Parton Distributions

Robert Thorne

August 23rd, 2011

University College London

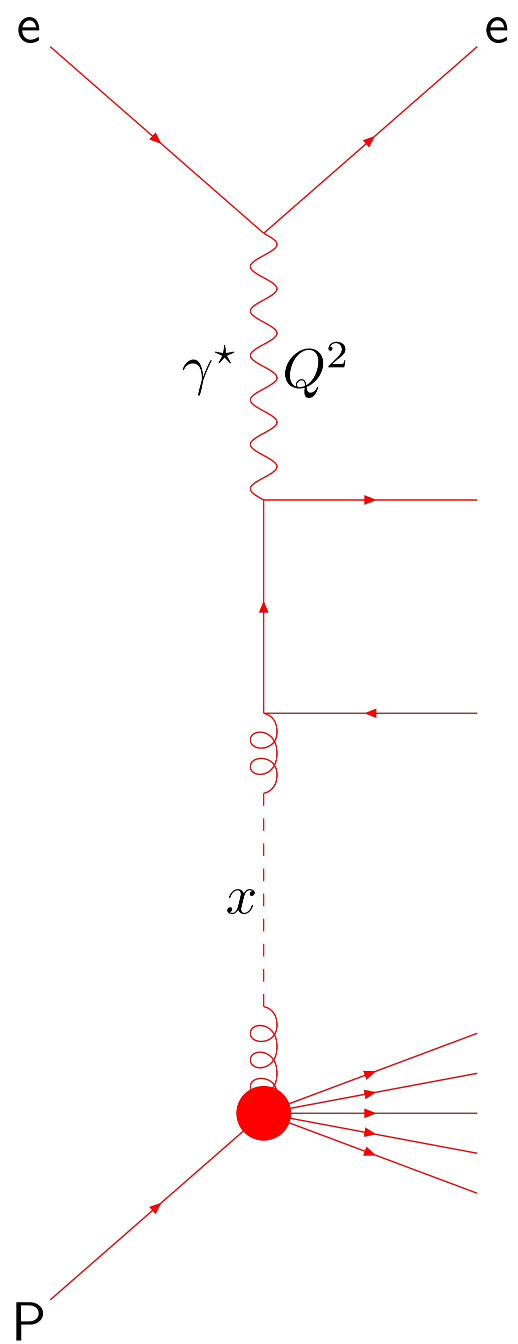
Strong force makes it difficult to perform analytic calculations of scattering processes involving hadronic particles.

The weakening of $\alpha_S(\mu^2)$ at higher scales $\longrightarrow$ the **Factorization Theorem**.

Hadron scattering with an electron factorizes.

Q^2 – Scale of scattering

$x = \frac{Q^2}{2m\nu}$ – Momentum fraction of Parton (ν=energy transfer)



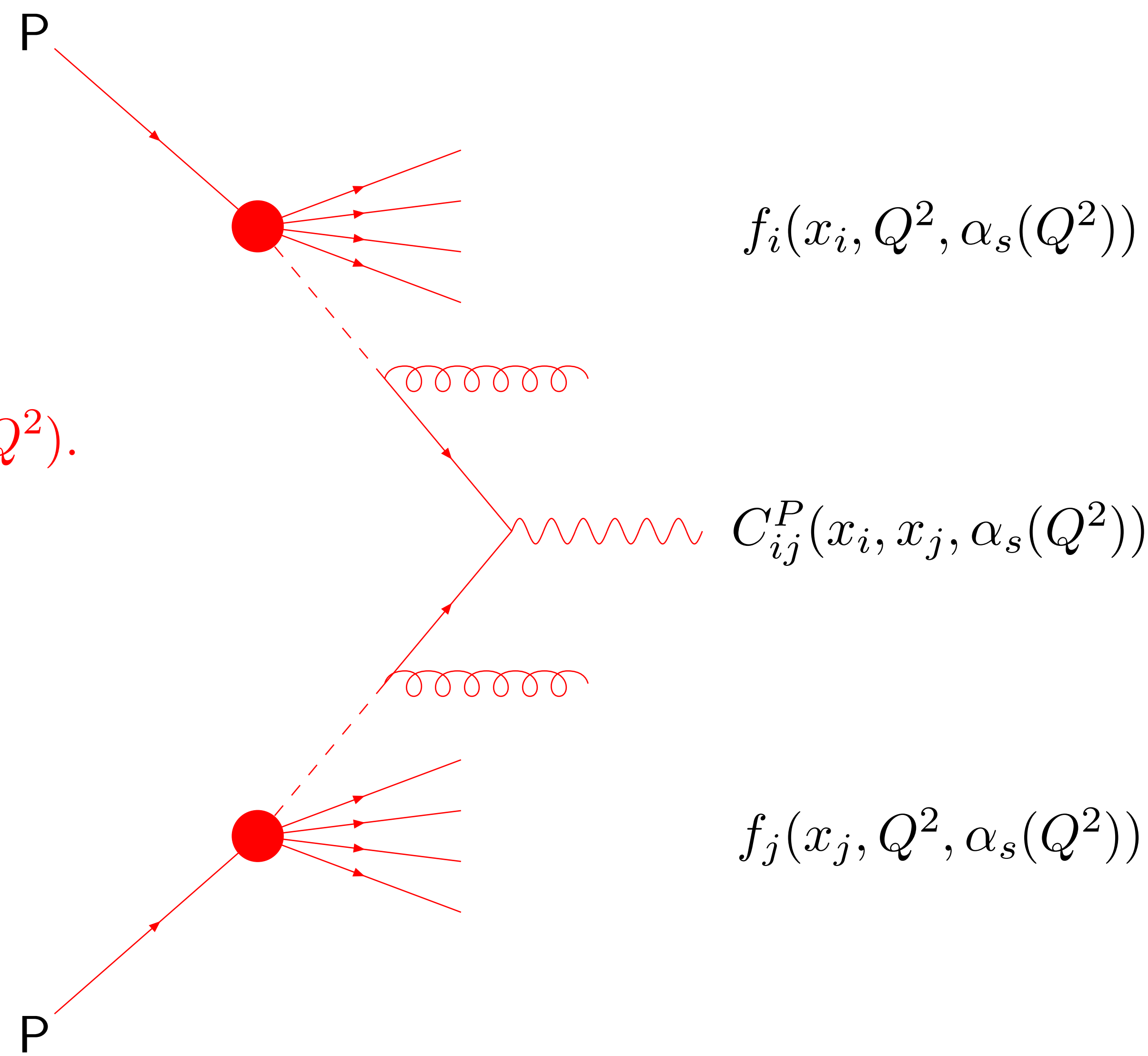
perturbative
calculable
coefficient function
$C_i^P(x, \alpha_s(Q^2))$

nonperturbative
incalculable
parton distribution
$f_i(x, Q^2, \alpha_s(Q^2))$

The coefficient functions $C_i^P(x, \alpha_s(Q^2))$ are process dependent (new physics) but are calculable as a power-series in $\alpha_s(Q^2)$.

$$C_i^P(x, \alpha_s(Q^2)) = \sum_k C_i^{P,k}(x)\alpha_s^k(Q^2).$$

Since the parton distributions $f_i(x, Q^2, \alpha_s(Q^2))$ are process-independent, i.e. **universal**, and evolution with scale is calculable, once they have been measured at one experiment, one can predict many other scattering processes.



Obtaining PDF sets – General procedure.

Start parton evolution at low scale $Q_0^2 \sim 1\mathrm{GeV}^2$. In principle 11 different partons to consider.

$$u, \bar{u}, \quad d, \bar{d}, \quad s, \bar{s}, \quad c, \bar{c}, \quad b, \bar{b}, \quad g$$

$m_c, m_b \gg \Lambda_{\mathrm{QCD}}$ so heavy parton distributions determined perturbatively. Leaves 7 independent combinations, or 6 if we assume $s = \bar{s}$ (just started not to).

$$u_V = u - \bar{u}, \quad d_V = d - \bar{d}, \quad \mathrm{sea} = 2 * (\bar{u} + \bar{d} + \bar{s}), \quad s + \bar{s} \quad \bar{d} - \bar{u}, \quad g.$$

Input partons parameterised as, e.g. MSTW, – much more general form for NNPDF, but same limits as $x \rightarrow, 0, 1$.

$$xf(x, Q_0^2) = (1-x)^\eta (1 + \epsilon x^{0.5} + \gamma x) x^\delta.$$

Evolve partons upwards using LO, NLO (or increasingly NNLO) DGLAP equations.

$$\frac{df_i(x, Q^2, \alpha_s(Q^2))}{d \ln Q^2} = \sum_j P_{ij}(x, \alpha_s(Q^2)) \otimes f_j(x, Q^2, \alpha_s(Q^2))$$

Fit data for scales above $2-5\mathrm{GeV}^2$. Need many different types for full determination.

- Lepton-proton collider HERA – (DIS) $\rightarrow$ small-x quarks (best below $x \sim 0.05$). Also gluons from evolution (same x), and now $F_L(x, Q^2)$. Also, jets $\rightarrow$ moderate-x gluon.Charged current data some limited info on flavour separation. Heavy flavour structure functions – gluon and charm, bottom distributions and masses.

- Fixed target DIS – higher x – leptons (BCDMS, NMC, ...) $\rightarrow$ up quark (proton) or down quark (deuterium) and neutrinos (CHORUS, NuTeV, CCFR) $\rightarrow$ valence or singlet combinations.

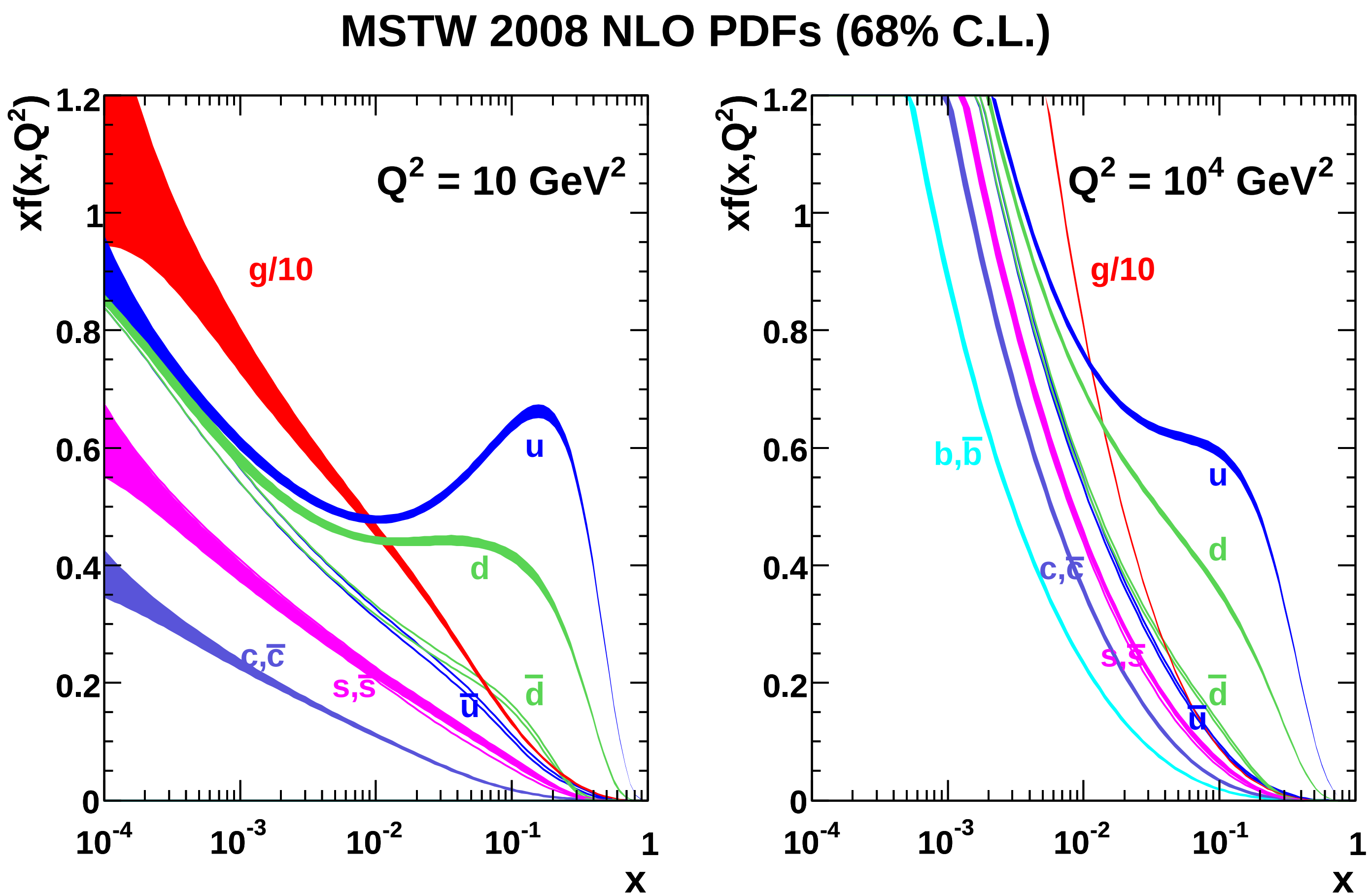
- Di-muon production in neutrino DIS – strange quarks and neutrino-antineutrino comparison $\rightarrow$ asymmetry . Only for $x > 0.01$.

- Drell-Yan production of dileptons – quark-antiquark annihilation (E605, E866) – high-x sea quarks. Deuterium target – $\bar{u}/\bar{d}$ asymmetry.

- High-p_T jets at colliders (Tevatron) – high-x gluon distribution – $x > 0.01$.

- W and Z production at colliders (Tevatron/LHC) – different quark contributions to DIS.

This procedure is generally successful and is part of a large-scale, ongoing project. Results in partons of the form shown.



Various choices of PDF – MSTW, CTEQ, NNPDF, AB(K)M, HERA, Jimenez-Delgado *et al etc.*. All LHC cross-sections rely on our understanding of these partons.

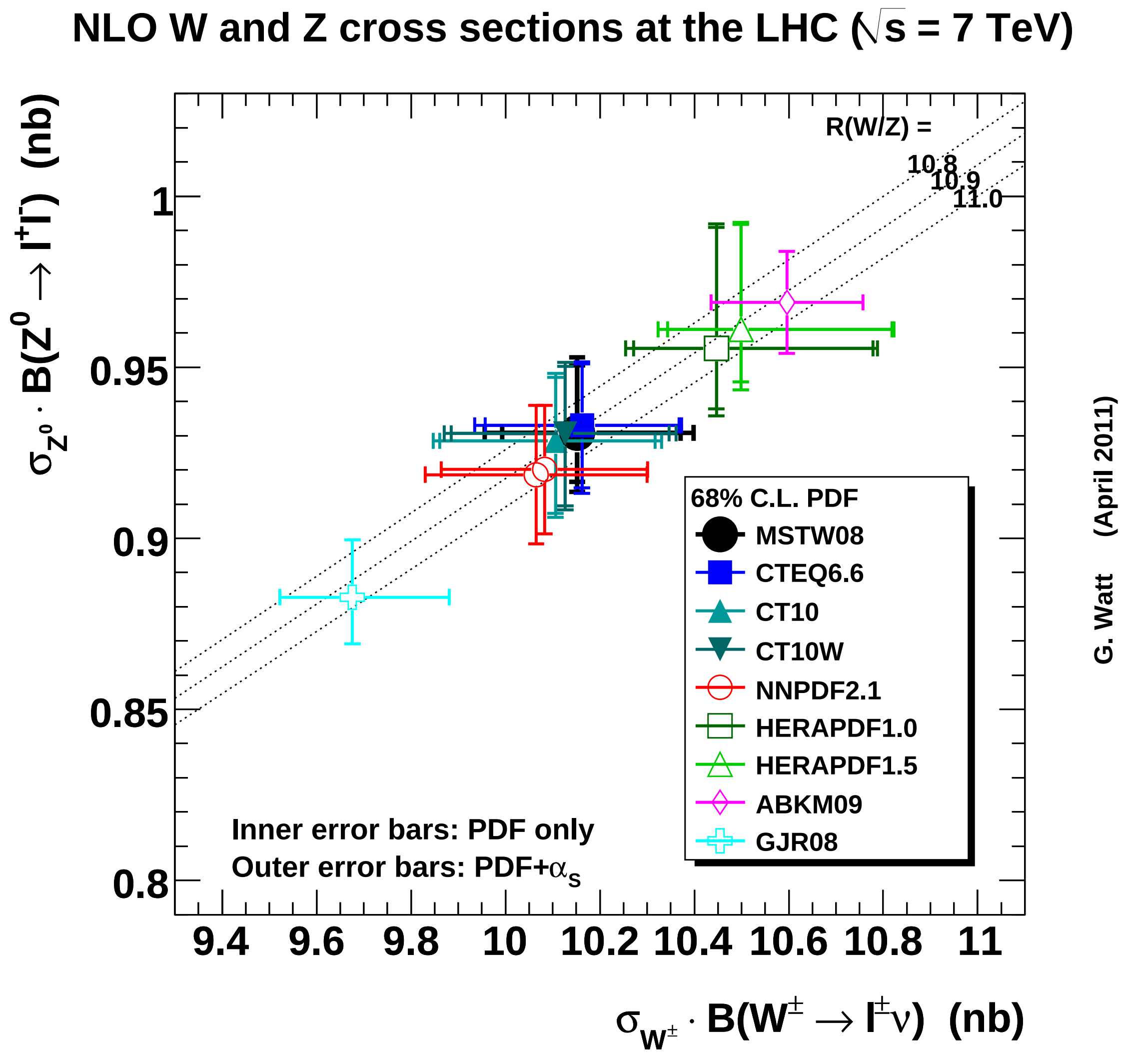
Predictions (Watt) for W and Z cross-sections for LHC with common NLO QCD and vector boson width effects, and common branching ratios, and at $7\mathrm{TeV}$.

Good agreement at NLO for variety of PDFs.

In fact comparing all groups get significant discrepancies between them even for this benchmark process.

Can understand some of the systematic differences.

Total W, Z total cross-sections best-case scenario – rapidities show more variation.



Sources of Variations/Uncertainty

It is vital to consider theoretical/assumption-dependent uncertainties:

- Methods of determining "best fit" and uncertainties.
- Underlying assumptions in procedure, e.g. parameterisations and data used.
- Treatment of heavy flavours.
- PDF and α_S correlations.

Responsible for differences between groups for extraction of fixed-order PDFs.

Different PDF sets

- MSTW08 – fit all previous types of data. Most up-to-date Tevatron jet data. Not most recent HERA combination of data. PDFs at LO, NLO and NNLO.

- CT10 – very similar. PDFs at NLO. CT10 include HERA combination and more Tevatron data though also run I jet data. Not large changes from CTEQ6.6. CT10W gives higher weight to Tevatron asymmetry data.

- NNPDF2.1 – include all except HERA jet data (not strong constraint). NNPDF2.1 improves on NNPDF2.0 by better heavy flavour treatment. PDFs at NLO and very recently NNLO and LO .

- HERAPDF1.0 – based on HERA inclusive structure functions, neutral and charged current. Use combined data. PDFs at NLO and (without uncertainties) NNLO.

- ABKM09 – fit to DIS and fixed target Drell-Yan data. PDFs at NLO and NNLO. Less conservative cuts at low W^2 than other groups – fit for higher twist corrections rather than attempt to avoid them.

- GJR08 – fit to DIS, fixed target Drell-Yan and Tevatron jet data (not at NNLO). PDFs at NLO and NNLO.

Various groups have provided preliminary updates or illustrations of variations due to inclusion of new data. Includes ...

