to \mu^-\mu^+ X$	$W^* s \to c$	s	$0.01 \lesssim x \lesssim 0.2$
$\bar{\nu}\, N \to \mu^+\mu^- X$	$W^* \bar{s} \to \bar{c}$	$\bar{s}$	$0.01 \lesssim x \lesssim 0.2$
$e^{\pm} p \to e^{\pm} X$	$\gamma^* q \to q$	$g, q, \bar{q}$	$0.0001 \lesssim x \lesssim 0.1$
$e^+ p \to \bar{\nu} X$	$W^+ \{d, s\} \to \{u, c\}$	d, s	$x \gtrsim 0.01$
$e^{\pm} p \to e^{\pm} c\bar{c} X$	$\gamma^* c \to c,\ \gamma^* g \to c\bar{c}$	$c,\ g$	$0.0001 \lesssim x \lesssim 0.01$
$e^{\pm} p \to \text{jet} + X$	$\gamma^* g \to q\bar{q}$	g	$0.01 \lesssim x \lesssim 0.1$
$p\bar{p} \to \text{jet} + X$	$gg, qg, qq \to 2j$	g, q	$0.01 \lesssim x \lesssim 0.5$
$p\bar{p} \to (W^{\pm} \to \ell^{\pm}\nu)\, X$	$ud \to W, \bar{u}\bar{d} \to W$	$u, d, \bar{u}, \bar{d}$	$x \gtrsim 0.05$
$p\bar{p} \to (Z \to \ell^+\ell^-)\, X$	$uu, dd \to Z$	d	$x \gtrsim 0.05$

- Nuclear corrections needed for n (deuteron) and N (Fe or Pb).

Which processes constrain different PDFs?

- Processes included in MSTW 2008 analysis [arXiv:0901.0002]:

Process	Subprocess	Partons	x range
$\ell^\pm\,\{p,n\} \to \ell^\pm\,X$	$\gamma^* q \to q$	$q, \bar{q}, g$	$x \gtrsim 0.01$
$\ell^\pm\,n/p \to \ell^\pm\,X$	$\gamma^*\,d/u \to d/u$	d/u	$x \gtrsim 0.01$
$pp \to \mu^+\mu^-\,X$	$u\bar{u}, d\bar{d} \to \gamma^*$	$\bar{q}$	$0.015 \lesssim x \lesssim 0.35$
$pn/pp \to \mu^+\mu^-\,X$	$(u\bar{d})/(u\bar{u}) \to \gamma^*$	$\bar{d}/\bar{u}$	$0.015 \lesssim x \lesssim 0.35$
$\nu(\bar{\nu})\,N \to \mu^-(\mu^+)\,X$	$W^* q \to q'$	$q, \bar{q}$	$0.01 \lesssim x \lesssim 0.5$
$\nu\,N \to \mu^-\mu^+\,X$	$W^* s \to c$	s	$0.01 \lesssim x \lesssim 0.2$
$\bar{\nu}\,N \to \mu^+\mu^-\,X$	$W^* \bar{s} \to \bar{c}$	$\bar{s}$	$0.01 \lesssim x \lesssim 0.2$
$e^\pm\,p \to e^\pm\,X$	$\gamma^* q \to q$	$g, q, \bar{q}$	$0.0001 \lesssim x \lesssim 0.1$
$e^+\,p \to \bar{\nu}\,X$	$W^+\,\{d,s\} \to \{u,c\}$	d, s	$x \gtrsim 0.01$
$e^\pm p \to e^\pm\,c\bar{c}\,X$	$\gamma^* c \to c,\ \gamma^* g \to c\bar{c}$	$c,\ g$	$0.0001 \lesssim x \lesssim 0.01$
$e^\pm p \to \text{jet} + X$	$\gamma^* g \to q\bar{q}$	g	$0.01 \lesssim x \lesssim 0.1$
$p\bar{p} \to \text{jet} + X$	$gg, qg, qq \to 2j$	g, q	$0.01 \lesssim x \lesssim 0.5$
$p\bar{p} \to (W^\pm \to \ell^\pm\nu)\,X$	$ud \to W, \bar{u}\bar{d} \to W$	$u, d, \bar{u}, \bar{d}$	$x \gtrsim 0.05$
$p\bar{p} \to (Z \to \ell^+\ell^-)\,X$	$uu, dd \to Z$	d	$x \gtrsim 0.05$

- Nuclear corrections needed for n (deuteron) and N (Fe or Pb).

Data sets fitted in MSTW 2008 NLO analysis [arXiv:0901.0002]

Data set	χ^2 / $N_{\rm pts.}$
H1 MB 99 e^+p NC	9 / 8
H1 MB 97 e^+p NC	42 / 64
H1 low Q^2 96–97 e^+p NC	44 / 80
H1 high Q^2 98–99 e^-p NC	122 / 126
H1 high Q^2 99–00 e^+p NC	131 / 147
ZEUS SVX 95 e^+p NC	35 / 30
ZEUS 96–97 e^+p NC	86 / 144
ZEUS 98–99 e^-p NC	54 / 92
ZEUS 99–00 e^+p NC	63 / 90
H1 99–00 e^+p CC	29 / 28
ZEUS 99–00 e^+p CC	38 / 30
H1/ZEUS $e^\pm p$ $F_2^{\rm charm}$	107 / 83
H1 99–00 e^+p **incl. jets**	19 / 24
ZEUS 96–97 e^+p **incl. jets**	30 / 30
ZEUS 98–00 $e^\pm p$ **incl. jets**	17 / 30
DØ II $p\bar{p}$ **incl. jets**	114 / 110
CDF II $p\bar{p}$ **incl. jets**	56 / 76
CDF II $W \to l\nu$ **asym.**	29 / 22
DØ II $W \to l\nu$ **asym.**	25 / 10
DØ II Z **rap.**	19 / 28
CDF II Z **rap.**	49 / 29

Data set	χ^2 / $N_{\rm pts.}$
BCDMS μp F_2	182 / 163
BCDMS μd F_2	190 / 151
NMC μp F_2	121 / 123
NMC μd F_2	102 / 123
NMC $\mu n/\mu p$	130 / 148
E665 μp F_2	57 / 53
E665 μd F_2	53 / 53
SLAC ep F_2	30 / 37
SLAC ed F_2	30 / 38
NMC/BCDMS/SLAC F_L	38 / 31
E866/NuSea pp DY	228 / 184
E866/NuSea pd/pp DY	14 / 15
NuTeV νN F_2	49 / 53
CHORUS νN F_2	26 / 42
NuTeV νN xF_3	40 / 45
CHORUS νN xF_3	31 / 33
CCFR $\nu N \to \mu\mu X$	66 / 86
NuTeV $\nu N \to \mu\mu X$	39 / 40
All data sets	**2543 / 2699**

- **Bold** = New w.r.t. MRST 2006 fit.
- Red = Needs nuclear corrections.

Data sets fitted in MSTW 2008 NLO analysis [arXiv:0901.0002]

Data set	χ^2 / $N_{\rm pts.}$
H1 MB 99 e^+p NC	9 / 8
H1 MB 97 e^+p NC	42 / 64
H1 low Q^2 96–97 e^+p NC	44 / 80
H1 high Q^2 98–99 e^-p NC	122 / 126
H1 high Q^2 99–00 e^+p NC	131 / 147
ZEUS SVX 95 e^+p NC	35 / 30
ZEUS 96–97 e^+p NC	86 / 144
ZEUS 98–99 e^-p NC	54 / 92
ZEUS 99–00 e^+p NC	63 / 90
H1 99–00 e^+p CC	29 / 28
ZEUS 99–00 e^+p CC	38 / 30
H1/ZEUS $e^\pm p$ $F_2^{\rm charm}$	107 / 83
H1 99–00 e^+p incl. jets	19 / 24
ZEUS 96–97 e^+p incl. jets	30 / 30
ZEUS 98–00 $e^\pm p$ incl. jets	17 / 30
DØ II $p\bar{p}$ incl. jets	114 / 110
CDF II $p\bar{p}$ incl. jets	56 / 76
CDF II $W \to l\nu$ asym.	29 / 22
DØ II $W \to l\nu$ asym.	25 / 10
DØ II Z rap.	19 / 28
CDF II Z rap.	49 / 29