HERAPDF have *preliminary* version HERAPDF1.5 with grids available at NLO and NNLO, both with uncertainties. However, based on as yet unpublished combined run II data and no official publication. Also versions 1.6 and 1.7 including combinations including HERA jet data, prelim. combined charm data, lower beam energy data.

MSTW have prelim. sets fit to combined HERA data, and looking at deuterium corrections – in DIS proceedings.

ABM have versions including combined HERA data and including a variety of Tevatron jet data sets – again see DIS proceedings.

Lots of other reports, e.g. sets for fits to Collider data only (NNPDF, MSTW), ...

Parton Fits and Uncertainties. Two main approaches.

Parton parameterization and Hessian (Error Matrix) approach first used by H1 and ZEUS, and extended by CTEQ.

$$\chi^2 - \chi^2_{min} \equiv \Delta\chi^2 = \sum_{i,j} H_{ij}(a_i - a_i^{(0)})(a_j - a_j^{(0)})$$

The Hessian matrix H is related to the covariance matrix of the parameters by

$$C_{ij}(a) = \Delta\chi^2 (H^{-1})_{ij}.$$

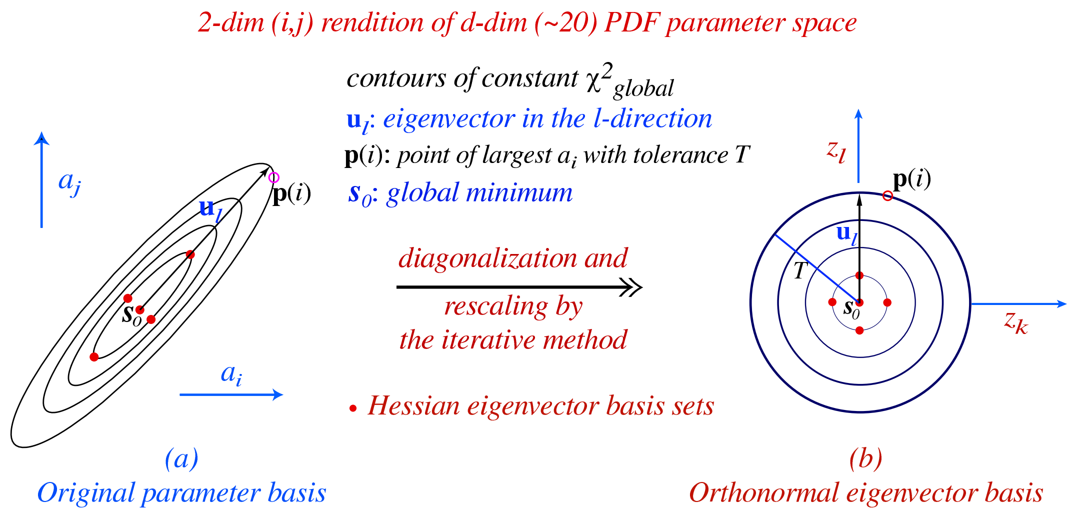
We can then use the standard formula for linear error propagation.

$$(\Delta F)^2 = \Delta\chi^2 \sum_{i,j} \frac{\partial F}{\partial a_i}(H)^{-1}_{ij}\frac{\partial F}{\partial a_j},$$

This is now the most common approach. Basis of e.g ABKM, GJR, where correlations are maintained in sets.

Can find and rescale eigenvectors of H leading to diagonal form

$$\Delta\chi^2 = \sum_i z_i^2$$



Implemented by CTEQ, then MRST/MSTW, HERAPDF. Uncertainty on physical quantity then given by

$$(\Delta F)^2 = \sum_i \left(F(S_i^{(+)}) - F(S_i^{(-)})\right)^2,$$

where $S_i^{(+)}$ and $S_i^{(-)}$ are PDF sets displaced along eigenvector direction.

Must choose "correct" $\Delta\chi^2$ given complication of errors in full fit and sometimes conflicting data sets.

Determination of best fit and uncertainties

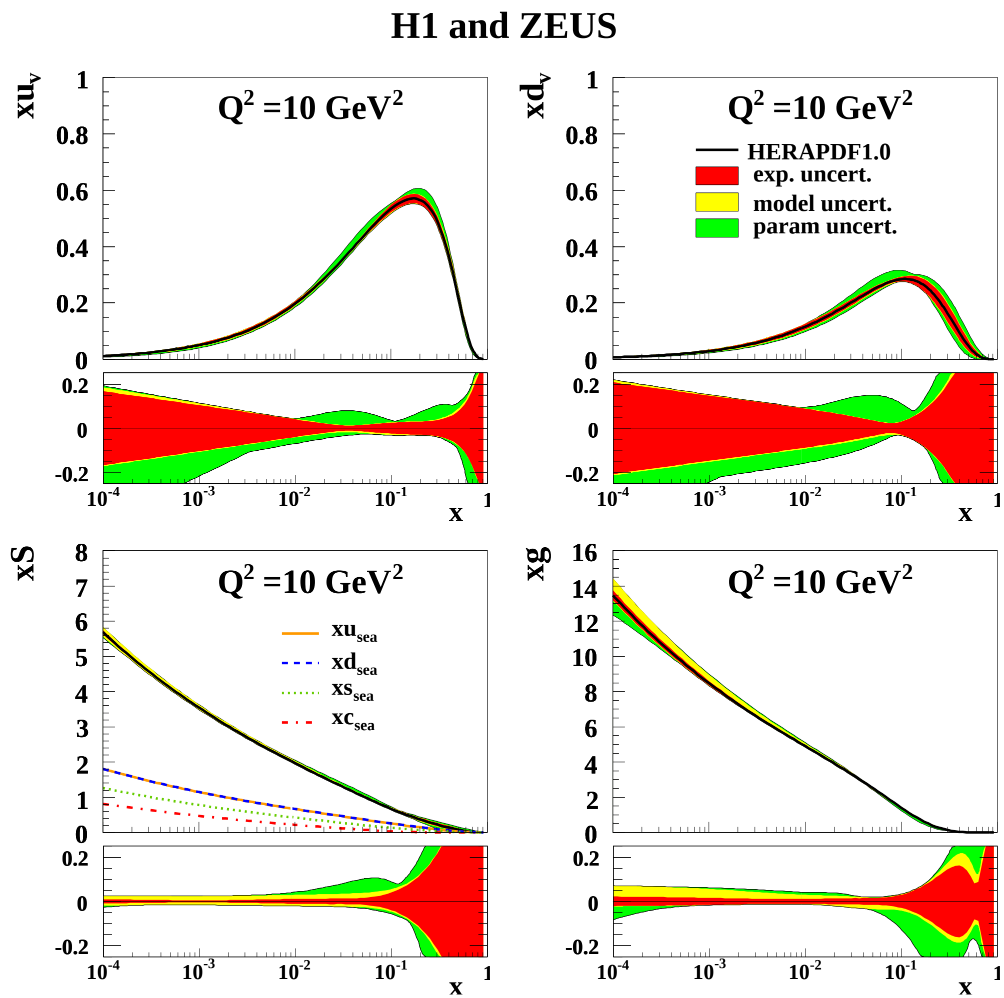
All but NNPDF minimise χ^2 and expand about best fit.

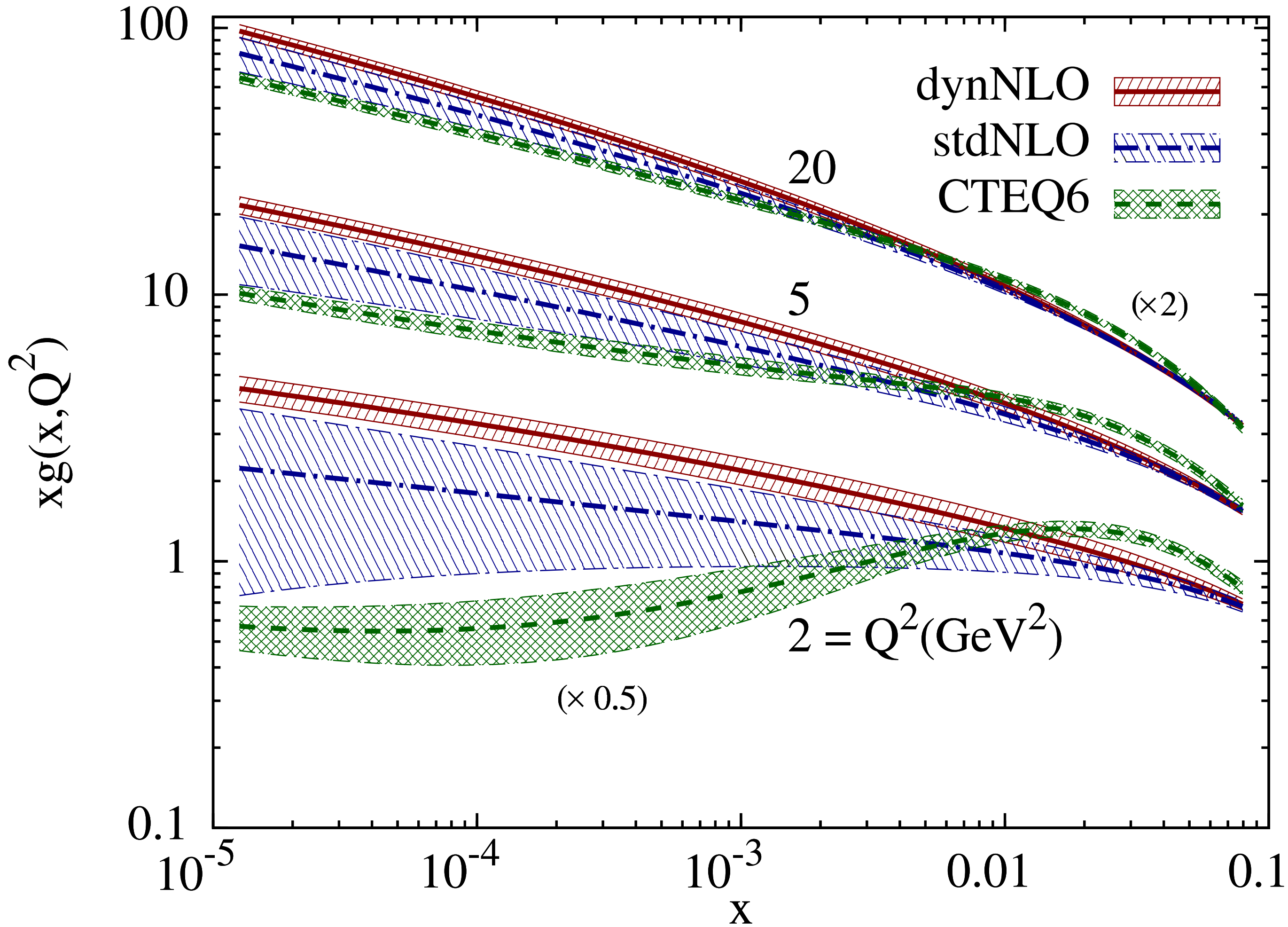
- MSTW08 – 28 parameters, 20 eigenvectors. Due to incompatibility of different sets and (perhaps to some extent) parameterisation inflexibility (little direct evidence for this) have inflated $\Delta\chi^2$ of $5-20$ for eigenvectors.
- CT10 – 26 eigenvectors, and some fixed parameters. Inflated $\Delta\chi^2$ of ~ 40 for 1-sigma for eigenvectors.
- HERAPDF2.0 – 10 eigenvectors. Use "$\Delta\chi^2 = 1$". Additional model and parameterisation uncertainties.
- ABKM09 – 21 parton parameters. Use $\Delta\chi^2 = 1$. Also α_S, m_c, m_b.
- GJR08 – 20 parton parameters (8 fixed for uncertainty) and α_S. Use $\Delta\chi^2 \approx 20$. Impose strong constraint on input form of PDFs.

Perhaps surprisingly all get rather similar uncertainties for PDFs cross-sections, though don't all mean the same.

Illustration of the HERAPDF1.0 parameterisation uncertainty, though start with fewer parameters.

Also model uncertainty from variation of starting scale Q_0^2, strange fraction at input, quark masses and Q^2 cuts.





The effect of the GJR dynamical generation of the gluon PDF via evolution from a valence-like form at very low scale $Q^2 = 0.5\mathrm{GeV}^2$, compared to the corresponding "standard" PDFs from a starting scale of $Q_0^2 = 2\mathrm{GeV}^2$, (and CTEQ6 - which is valence-like at $Q_0^2 = 1.69\mathrm{GeV}^2$).

Neural Network group (Ball *et al.*) limit parameterization dependence. Leads to alternative approach to "best fit" and uncertainties.

First part of approach, no longer perturb about best fit. Construct a set of Monte Carlo replicas $F_{i,p}^{art,k}$ of the original data set $F_{i,p}^{exp,(k)}$.

- REPLICAS FLUCTUATE ABOUT CENTRAL DATA:
  $F_{i,p}^{(art)(k)} = S_{p,N}^{(k)} F_{i,p}^{\mathrm{exp}} \left(1 + r_p^{(k)} \sigma_p^{\mathrm{stat}} + \sum_{j=1}^{N_{\mathrm{sys}}} r_{p,j}^{(k)} \sigma_{p,j}^{\mathrm{sys}}\right)$

Where $r_p^{(k)}$ are random numbers following Gaussian distribution, and $S_{p,N}^{(k)}$ is the analogous normalization shift of the of the replica depending on $1 + r_{p,n}^{(k)}\sigma_p^{norm}$. Hence, include information about measurements and errors in distribution of $F_{i,p}^{art,(k)}$.

Fit to the data replicas obtaining PDF replicas $q_i^{(net)(k)}$ (follows Giele *et al.*)

Mean μ_O and deviation σ_O of observable O then given by

$$\mu_O = \frac{1}{N_{rep}} \sum_1^{N_{rep}} O[q_i^{(net)(k)}], \quad \sigma_O^2 = \frac{1}{N_{rep}} \sum_1^{N_{rep}} (O[q_i^{(net)(k)}] - \mu_O)^2.$$

Eliminates parameterisation dependence by using a neural net which undergoes a series of (mutations via genetic algorithm) to find the best fit. In effect is a much larger sets of parameters – ~ 37 per distribution.

However, does include pre-processing exponents as $x \to 1$ and $x \to 0$ to aid convergence of fit,

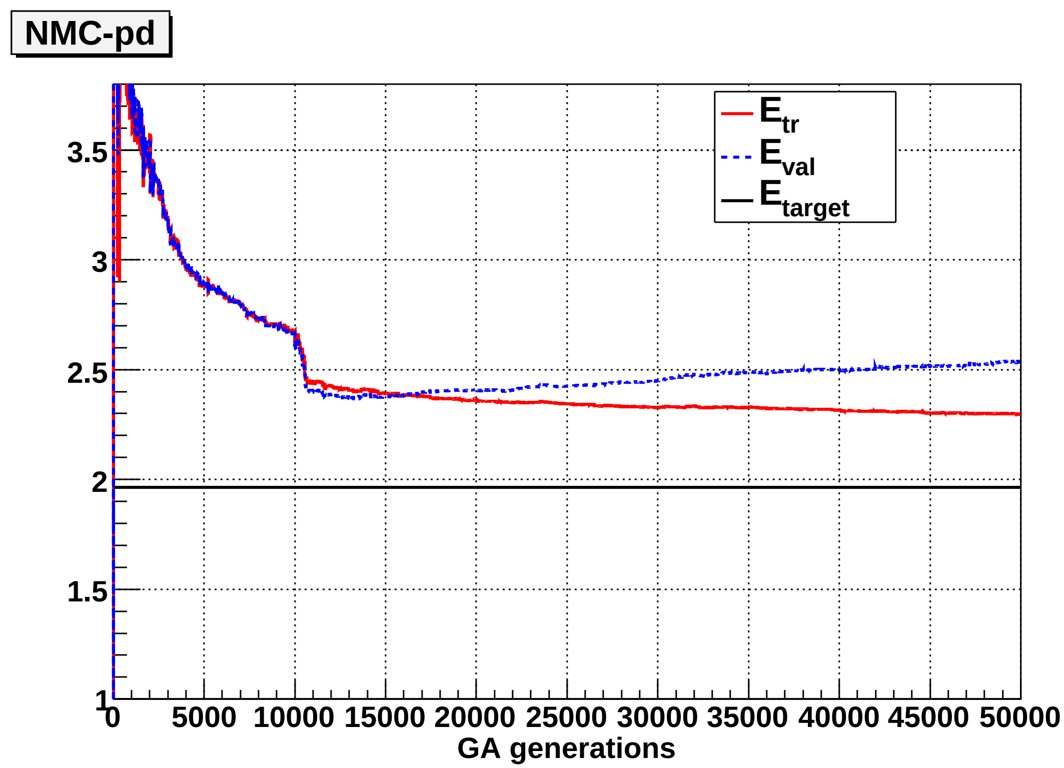
$$f(x, Q_0^2) = A(1-x)^m x^{-n} NN(x)$$

where n, m are in fairly narrow ranges, so overall behaviour guided at these extremes where data constraints vanish.

Split data sets randomly into equal size *training* and *validation* sets.

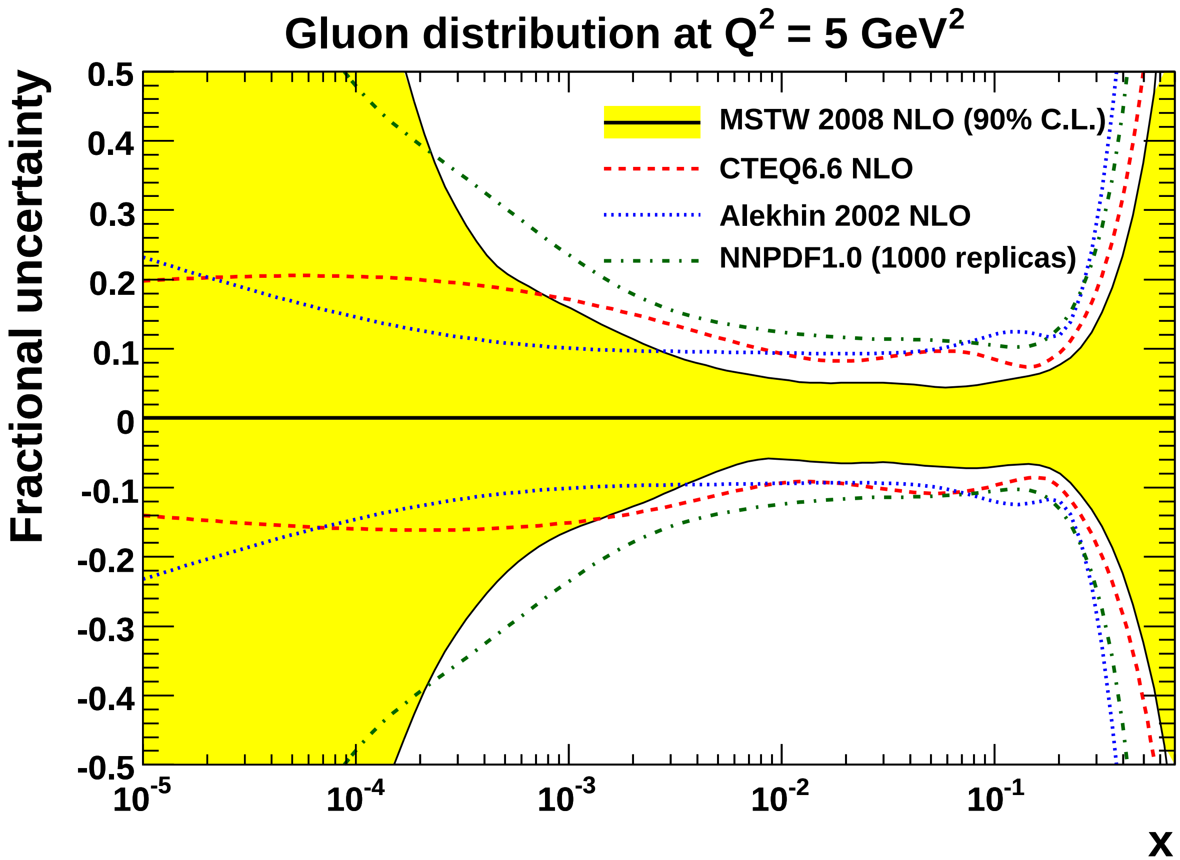
Fit until quality of fit to validation set starts to go up, even though training set still (hopefully slowly) improving.