Data set	χ^2 / $N_{\rm pts.}$
BCDMS μp F_2	182 / 163
BCDMS μd F_2	190 / 151
NMC μp F_2	121 / 123
NMC μd F_2	102 / 123
NMC $\mu n/\mu p$	130 / 148
E665 μp F_2	57 / 53
E665 μd F_2	53 / 53
SLAC ep F_2	30 / 37
SLAC ed F_2	30 / 38
NMC/BCDMS/SLAC F_L	38 / 31
E866/NuSea pp DY	228 / 184
E866/NuSea pd/pp DY	14 / 15
NuTeV νN F_2	49 / 53
CHORUS νN F_2	26 / 42
NuTeV νN xF_3	40 / 45
CHORUS νN xF_3	31 / 33
CCFR $\nu N \to \mu\mu X$	66 / 86
NuTeV $\nu N \to \mu\mu X$	39 / 40
All data sets	**2543 / 2699**

- **Bold** = New w.r.t. MRST 2006 fit.
- Red = Needs nuclear corrections.

Why do we use data taken on nuclear targets?

- Assume isospin symmetry ($u \equiv u^p = d^n$, $d \equiv d^p = u^n$) and isoscalar target ($N = (p+n)/2$), then structure functions are:

Neutral-current DIS in charged-lepton–nucleon scattering

$$F_2^{\ell^\pm p} = x\left[\frac{4}{9}\left(u+\bar{u}+\ldots\right)+\frac{1}{9}\left(d+\bar{d}+\ldots\right)\right]$$

$$F_2^{\ell^\pm d} = (F_2^{\ell^\pm p}+F_2^{\ell^\pm n})/2 = \frac{5}{18}x\left(u+\bar{u}+d+\bar{d}\right)+\ldots$$

Charged-current DIS in neutrino–nucleus scattering

$$F_2 \equiv (F_2^{\nu N}+F_2^{\bar{\nu}N})/2 = x\left(u+\bar{u}+d+\bar{d}+\ldots\right)$$

$$xF_3 \equiv (xF_3^{\nu N}+xF_3^{\bar{\nu}N})/2 = x\left(u-\bar{u}+d-\bar{d}+\ldots\right)$$

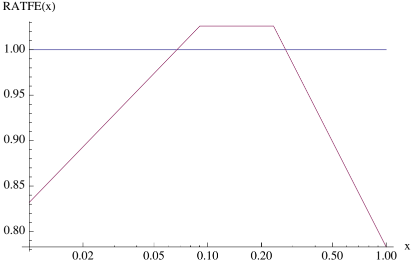
- $\nu N \to \mu\mu X$ constrains s; $\bar{\nu}N \to \mu\mu X$ constrains $\bar{s}$.

Deuterium and neutrino DIS data provide **flavour separation**!

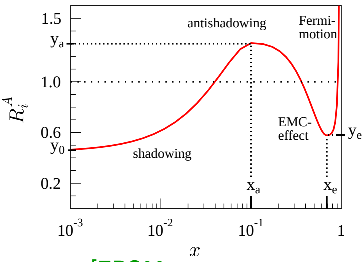
Treatment of nuclear corrections by MRST (1998–2006)

- Fit νN data on F_2 and xF_3 from CCFR (iron target).
- Empirical nuclear correction "RATFE" derived from μ–Fe data.
- Correction assumed to be Q^2-/process-/flavour- independent.

Correction used by MRST:



Compare to generic correction:



[EPS09, arXiv:0902.4154]

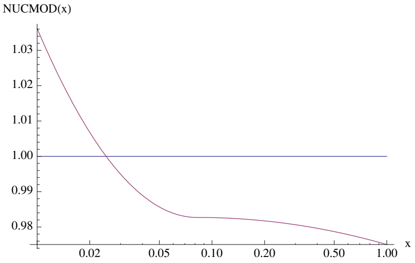
- Similar treatment in standard CTEQ fits (→ I. Schienbein).
- NNPDF use νN data (Fe and Pb) but **without** nuclear effects.

Improvements in MRST 2006 $\longrightarrow$ MSTW 2008

- CCFR F_2 and xF_3 data $\rightarrow$ NuTeV (Fe) and CHORUS (Pb).
- Apply x-, Q^2-, A- and flavour- **dependent** nuclear correction:

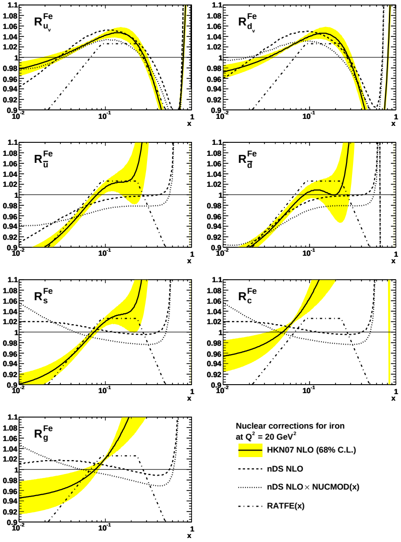
$$R_f^A(x, Q^2) = f^A(x, Q^2)/f(x, Q^2).$$

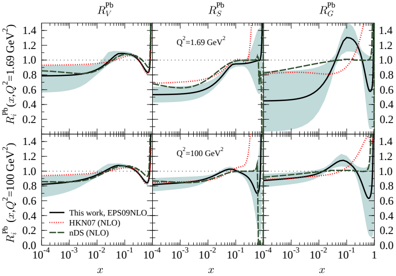
- Only NLO fit available in 2006 was nDS [de Florian and Sassot, '03].
- Determination of R_f^A used free proton PDFs as input (GRV98). In principle, some degree of circularity ($\rightarrow$ I. Schienbein).



- Allow for **uncertainty** with extra modification function (Q^2-, A-, f- independent).
- Three parameters allowed to go free in global fits, which enter covariance matrix used for final error propagation.

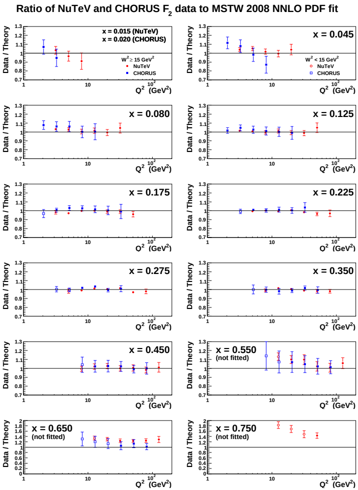
Comparison of nuclear corrections for *Fe* and *Pb*

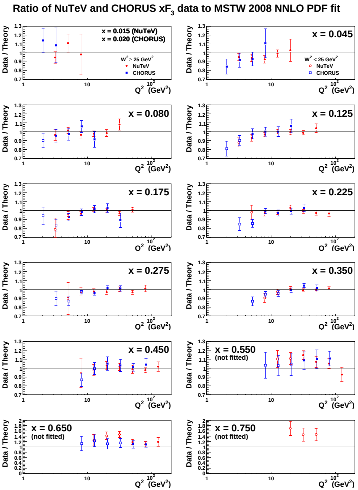




- Two later NLO fits with uncertainties (HKN07, EPS09).
- Reasonable agreement for most relevant R_u^A and R_d^A.
- Larger uncertainties present for R_s^A and (less relevant) R_g^A.

Description of NuTeV and CHORUS data on F_2 and xF_3

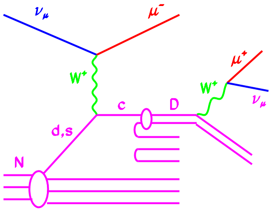




χ^2 for F_2 data	LO	NLO	NNLO
NuTeV (49 pts.)	49	49	46
CHORUS (37 pts.)	21	26	29

χ^2 for xF_3 data	LO	NLO	NNLO
NuTeV (45 pts.)	62	40	34
CHORUS (33 pts.)	44	31	26

NuTeV/CCFR dimuon cross sections and strangeness



$$\frac{\mathrm{d}\sigma}{\mathrm{d}x\mathrm{d}y}(\nu_\mu N \to \mu^+\mu^- X) = B_c\,\mathcal{A}\,\frac{\mathrm{d}\sigma}{\mathrm{d}x\mathrm{d}y}(\nu_\mu N \to \mu^- c\, X)$$
$$\propto |V_{cs}|^2 \xi s^A(\xi, Q^2) + |V_{cd}|^2 \ldots$$

- ν_μ and $\bar{\nu}_\mu$ cross sections constrain s^A and $\bar{s}^A$.
- Again, use R_f^A (nDS) to relate to **proton** PDFs.