Criterion for stopping the fit depends on different data sets.



Uncertainty has depended on stopping criteria.

Parameterisations - for the gluon at small x different parameterisations lead to very different uncertainty for small x gluon.



Most assume single power x^{λ} at input $\rightarrow$ limited uncertainty. If input at low Q^2 λ positive and small-x input gluon *fine-tuned* to ~ 0. Artificially small uncertainty.

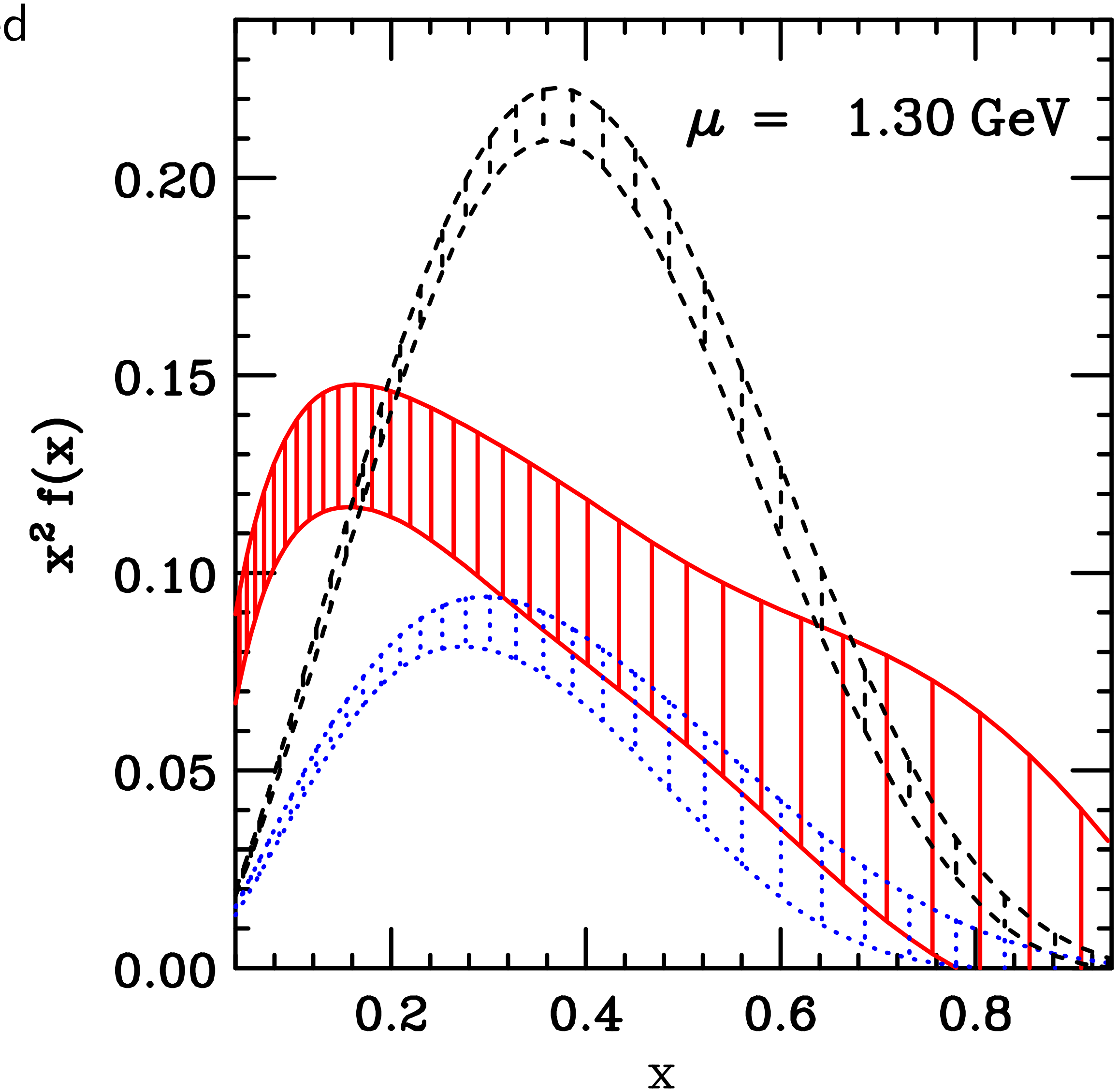
If $g(x) \propto x^{\lambda \pm \Delta\lambda}$ then $\Delta g(x) = \Delta\lambda \ln(1/x) * g(x)$.

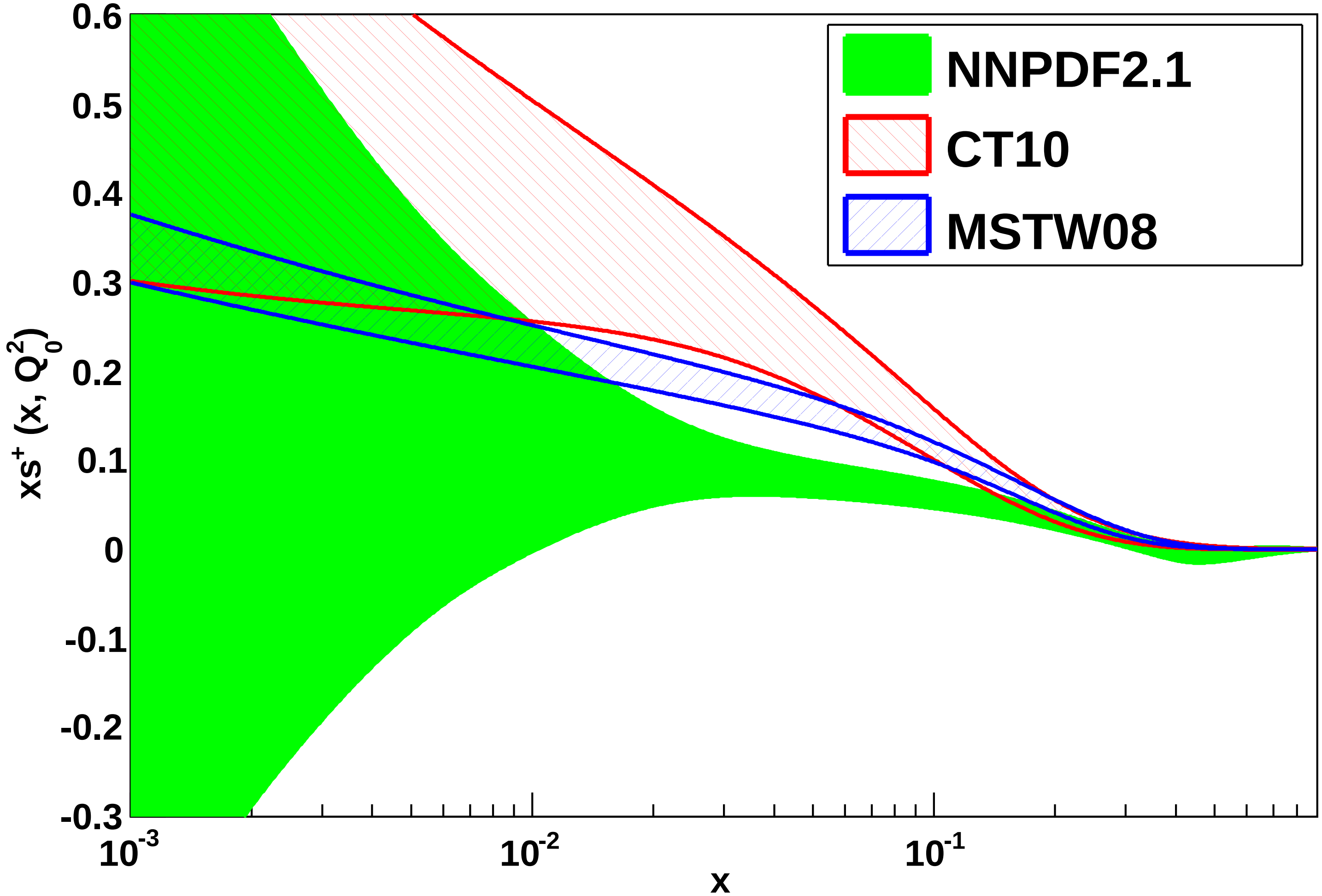
MSTW and NNPDF more flexible (can be negative) $\rightarrow$ rapid expansion of uncertainty where data runs out. CT10)HERAPDF1.5f) more flexible than previous versions.

Generally high-x PDFs parameterised so will behave like $(1-x)^\eta$ as $x \to 1$. More flexibility in CTEQ.

Very hard high-x gluon distribution (more-so even than NNPDF uncertainties).

However, is gluon, which is radiated from quarks, harder than the up valence distribution for $x \to 1$?





MSTW has theory assumption on strange at small x, CT10 less strong and NNPDF fully flexible.

Variation near $x = 0.05$ where data exists likely due to heavy flavour definitions/nuclear corrections.

Heavy Quarks – Essential to treat these correctly. Two distinct regimes:

Near threshold $Q^2 \sim m_H^2$ massive quarks not partons. Created in final state. Described using **Fixed Flavour Number Scheme** (FFNS), known fully to NLO.

$$F(x,Q^2) = C_k^{FF}(Q^2/m_H^2) \otimes f_k^{n_f}(Q^2)$$

Does not sum $\ln^n(Q^2/m_H^2)$ terms, and not calculated for many processes beyond LO. Used by AB(K)M and (G)JR. Sometimes final state details in this scheme only.

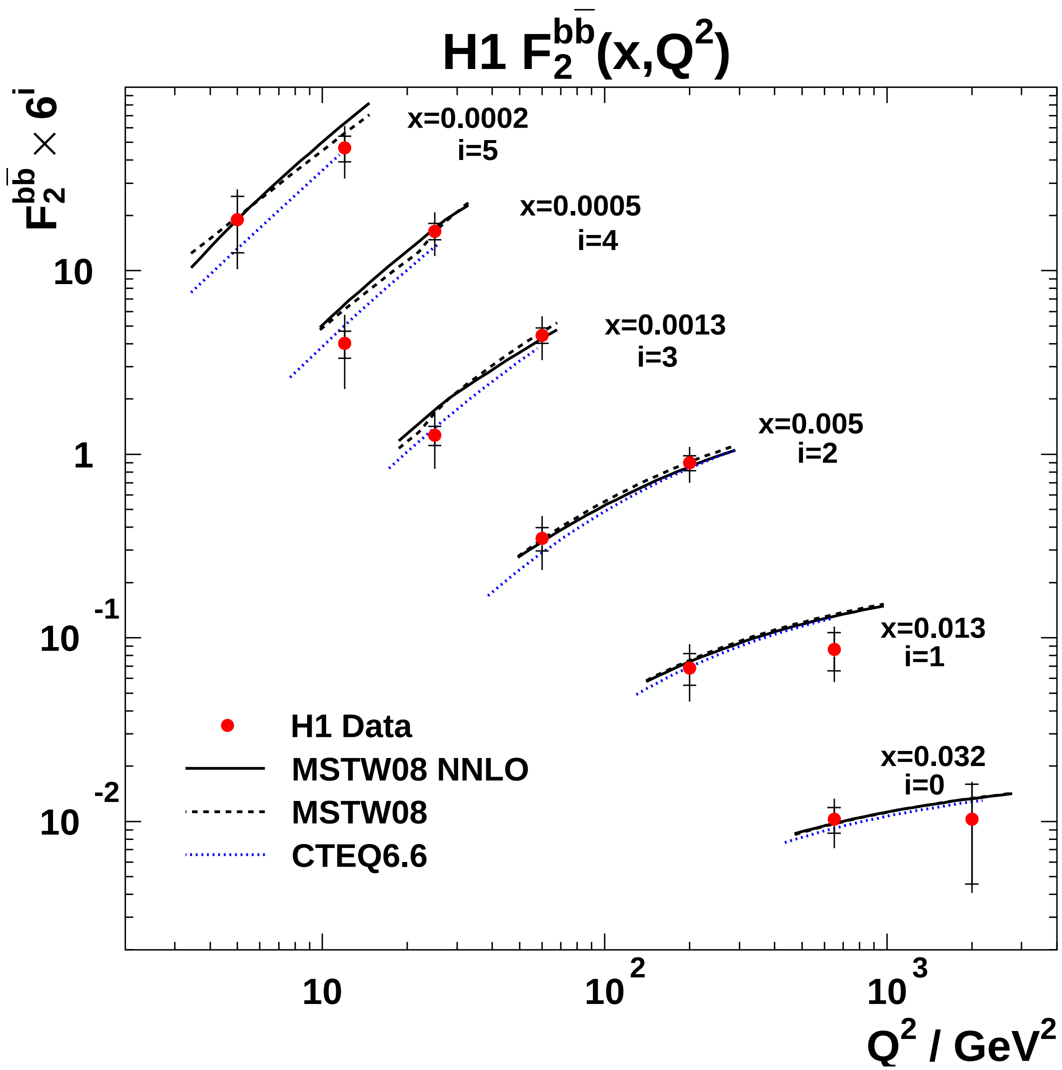
Alternative, at high scales $Q^2 \gg m_H^2$ heavy quarks like massless partons. Behave like up, down, strange. Sum $\ln(Q^2/m_H^2)$ terms via evolution. **Zero Mass Variable Flavour Number Scheme** (ZM-VFNS). Normal assumption in calculations. Ignores $\mathcal{O}(m_H^2/Q^2)$ corrections. No longer used.

$$F(x,Q^2) = C_j^{ZMVF} \otimes f_j^{n_f+1}(Q^2).$$

Advocate a **General Mass Variable Flavour Number Scheme** (GM-VFNS) interpolating between the two well-defined limits of $Q^2 \leq m_H^2$ and $Q^2 \gg m_H^2$. Used by MRST/MSTW and more recently (as default) by CTEQ, and now also by HERAPDF and NNPDF.

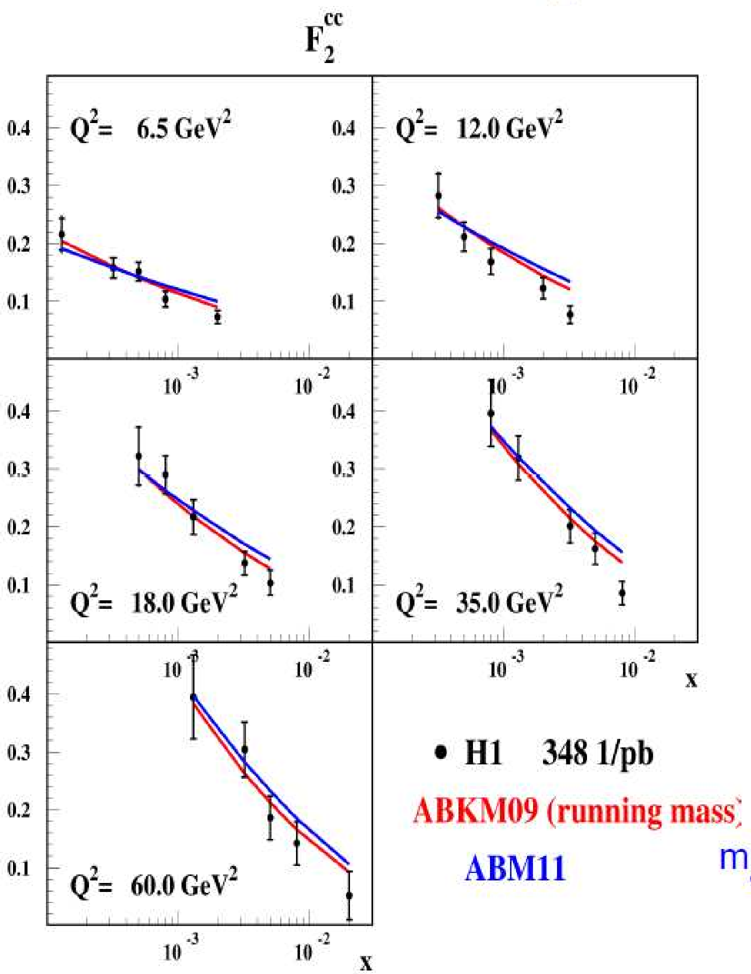
Various definitions possible. Versions used by MSTW (RT) and CTEQ (ACOT) have converged somewhat.

Various significant differences still exist as illustrated by comparison to most recent H1 data on bottom production.



ABM have improved the FFNS with an NNLO approx. and also looked at using the $\overline{MS}$ definition of m_c, with some advantages (Alekhin - PDF4LHC July).

New H1 data on F_2^{cc} and running-mass scheme



H1 Collaboration [hep-ex 1106.1028]

The choice of μ_R=m_h is close to the data kinematic → better perturbative convergence and reduced scale dependence

Laengenfeld, Moch, Uwer PRD 80, 054009 (2009)

The heavy-quark electroproductoin in the approximate NNLO (full NLO + NNLO threshold resummation)

sa, Moch PLB 699, 345 (2011

ABKM09 (running mass):

$m_c(m_c)$=1.18±0.06 GeV (incl.F_2 +PDG)

ABM11: (prel.)

$m_c(m_c)$=1.10±0.04 GeV (incl.F_2 + F_2^{cc})

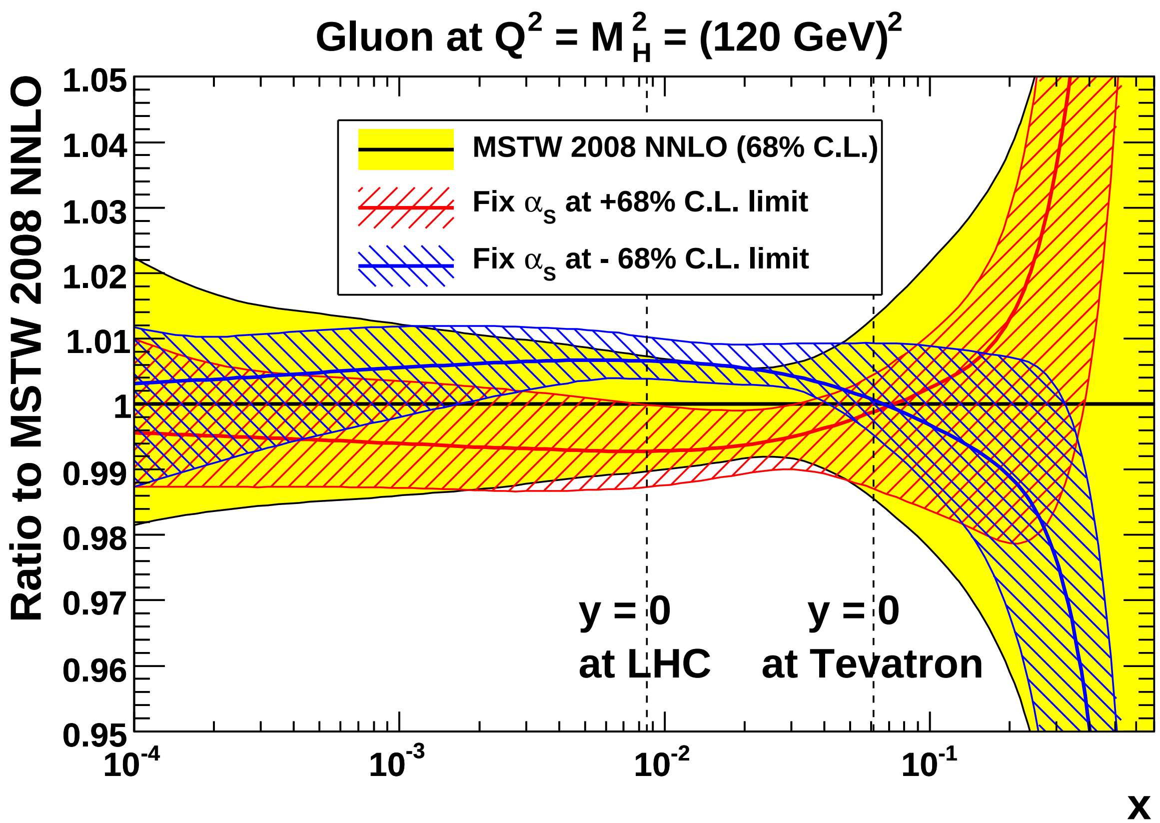
$m_c(m_c)$=1.14±0.04 GeV (incl.F_2 + F_2^{cc} + PDG)

$m_c(m_c)$=1.27±0.08 GeV (PDG '10)

PDF correlation with α_S.