- Can **relax assumption** made in previous fits that

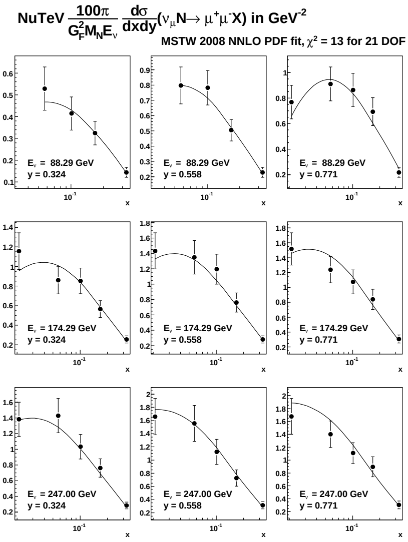
$$s(x, Q_0^2) = \bar{s}(x, Q_0^2) = \frac{\kappa}{2}\left[\bar{u}(x, Q_0^2) + \bar{d}(x, Q_0^2)\right], \text{ with } \kappa \approx 0.5.$$

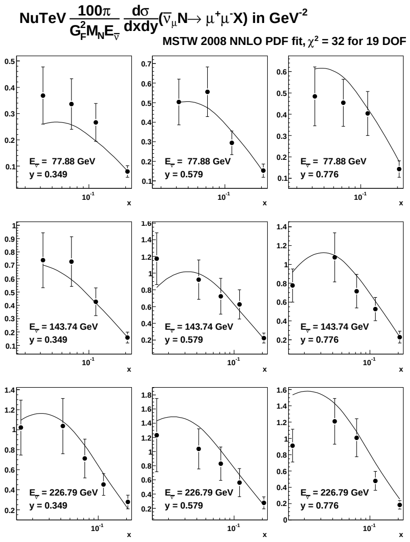
- MSTW **parameterise** at input scale of $Q_0^2 = 1\ \mathrm{GeV}^2$ in the form:

$$xs(x, Q_0^2) + x\bar{s}(x, Q_0^2) = A_+\,(1-x)^{\eta_+}\,xS(x, Q_0^2),$$
$$xs(x, Q_0^2) - x\bar{s}(x, Q_0^2) = A_-\,x^{0.2}\,(1-x)^{\eta_-}\,(1 - x/x_0).$$

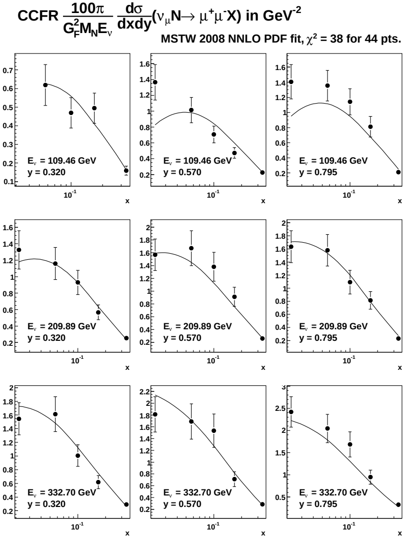
- x_0 fixed by zero strangeness: $\int_0^1 \mathrm{d}x\,\left[s(x, Q_0^2) - \bar{s}(x, Q_0^2)\right] = 0.$

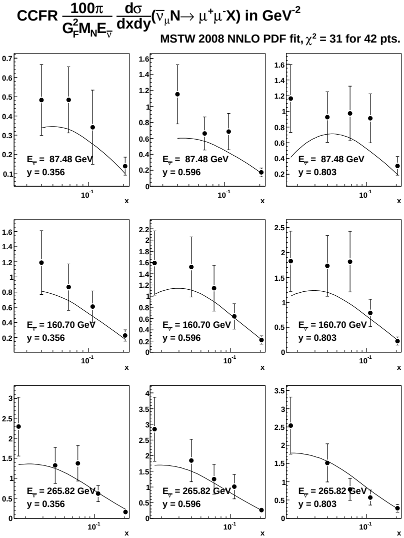
Description of NuTeV dimuon cross sections





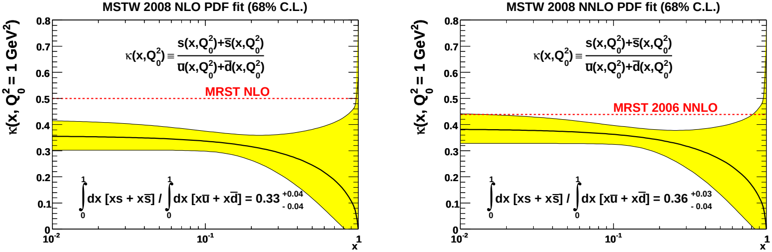
Description of CCFR dimuon cross sections



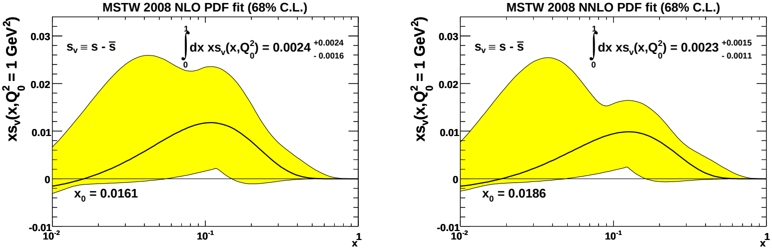


Strange quark and antiquark distributions

- Ratio of strange sea to non-strange sea: $(s+\bar{s})/(\bar{u}+\bar{d})$.



- Strange sea asymmetry: $xs - x\bar{s}$.



Resolutions to NuTeV $\sin^2\theta_W$ "anomaly"

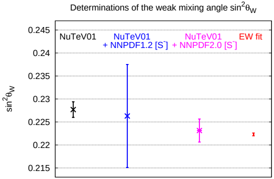
- NuTeV extraction [hep-ex/0110059] of $\sin^2\theta_W$ from

$$R^- \equiv \frac{\sigma(\nu_\mu N \to \nu_\mu X) - \sigma(\bar{\nu}_\mu N \to \bar{\nu}_\mu X)}{\sigma(\nu_\mu N \to \mu^- X) - \sigma(\bar{\nu}_\mu N \to \mu^+ X)} \approx \frac{1}{2} - \sin^2\theta_W$$

  was $\sim 3\sigma$ above the Standard Model prediction.

- New Physics? Or neglected PDF-related uncertainties, e.g.

1. **Strange sea asymmetry?** ($s \neq \bar{s}$)
   - MSTW $s - \bar{s}$ removes 1–2σ of NuTeV $\sin^2\theta_W$ "anomaly".
   - Unbiased parameterisation (NNPDF) removes **all** of it.



2. **Isospin violation?** ($u^p \neq d^n$, $d^p \neq u^n$)
   - Generated by QED corrections to parton evolution.
   - Found by MRST [hep-ph/0411040] to remove a little more than 1σ of the total discrepancy in the NuTeV $\sin^2\theta_W$.
3. **Nuclear corrections?** ($R^A_{u_v} \neq R^A_{d_v}$) [Eskola and Paukkunen, '06]

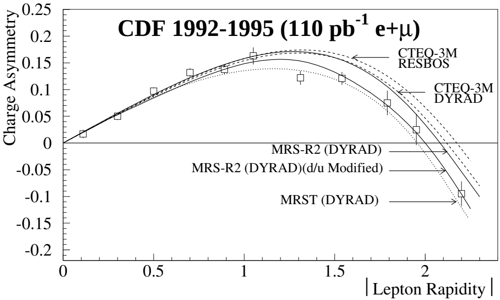
Nuclear modifications in the deuteron

- MSTW08 fit used F_2^d data from BCDMS, NMC, E665, SLAC.
- Deuteron corrections expected to be small (few % at most).
- Neglected by most proton PDF groups (e.g. CTEQ, NNPDF) and in most fits of nuclear PDFs (exception: HKN07).
- F_2^d data used in all MRST/MSTW fits were **corrected** for nuclear shadowing using a model [Badelek and Kwiecinski, '94]:
    - Calculated using a vector-meson-dominance mechanism.
    - Correction of 1% at $x = 0.01$ and $Q^2 = 4$ GeV^2.
    - **No** nuclear modification at medium/large x.
- Recently we re-examined deuteron corrections, prompted by apparent discrepancies with new precise Tevatron data on the $W \to \ell\nu$ charge asymmetry. **Preliminary** studies.