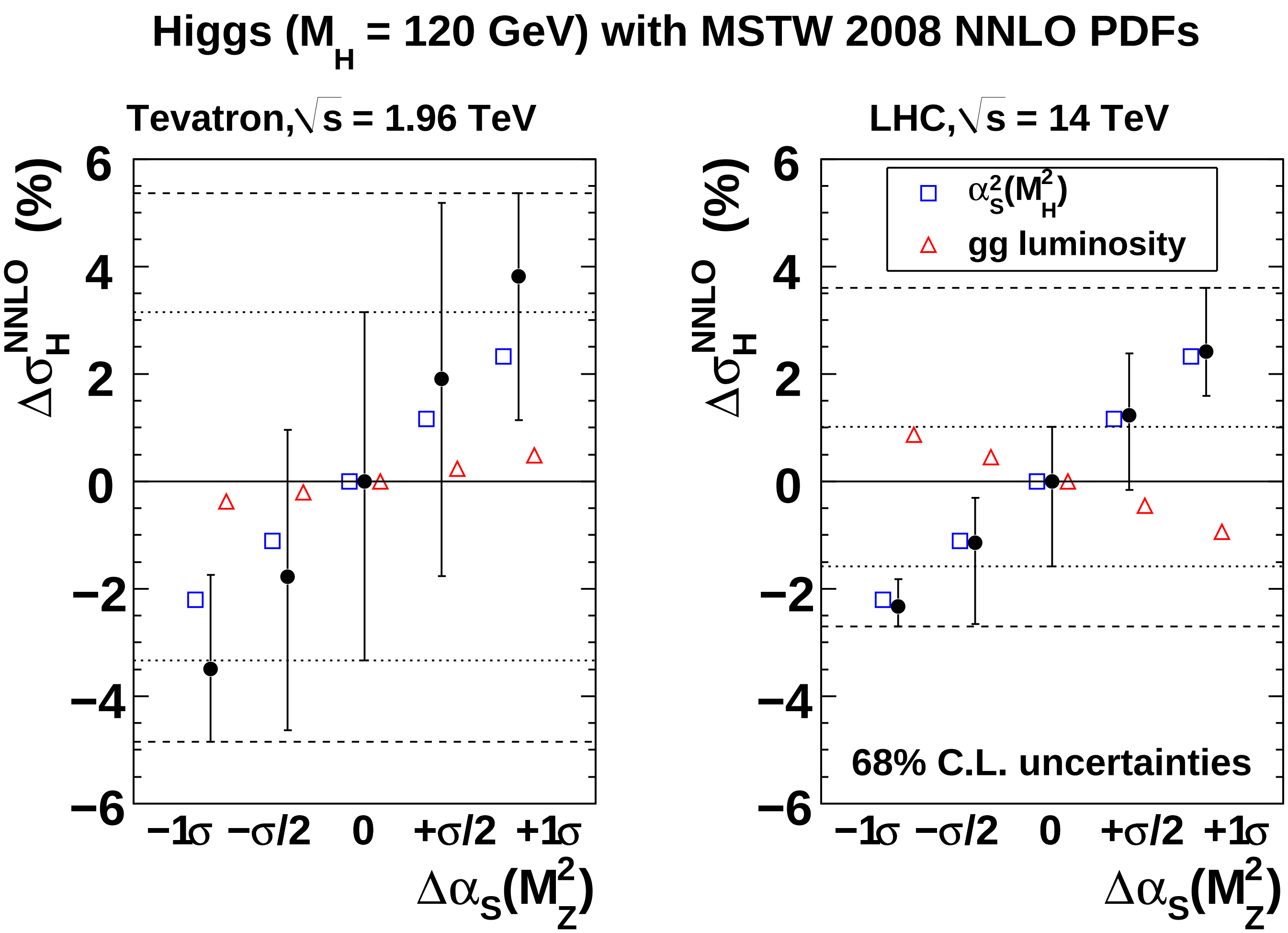
Can also look at PDF changes and uncertainties at different $\alpha_S(M_Z^2)$. Fully included (difficult to disentangle) in ABKM, (G)JR), but often only for one fixed $\alpha_S(M_Z^2)$.

MSTW produce sets for limits of α_S uncertainty – PDF uncertainties reduced since quality of fit already worse than best fit.



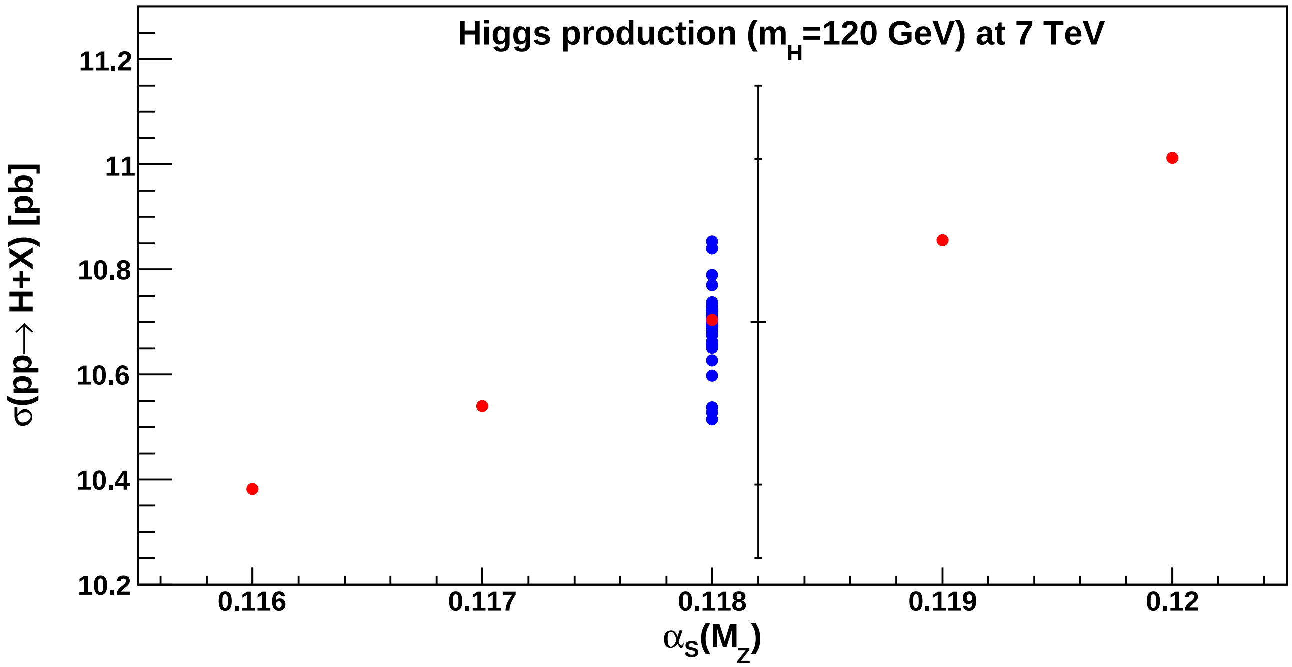
Expected gluon–$\alpha_S(M_Z^2)$ small–x anti-correlation $\rightarrow$ high-x correlation from sum rule.

NNLO predictions for Higgs $(120\mathrm{GeV})$ production for different allowed $\alpha_S(M_Z^2)$ values and their uncertainties.



Increases by a factor of $2-3$ (up more than down) at LHC. Direct $\alpha_S(M_Z^2)$ dependence mitigated somewhat by anti-correlated small-x gluon (asymmetry feature of *minor* problems in fit to HERA data). At Tevatron intrinsic gluon uncertainty dominates.

CTEQ have shown that up to Gaussian approx. for uncertainties (and some other caveats) α_S uncertainty accounted for by adding deviation from PDFs with upper and lower α_S limits (red) in quadrature with all other PDF eigenvectors (blue), seen below.

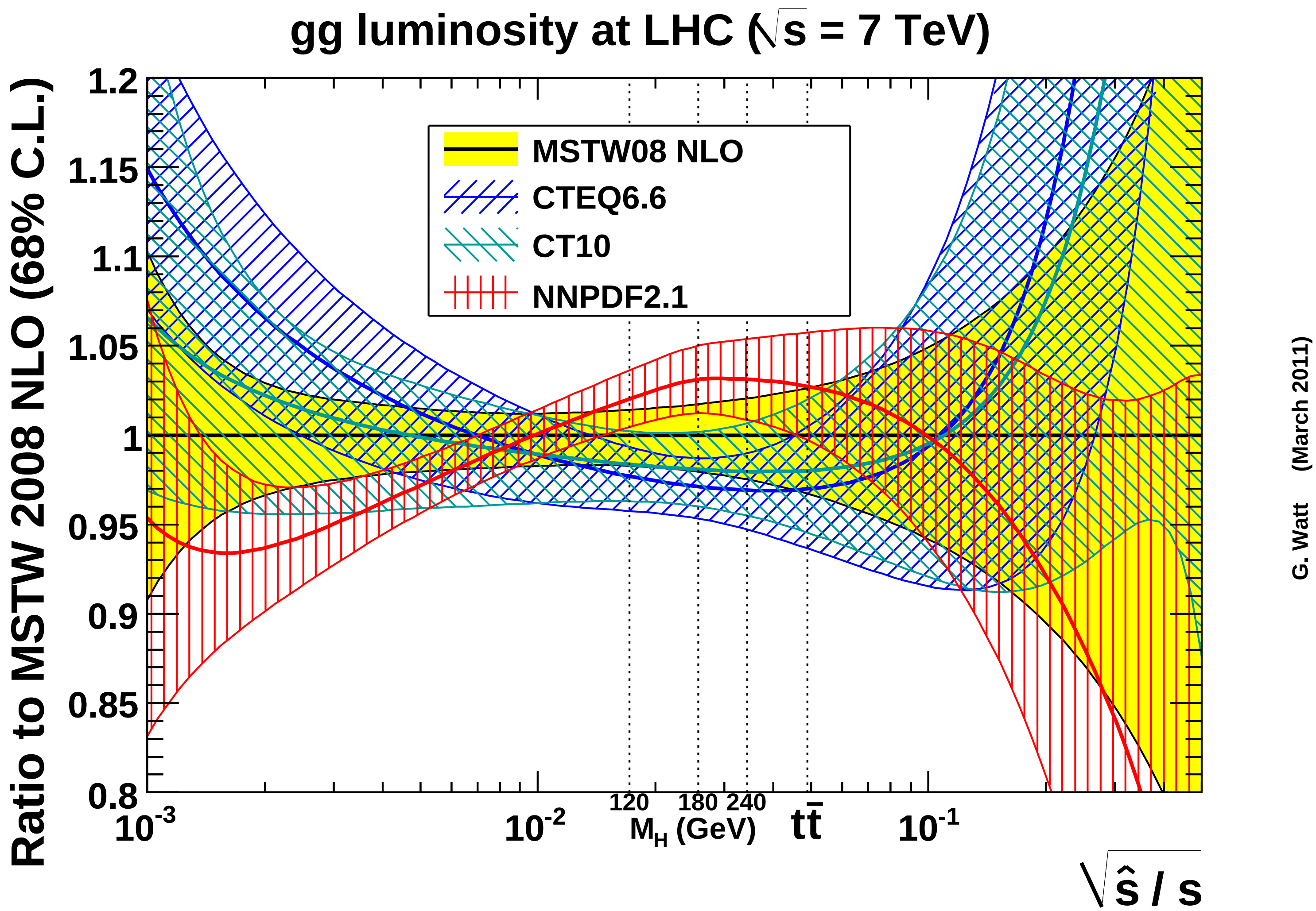


NNPDF advocate distributing PDF replicas according to probability of $\alpha_S(m_Z^2)$ taking that value based on some assumed central value and uncertainty, i.e.

$$N_{\rm rep}^{\alpha_S} \propto \exp\left(-\frac{(\alpha_S - \alpha_S^{(0)})^2}{2(\delta\alpha_S^{(68)})^2}\right),$$

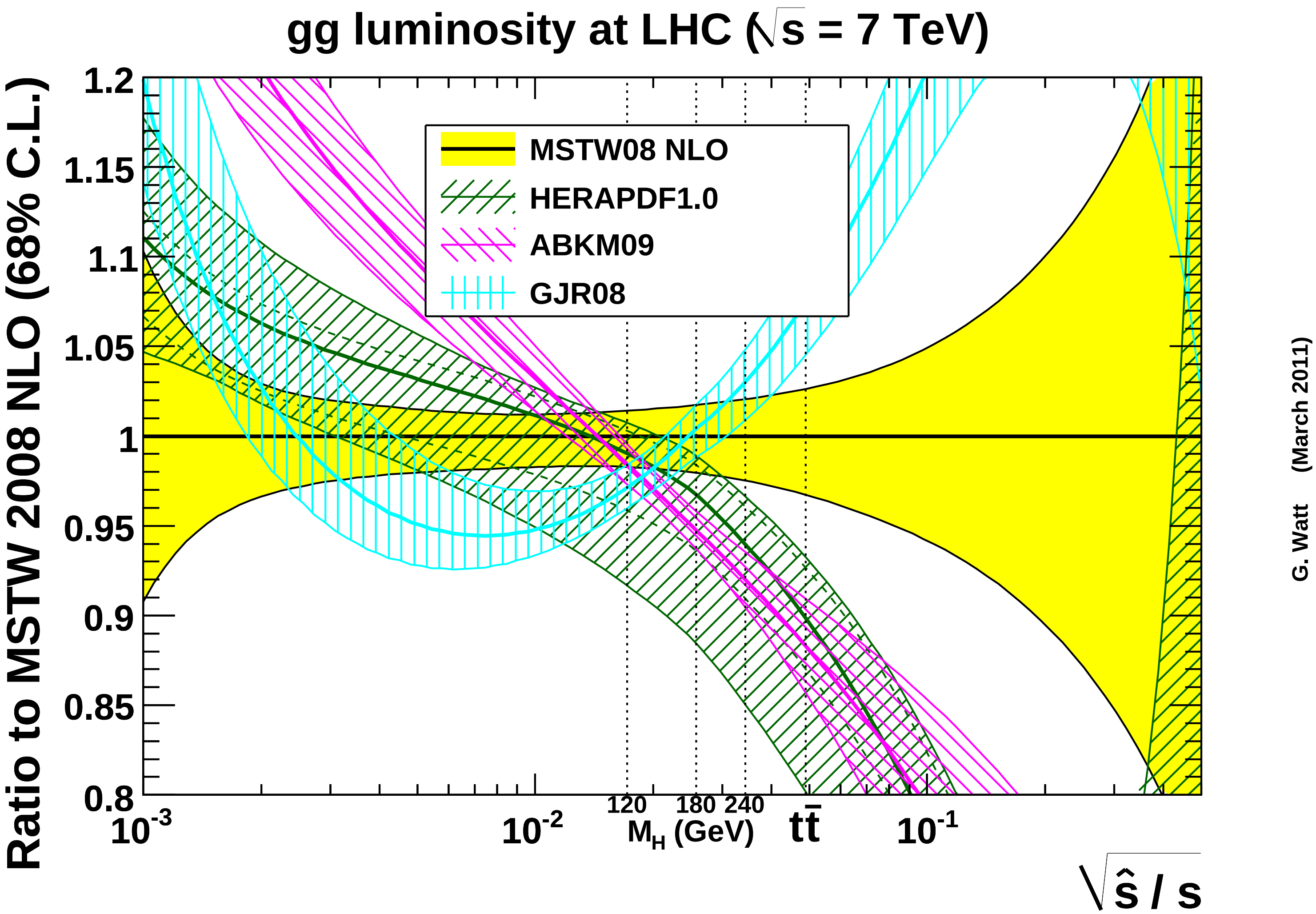
All lead to roughly same results Vicini *et al*.

Predictions by various groups - parton luminosities – NLO. Plots by G. Watt.



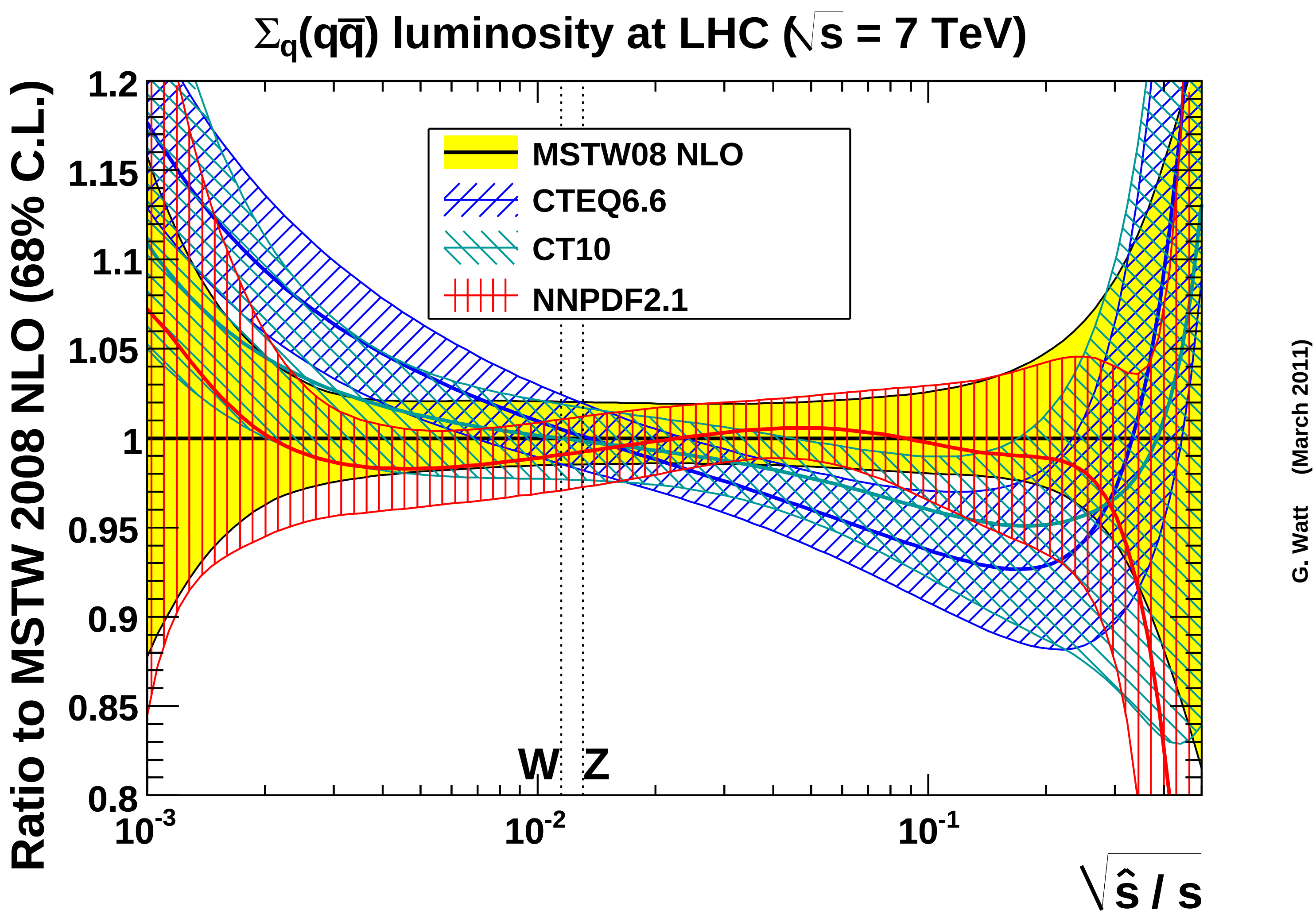
Cross-section for $t\bar{t}$ almost identical in PDF terms to $450\mathrm{GeV}$ Higgs.

Also $H + t\bar{t}$ at $\sqrt{\hat{s}/s} \sim 0.1$.

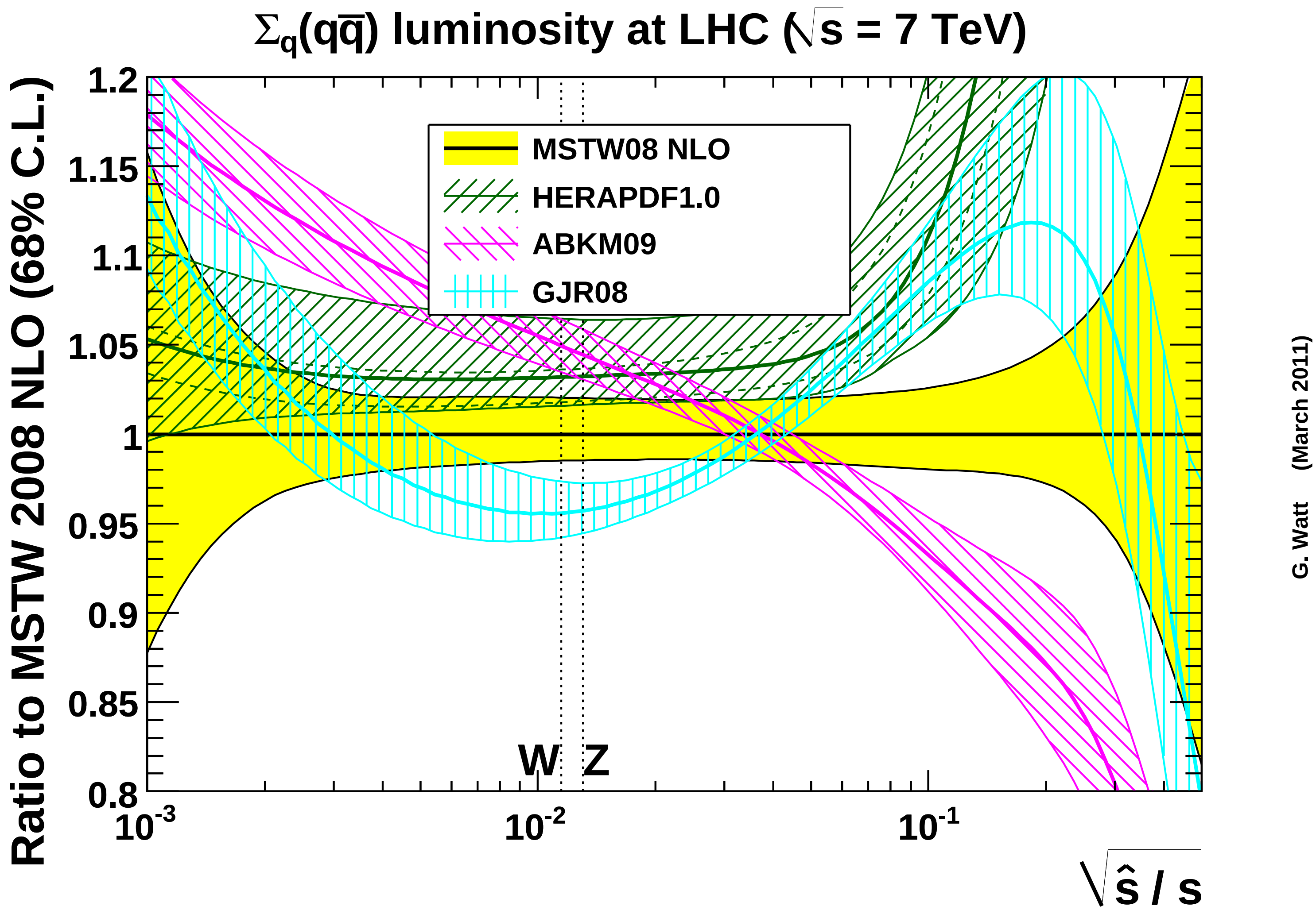


Clearly some distinct variation between groups. Much can be understood in terms of previous differences in approaches.

Uncertainties not completely comparable.



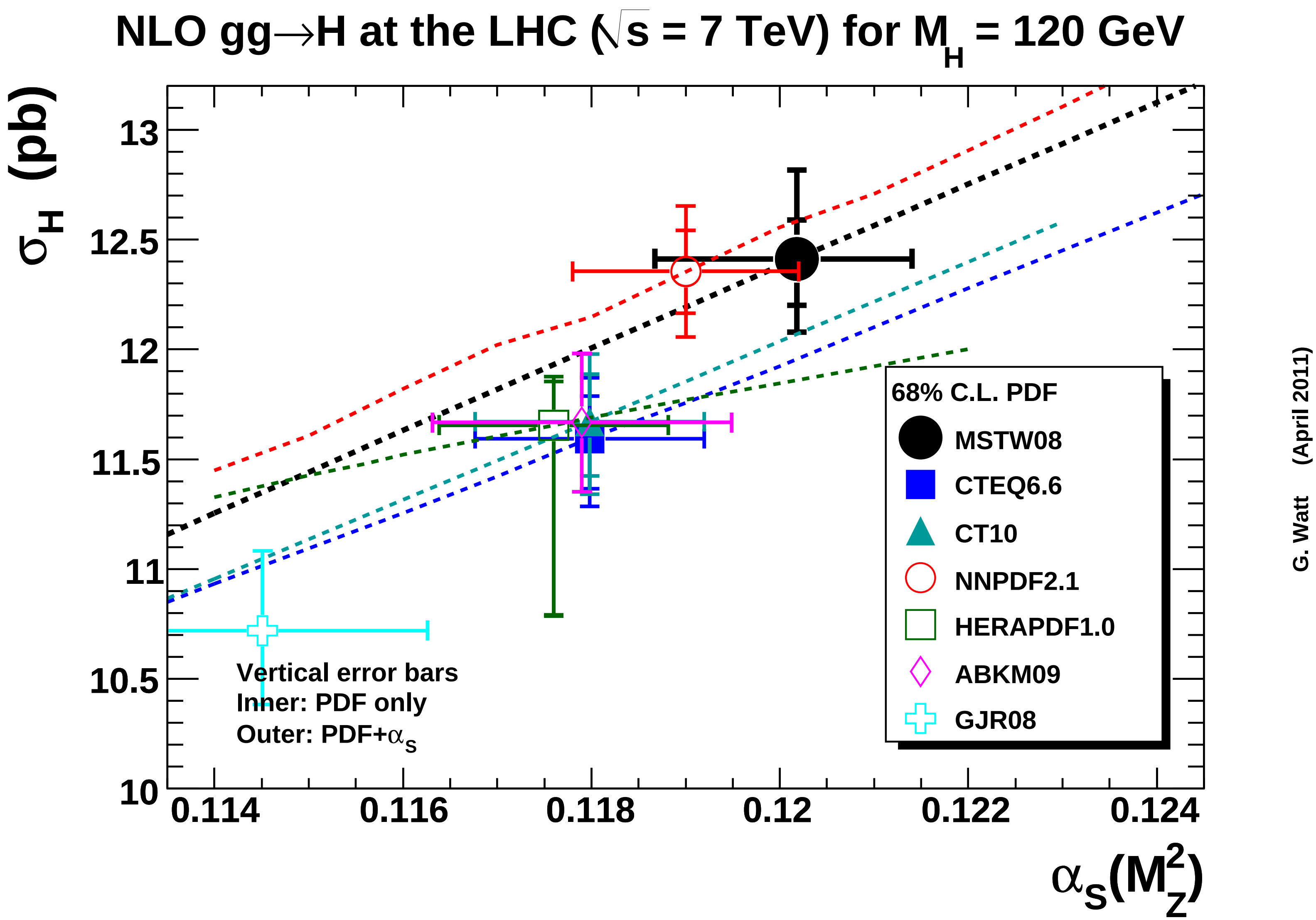
Many of the same general features for quark-antiquark luminosity. Some differences mainly at higher x.



Canonical example W, Z production, but higher $\hat{s}/s$ relevant for WH or vector boson fusion.

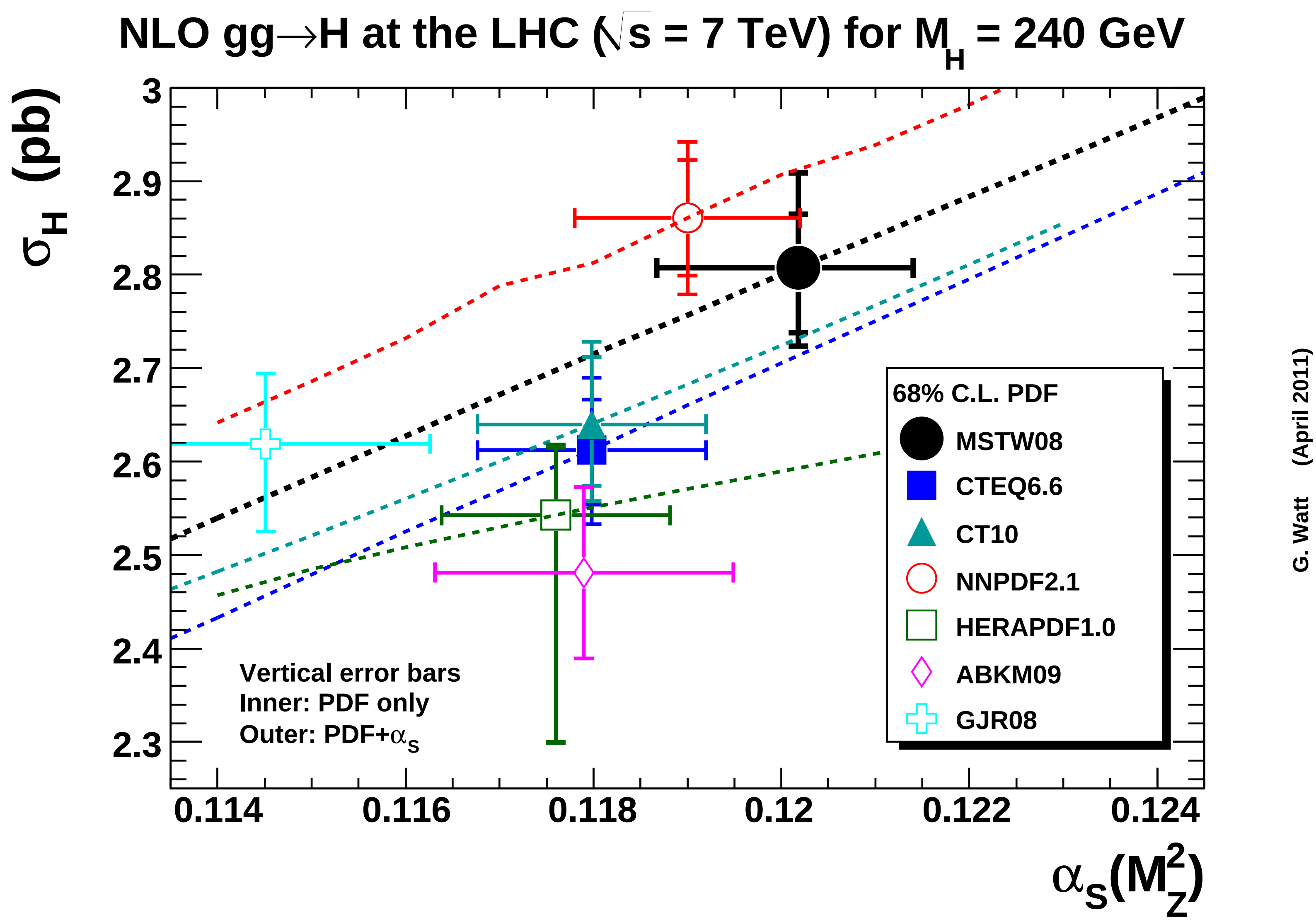
All plots and more at http://projects.hepforge.org/mstwpdf/pdf4lhc

Variations in Cross-Section Predictions – NLO

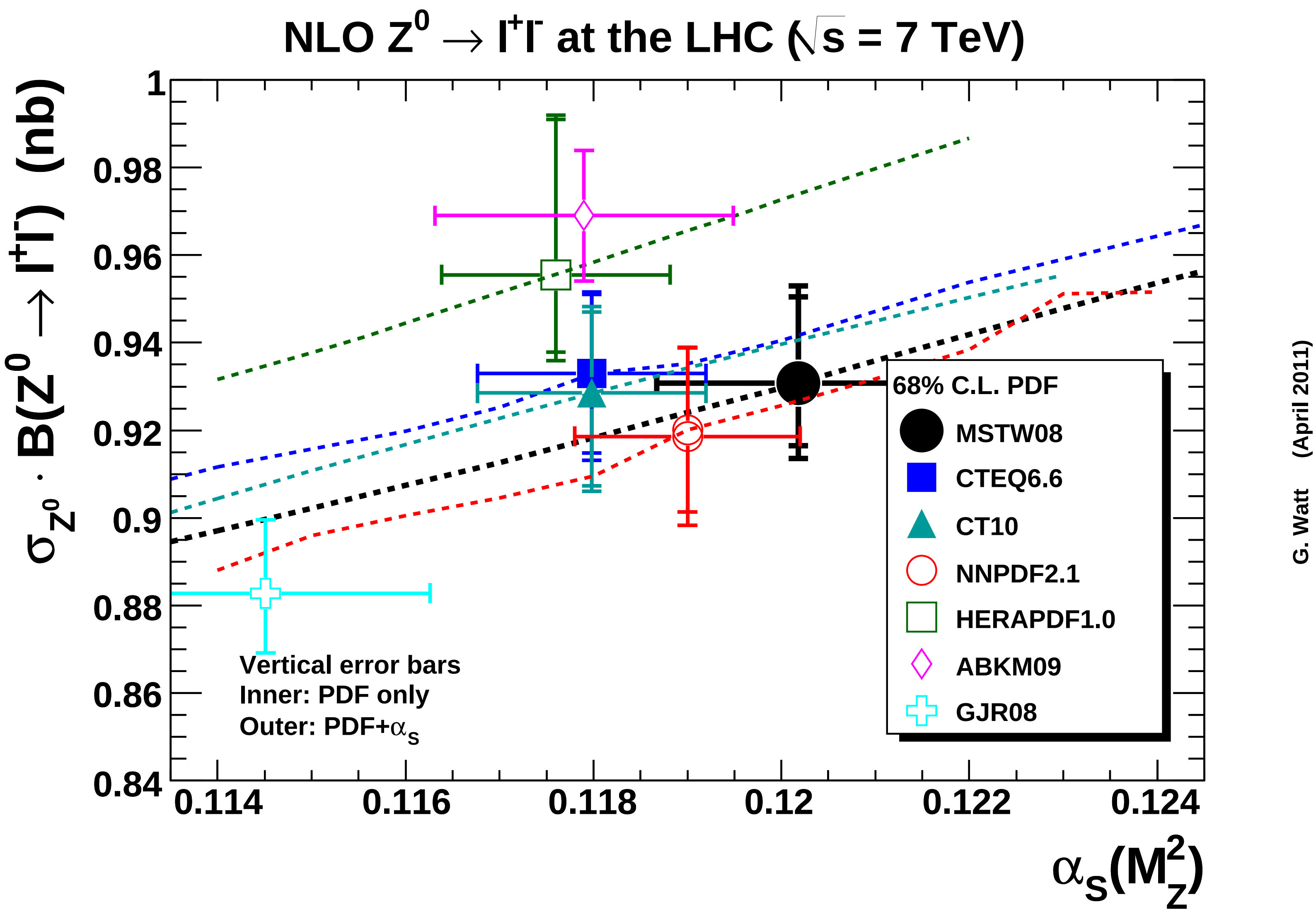


Dotted lines show how central PDF predictions vary with $\alpha_S(M_Z^2)$.

Again plots by G Watt using PDF4LHC benchmark criteria.

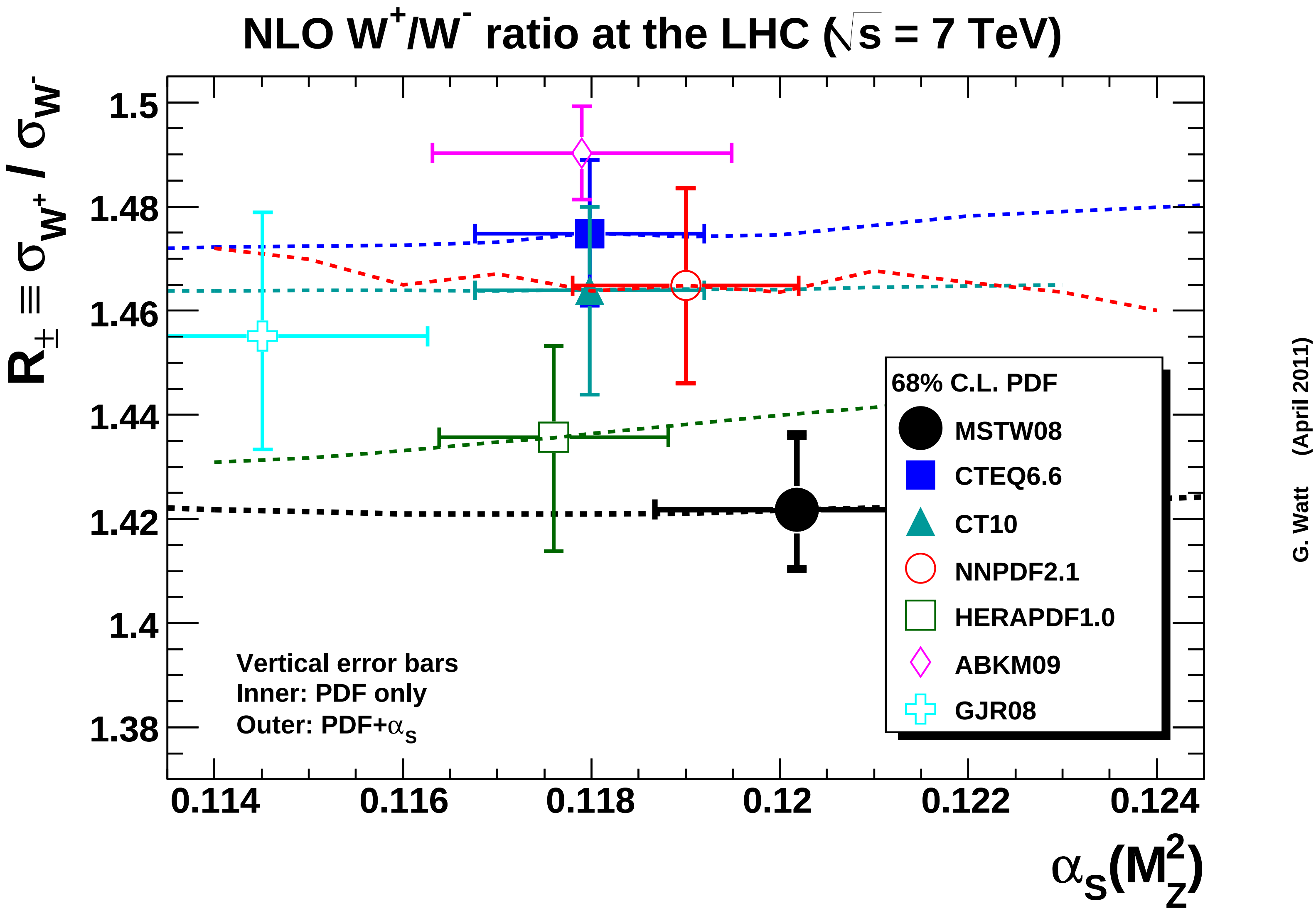


Excluding GJR08 amount of difference due to $\alpha_S(M_Z^2)$ variations $3-4\%$.