$W \to \ell\nu$ charge asymmetry from Tevatron Run I

$$A_W(y_W) = \frac{\mathrm{d}\sigma(W^+)/\mathrm{d}y_W - \mathrm{d}\sigma(W^-)/\mathrm{d}y_W}{\mathrm{d}\sigma(W^+)/\mathrm{d}y_W + \mathrm{d}\sigma(W^-)/\mathrm{d}y_W} \approx \frac{u(x_1)d(x_2) - d(x_1)u(x_2)}{u(x_1)d(x_2) + d(x_1)u(x_2)}$$

But **measure** $A_\ell(\eta_\ell) = \frac{\mathrm{d}\sigma(\ell^+)/\mathrm{d}\eta_\ell - \mathrm{d}\sigma(\ell^-)/\mathrm{d}\eta_\ell}{\mathrm{d}\sigma(\ell^+)/\mathrm{d}\eta_\ell + \mathrm{d}\sigma(\ell^-)/\mathrm{d}\eta_\ell}$, constrains d/u ratio.

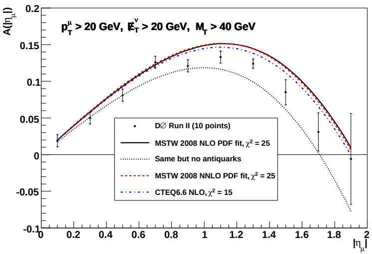


[hep-ex/9809001]

$W \to \ell\nu$ charge asymmetry from Tevatron Run II

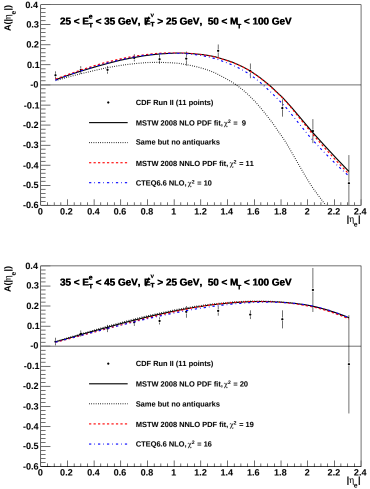
- Run II data in MSTW 2008 fit.
- Mainly constraint on **down** quark.
- Antiquarks important at low p_T^ℓ.

DØ data on lepton charge asymmetry from W→μν decays



[Data: arXiv:0709.4254]

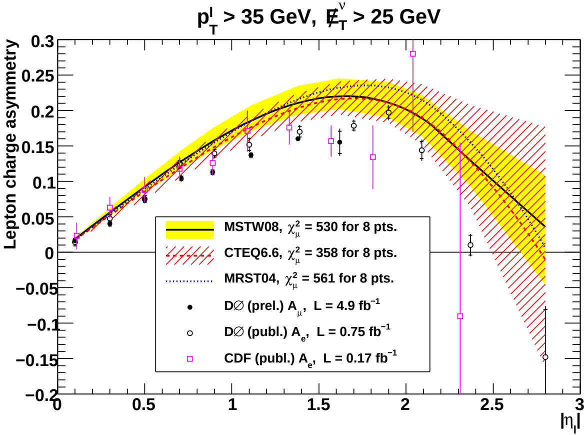
CDF data on lepton charge asymmetry from W→eν decays



[Data: hep-ex/0501023]

Latest DØ data on $W \to \ell\nu$ charge asymmetry

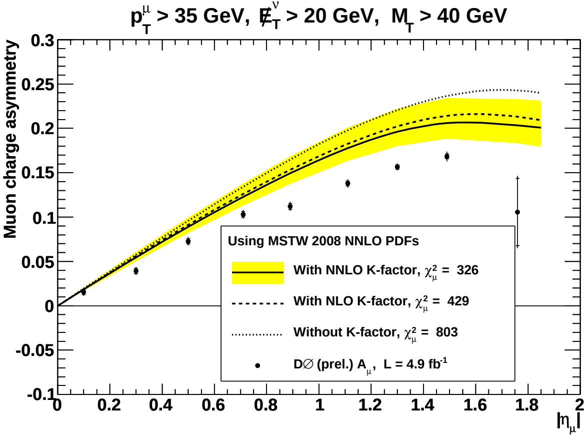
New DØ data: A_e [arXiv:0807.3367] and A_μ [DØ Note 5976-CONF]



- Problems describing new data at NLO, especially for $p_T^\ell > 35$ GeV.

Latest DØ data on $W \to \ell\nu$ charge asymmetry

- Effect of NNLO (or p_T^W-resummation, RESBOS) is small.

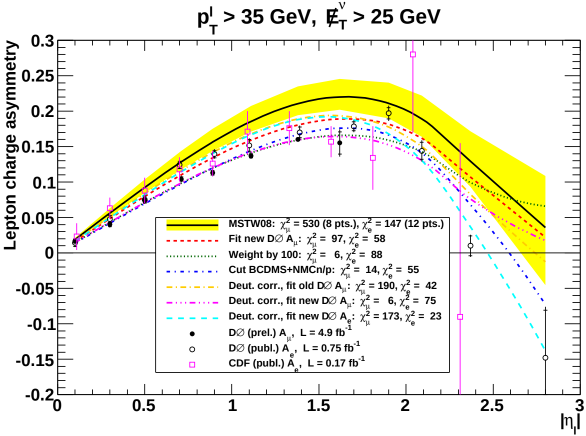


NNLO: Catani, Cieri, Ferrera, de Florian, Grazzini, arXiv:0903.2120

(Previous calculation: Melnikov, Petriello, hep-ph/0609070, FEWZ)

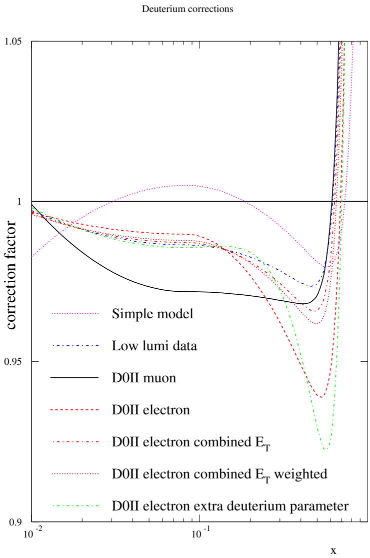
Latest DØ data on $W \to \ell\nu$ charge asymmetry

- Can the PDFs be refitted to describe the new data?



- Allowing free deuteron corrections aids description of DØ data.
- But not possible to describe both A_μ and A_e simultaneously.

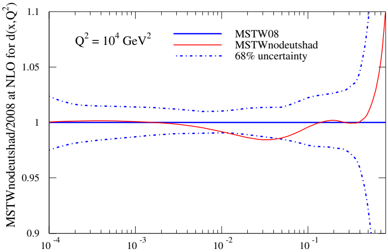
Deuteron corrections determined from global fits

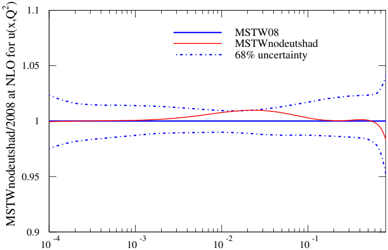


- **Parameterise** deuteron corrections as a smooth Q^2-independent function with 4 free parameters.
- Helps resolve tension between F_2^d data and Tevatron $W \to \ell\nu$ data.
- But corrections are **larger** and a different shape than theoretical expectations ("Simple model").
- Particularly large corrections needed for DØ $W \to \mu\nu$ charge asymmetry data.

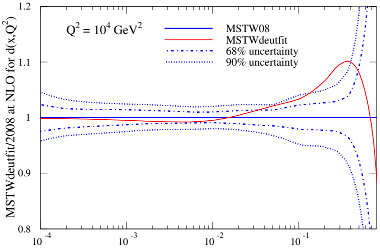
Effect of deuteron corrections on *d*- and *u*-quark PDFs

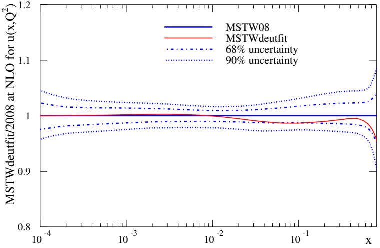
No deuteron shadowing:





Deuteron corrections fitted:





Summary

Status of proton PDFs (**global** fits in GM-VFNS)

- MSTW08/CTEQ6.6 at NLO, but only MSTW08 at NNLO.
- NNPDF a viable alternative **after** GM-VFNS implemented.

Rôle of nuclear corrections in global fits of proton PDFs

- Neutrino DIS data and deuteron structure functions are important to provide flavour separation of proton PDFs.
- Plenty of room for improvement in treatment of nuclear corrections (→ next talks by H. Paukkunen and I. Schienbein).
- Need reliable estimate of nuclear corrections (including deuteron), and propagation of uncertainties through to proton PDFs, to be used for precision LHC physics.