$\alpha_S(M_Z^2)$ dependence now more due to PDF variation with $\alpha_S(M_Z^2)$.

Again variations somewhat bigger than individual uncertainties.



Quite a variation in ratio. Shows variations in flavour and quark-antiquark decompositions.

All plots and more at http://projects.hepforge.org/mstwpdf/pdf4lhc

Deviations in predictions clearly much more than uncertainty claimed by each.

In some cases clear reason why central values might differ, e.g. lack of some constraining data, though uncertainties then do not reflect true uncertainty.

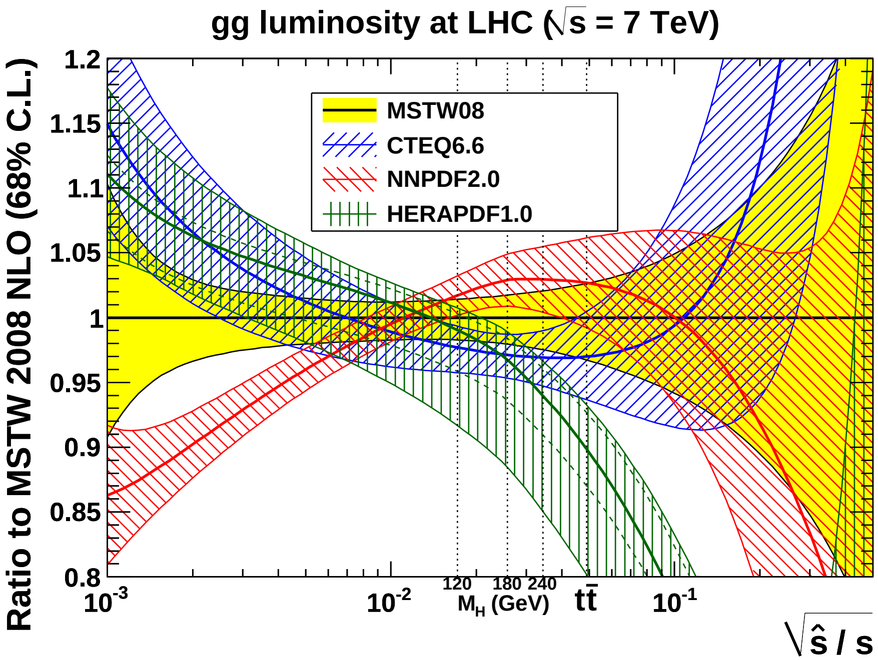
Sometimes no good understanding, or due to difference in procedure which is simply a matter of disagreement, e.g. gluon parameterisation at small x affects predicted Higgs cross-section.

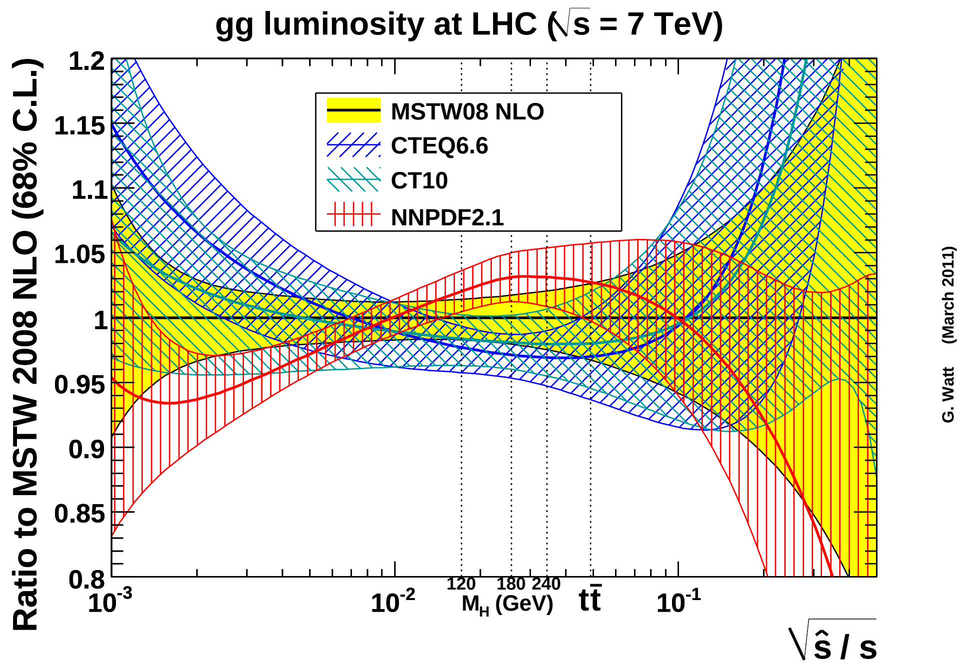
What is true uncertainty for comparing to unknown production cross section. Task asked of PDF4LHC group.

Interim recommendation take envelope of *global* sets, MSTW, CTEQ NNPDF (check other sets) and take central point as uncertainty.

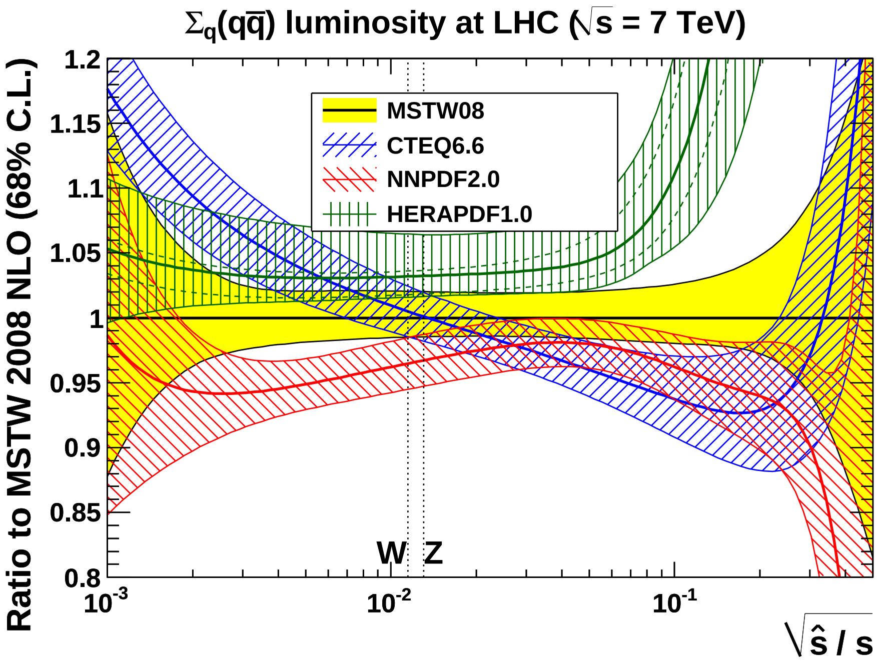
Not very satisfactory, but not clear what would be an improvement, especially as a general rule.

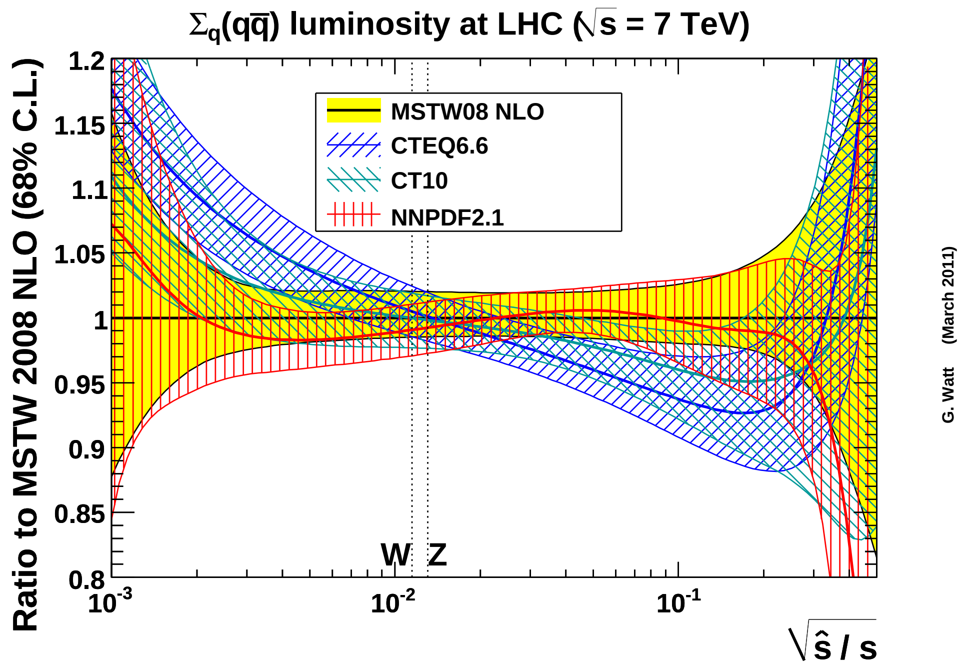
Usually not a big disagreement, and factor of about 2 expansion of MSTW uncertainty.





MSTW, NNPDF and CTEQ are converging somewhat.





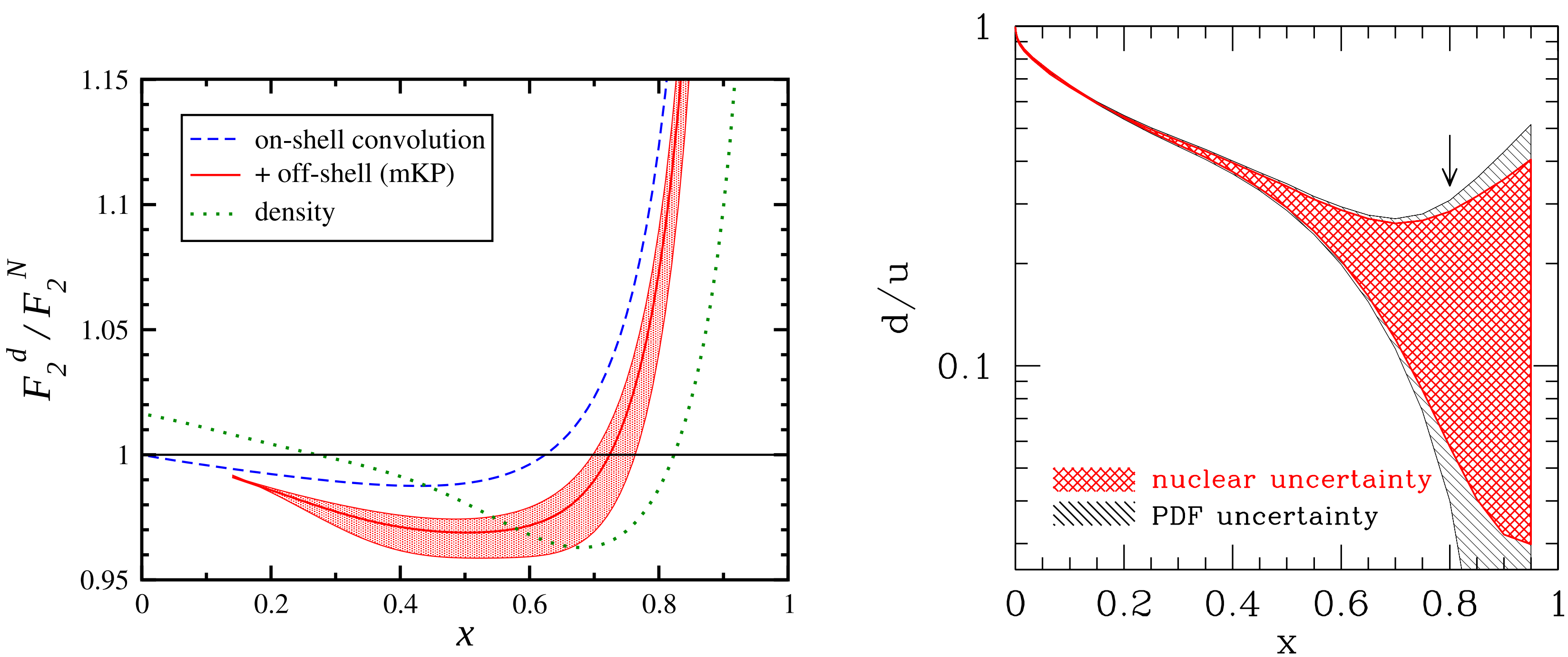
Same for quark-antiquark luminosities.

Other sources of Uncertainty.

Also other sources which (mainly) lead to inaccuracies common to all fixed-order extractions.

- Standard higher orders NNLO. Many sets available here, soon all of them.
- QED and Weak (comparable to NNLO ?) ($\alpha_s^3 \sim \alpha$). Sometime enhancements.
- Nuclear/deuterium corrections to structure functions.
- Resummations, e.g. small x ($\alpha_s^n \ln^{n-1}(1/x)$), or large x ($\alpha_s^n \ln^{2n-1}(1-x)$).
- low Q^2 (higher twist), saturation.

Deuterium corrections.



Variation in W^+/W^- ratio probably partially related to the issue of deuterium corrections.

Recent study (Accardi *et al*) suggests these may be large (also some investigations by MSTW).

Uncertainty in correction as large as PDF uncertainty, but size of corrections can be larger.

PDFs at NNLO

NNLO splitting functions (Moch, Vermaseren and Vogt) allow essentially full NNLO determination of partons now being performed (MSTW, ABKM,GJR,HERA, NNPDF), though heavy flavour not fully worked out in the fixed-flavour number scheme (FFNS) and jet cross-sections are only approximate. Improves consistency of fit very slightly, and reduces α_S.

Surely this is best, i.e. most accurate.

Yes, but ...... only know some hard cross-sections at NNLO.

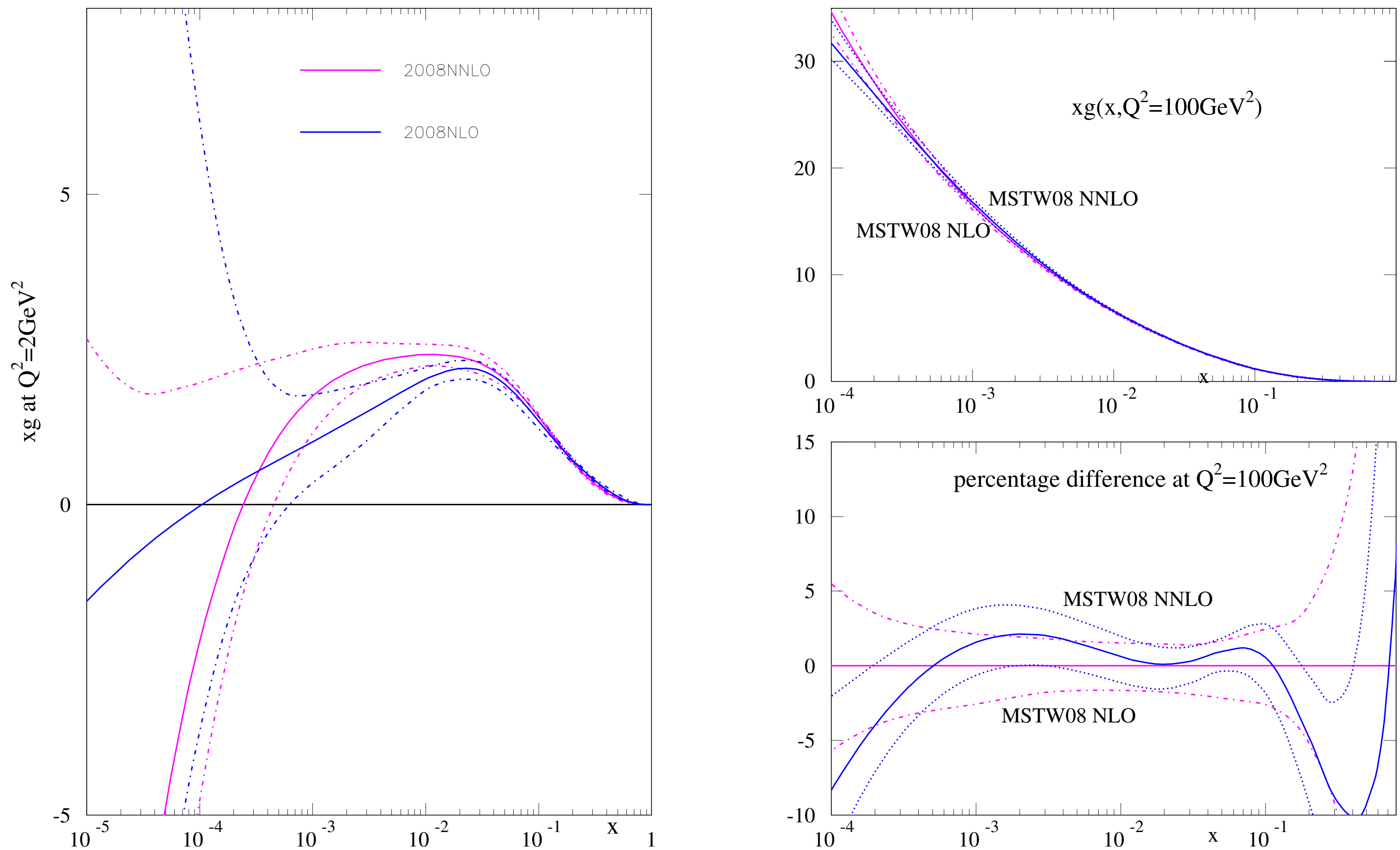
Processes with two strongly interacting particles largely completed

DIS coefficient functions and sum rules

$pp(\bar{p}) \rightarrow \gamma^\star, W, Z$ (including rapidity dist.), H, A^0, WH, ZH.

But for many other final states NNLO not known. NLO still more appropriate.

How do NNLO PDFs compare to NLO?



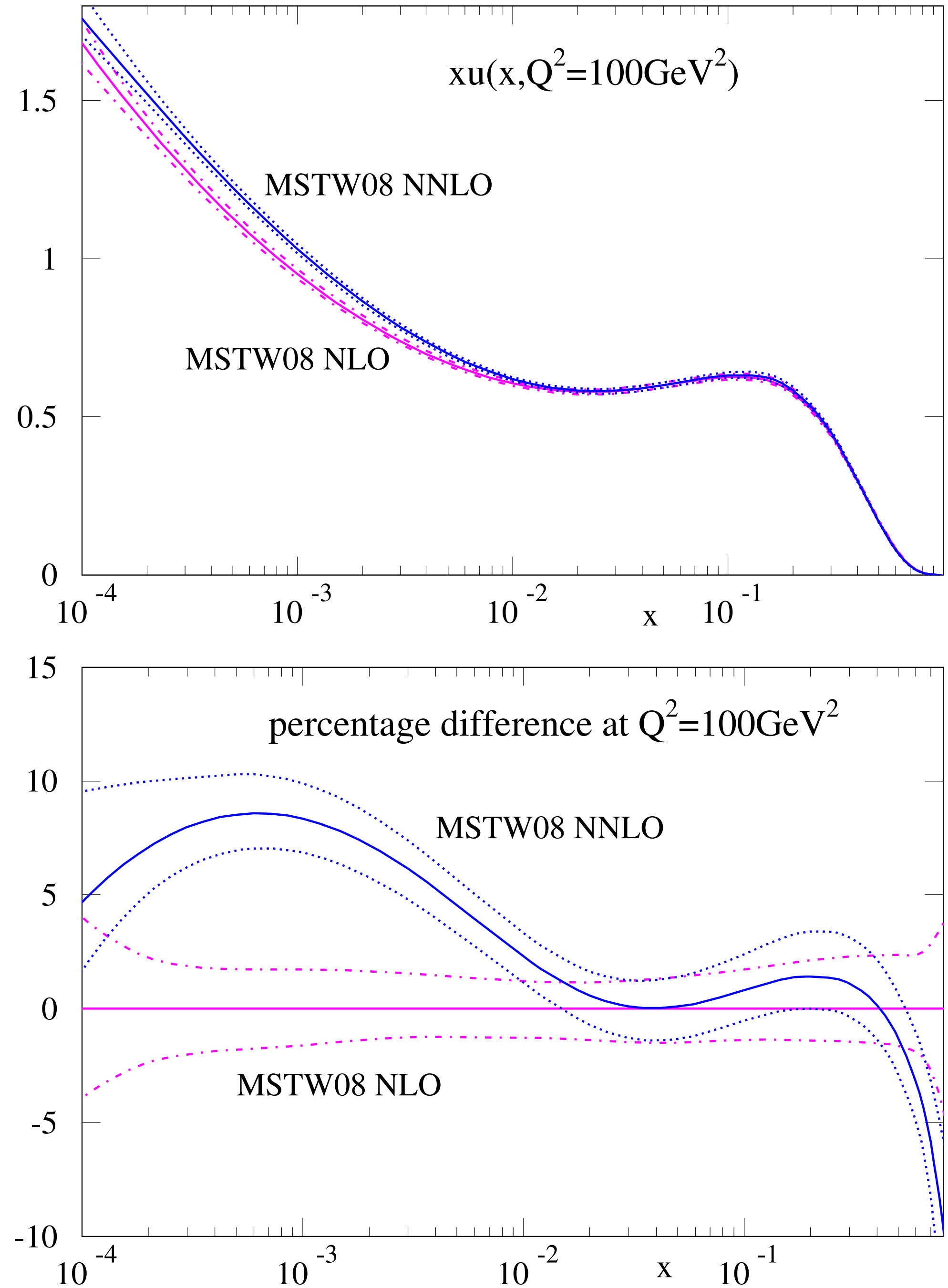
Gluons different at NLO and NNLO at low Q^2. Largely washed out by evolution, but only because of different α_S.

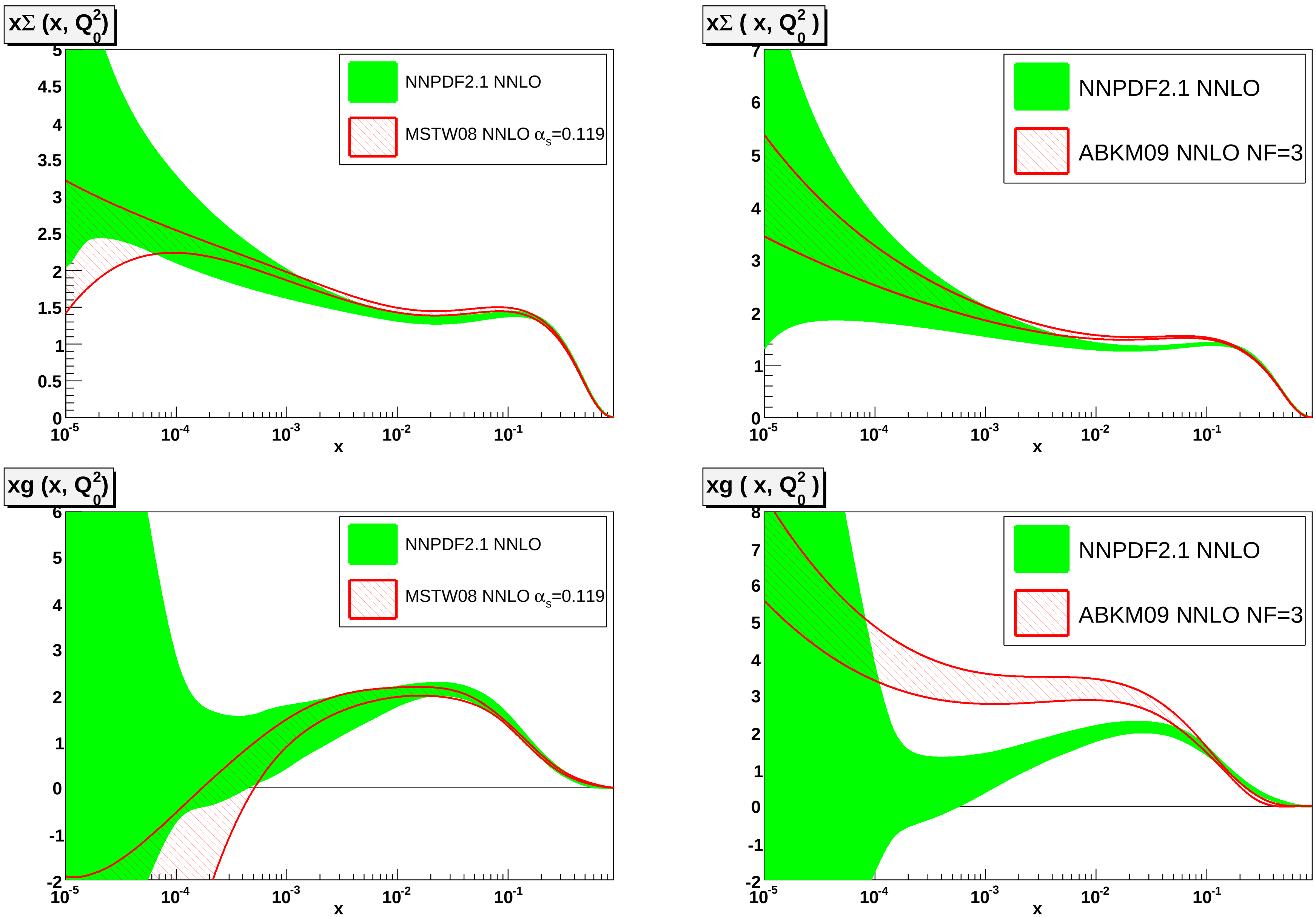
Sometimes vital to use NNLO PDFs if calculating at NNLO.

Systematic difference between PDF defined at NLO and at NNLO.

Due to large (negative) gluon coefficient function at not too small x.

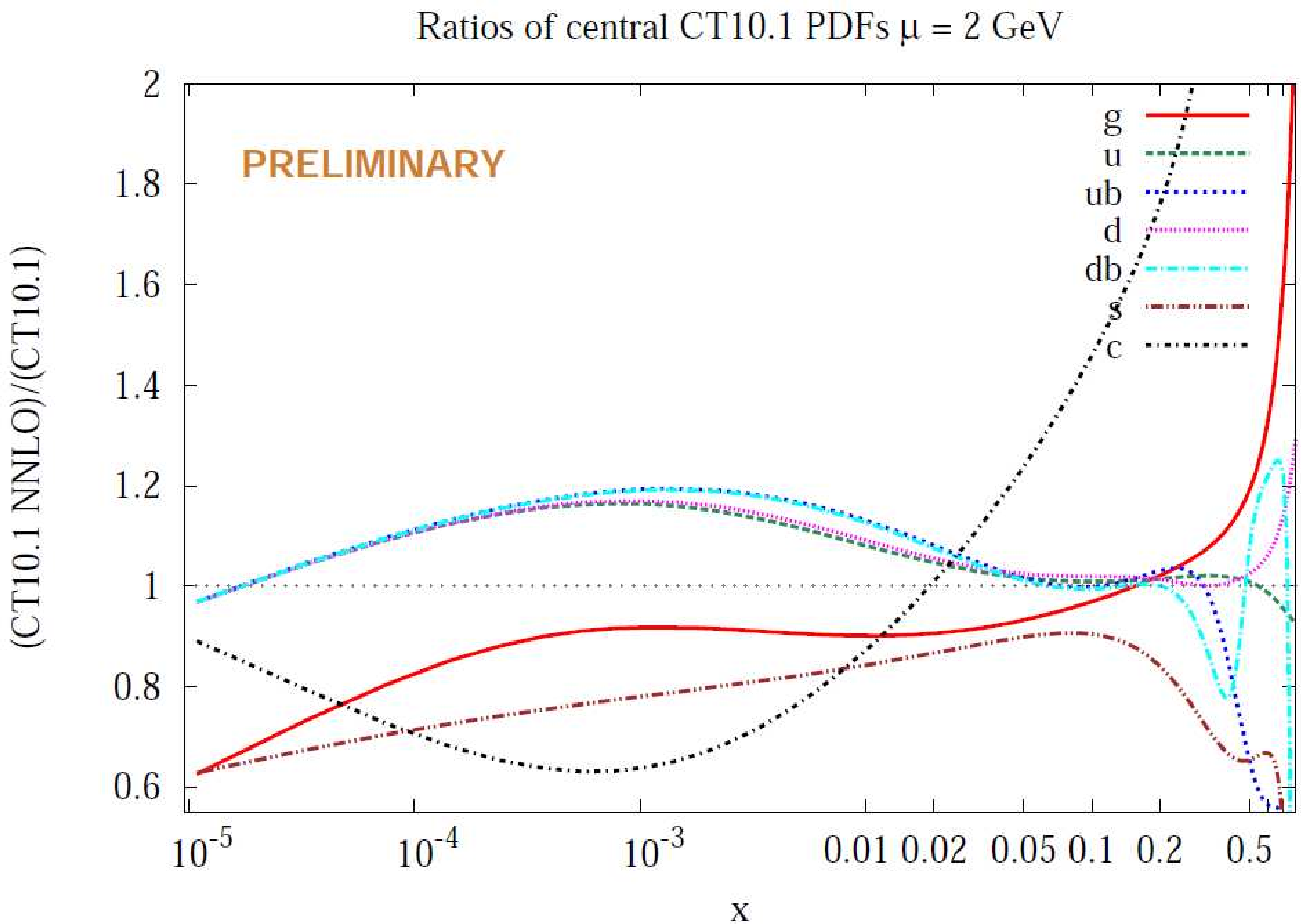
Systematic difference between PDF defined at NLO and at NNLO.





New NNPDF NNLO sets show similar trends to MSTW, some more differences to ABKM.

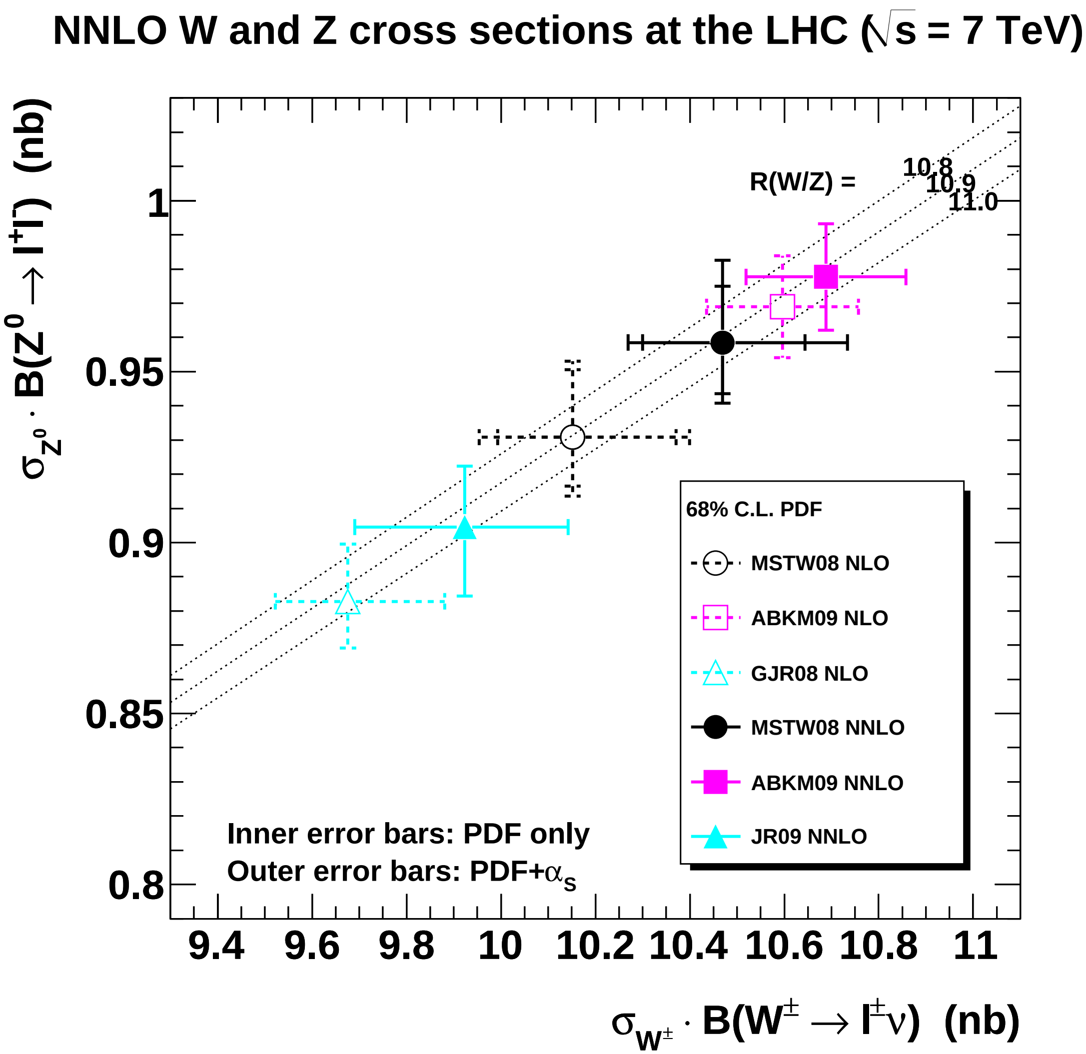
Candidate NNLO fit (compared to CT10.1 NLO)

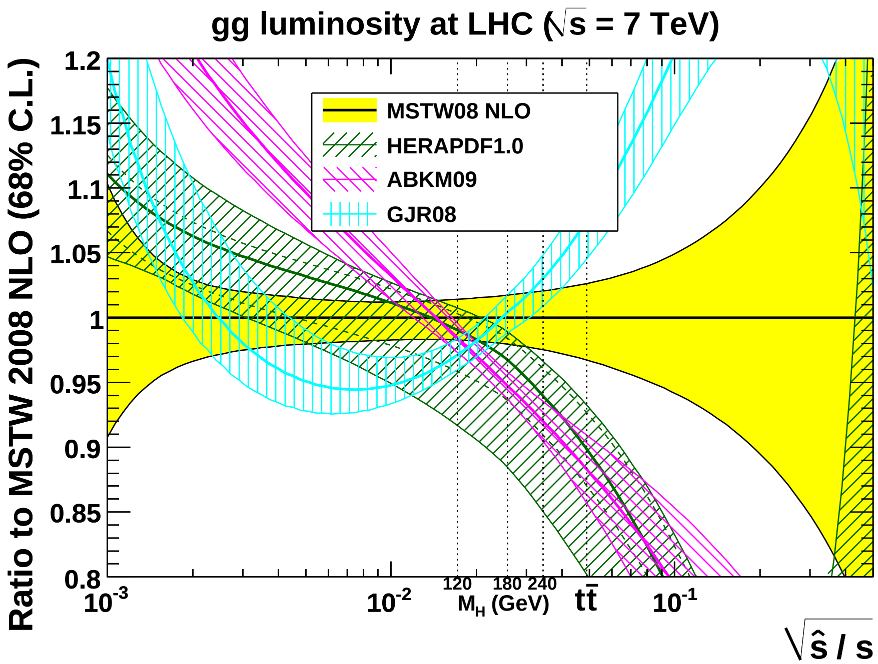


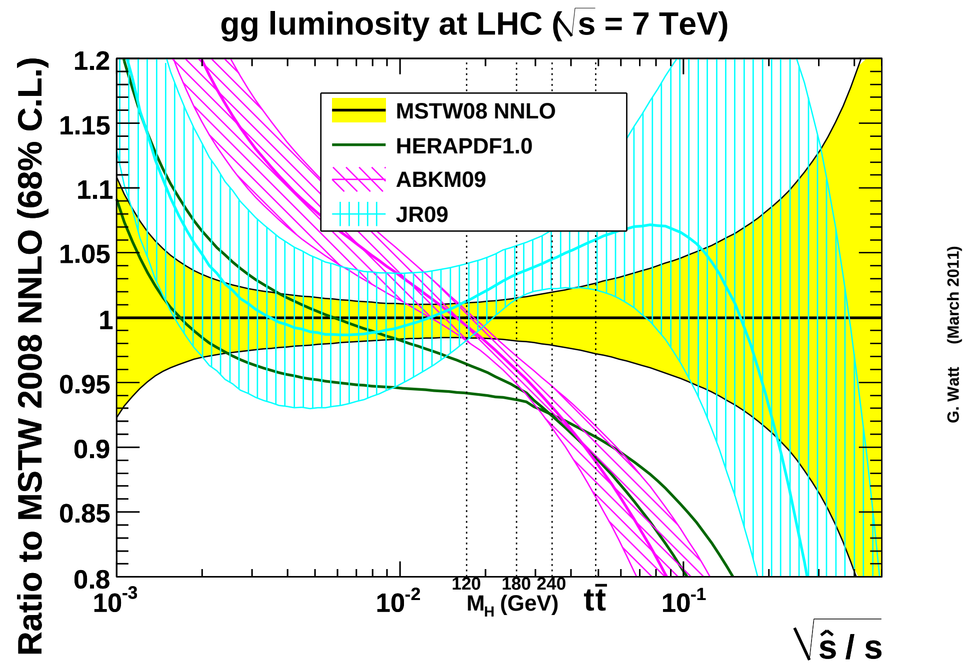
Prelim. CT10 sets show same trends as above (Nadolsky – PDF4LHC July).

Differences in predictions at NNLO compared to NLO (Watt).

Differences very much the same as they are comparing at NLO.







Luminosity differences for the gluon also largely the same at NNLO as at NLO.

Differences between different sets not likely to be due to theory choices which would diminish at higher orders, or approx. at NNLO which would change relative NLO and NNLO differences.

Considerations of differences and of NNLO

There is a significant systematic change in value from fit as one goes from NLO to NNLO.

Converging on general agreement that the NNLO values of α_S are $0.0002 - 0.0003$ smaller than the NLO values of α_S?

MSTW08 – $\alpha_S(M_Z^2) = 0.1202 \rightarrow 0.1171$.

ABKM09 – $\alpha_S(M_Z^2) = 0.1179 \rightarrow 0.1135$.

GJR/JR – $\alpha_S(M_Z^2) = 0.1145 \rightarrow 0.1124$.

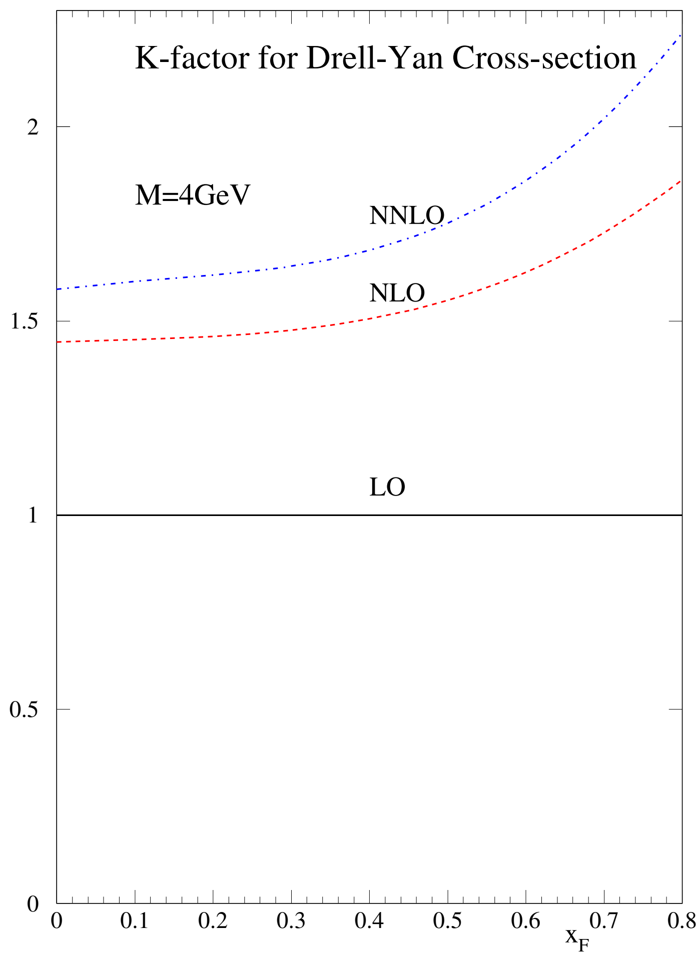
NNPDF2.1 – $\alpha_S(M_Z^2) = 0.1191 \rightarrow 0.1172$(prelim).

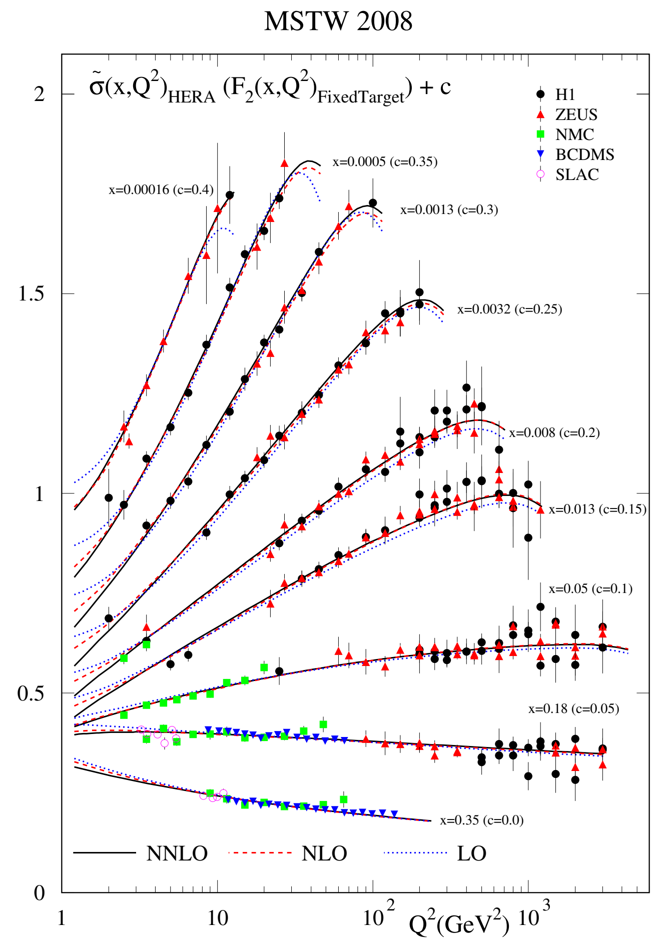
CT10.1 – $\alpha_S(M_Z^2) = 0.1196 \rightarrow 0.1180$(both prelim – PDF4LHC, DESY July).

HERAPDF1.6 – $\alpha_S(M_Z^2) = 0.1202$ at NLO and general preference for ~ 0.1176 at NNLO.

$\alpha_S(M_Z^2)$ is not a physical quantity. In (nearly) all PDF related quantities (and many others) shows tendency to decrease from order to order.

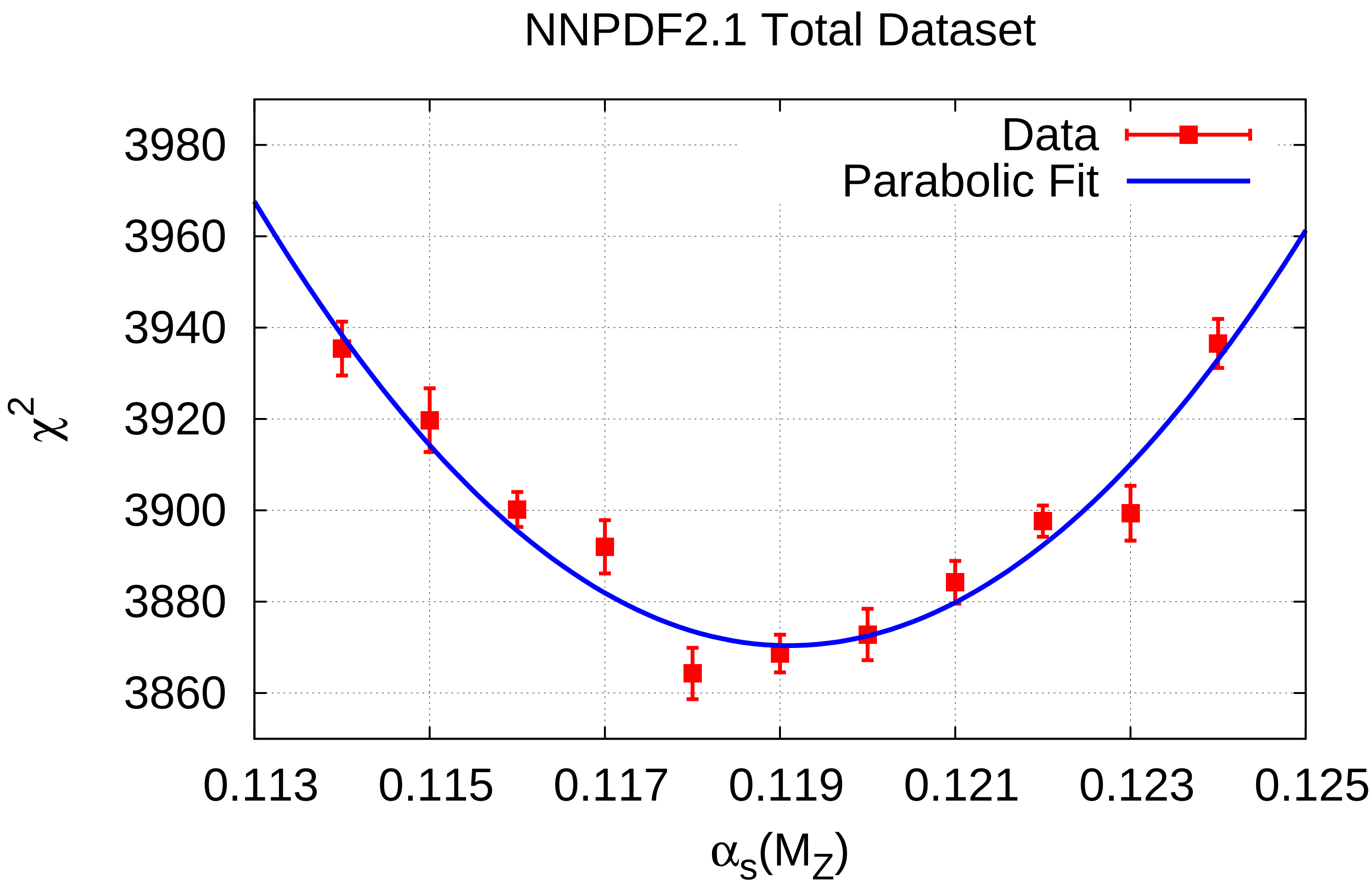
Central values differ far more than NLO $\rightarrow$ NNLO trend.





In general NNLO corrections either positive for cross sections, e.g. Drell Yan, or for evolution in structure functions.

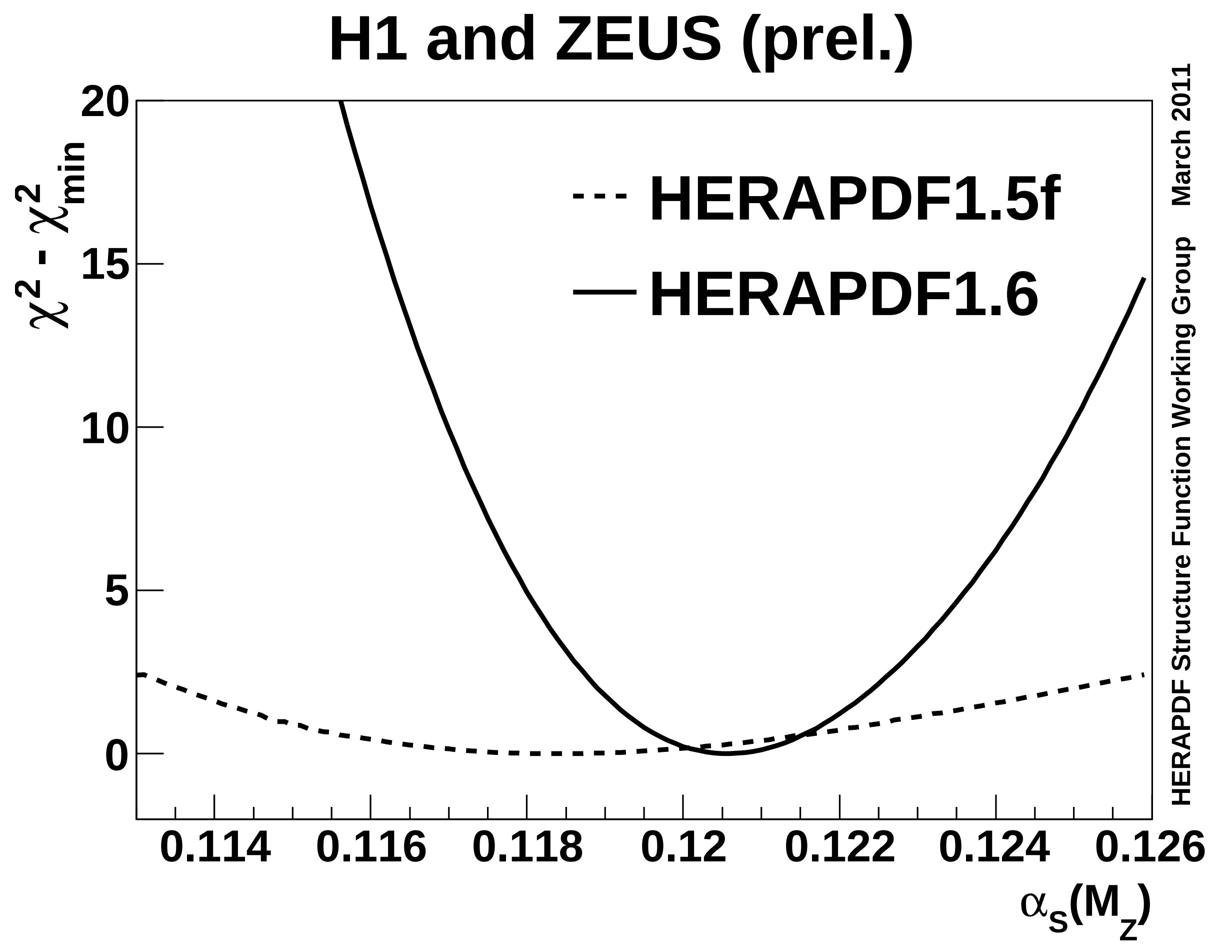
Automatically leads to lower $\alpha_S(M_Z^2)$ at NNLO than at NLO. Difference between two quite stable.



In full fit NNPDF now get precise value – 0.1191 ± 0.0006 from χ^2 profile.

Similar uncertainty to MSTW if $\Delta\chi^2 = 1$ used as criterion.

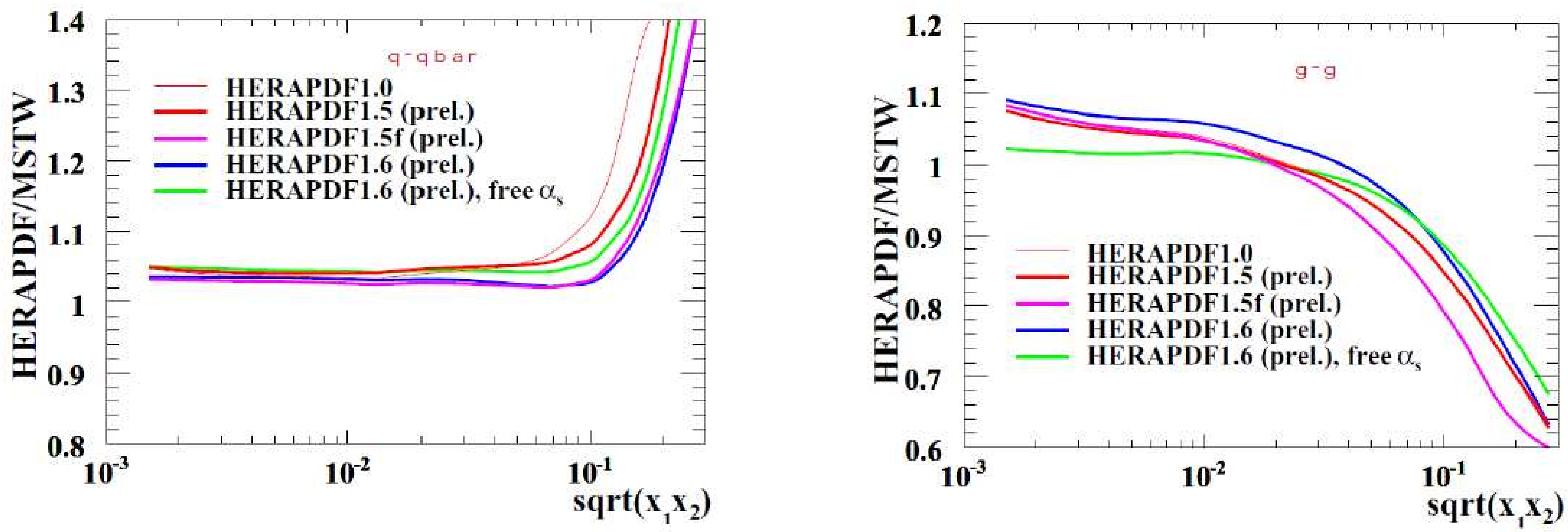
Similar $\Delta\chi^2$ profile to CT10.



HERAPDF have little constraint on $\alpha_S(M_Z^2)$ when fitting only to their structure function data. Inclusion of their jet data ties down α_S and consequently (as well as directly) constrains the gluon.

Lack of precise $\alpha_S(M_Z^2)$ determination from HERA or to some extent full DIS data noticed by MSTW, NNPDF.

LHC at 7 TeV parton-parton luminosity plots for HERAPDF1.5 in ratio to MSTW2008



The q-qbar luminosity at NLO

HERAPDF1.5 is softer than 1.0 at high-x and 1.5f is even softer

Adding the jets to make it 1.6 makes very little difference

Letting alphas be free so that $\alpha_S(M_Z)$=0.1202 rather than 0.1176 hardens the high-x quark distribution marginally

The g-g luminosity at NLO

HERAPDF1.5 is on top of 1.0 and 1.5f is slightly softer

Adding the jets to make it 1.6 hardens the high-x gluon

Letting alphas be free so that $\alpha_S(M_Z)$=0.1202 rather than 0.1176 also reduces the low-x gluon

9

Also find adding HERA jet data and improved charged current data softens antiquarks at high x and hardens gluon. (Cooper-Sarkar – PDF4LHC July).

ABM look at including individual jet data sets in the fit. Generally raises $\alpha_S(M_Z^2)$ a little, the high-x gluon and Higgs cross section predictions.

TABLE 1. The values of the strong coupling $\alpha_s(M_Z)$ obtained in global fits of PDFs at various orders of perturbation theory as indicated in the first column. The second column gives the results of the ABKM09 fit [11], the other columns are obtained from variants of the ABKM09 fit including data either for 1-jet inclusive or for di-jet production from the collaborations D0 [4, 5] or CDF [2, 3]. The value in bold corresponds to the published result in [11].

$\alpha_s(M_Z)$	ABKM09	D0 1-jet inc.	D0 di-jet	CDF 1-jet inc. (cone)	CDF 1-jet inc. (k_T)
NLO	0.1179(16)	0.1190(11)	0.1174(9)	0.1181(9)	0.1181(10)
NNLO	**0.1135(14)**	0.1149(12)	0.1145(9)	0.1134(9)	0.1143(9)

TABLE 2. The predicted cross sections for Higgs boson production in gluon-gluon fusion with $M_H = 165$ GeV at Tevatron ($\sqrt{s} = 1.96$ TeV) from variants of the ABKM09 fit [11] corresponding to Tab. 1. The uncertainty in brackets refers to 1 σ standard deviation for the combined uncertainty on the PDFs and the value of $\alpha_s(M_Z)$. The value in bold corresponds to the published result [1].

$\sigma(H)[pb]$	ABKM09	D0 1-jet inc.	D0 di-jet	CDF 1-jet inc. (cone)	CDF 1-jet inc. (k_T)
NLO	0.206(17)	0.235(10)	0.212(10)	0.229(8)	0.229(8)
NNLO	**0.253(22)**	0.297(12)	0.281(12)	0.283(10)	0.292(10)

TABLE 3. Same as Tab. 2 for the LHC ($\sqrt{s} = 7$ TeV).

$\sigma(H)[pb]$	ABKM09	D0 1-jet inc.	D0 di-jet	CDF 1-jet inc. (cone)	CDF 1-jet inc. (k_T)
NLO	5.73(17)	5.89(13)	5.76(10)	5.76(12)	5.77(11)
NNLO	**7.05(23)**	7.30(15)	7.28(14)	7.02(14)	7.18(14)

Also reduces uncertainties (by at most a factor of 2).

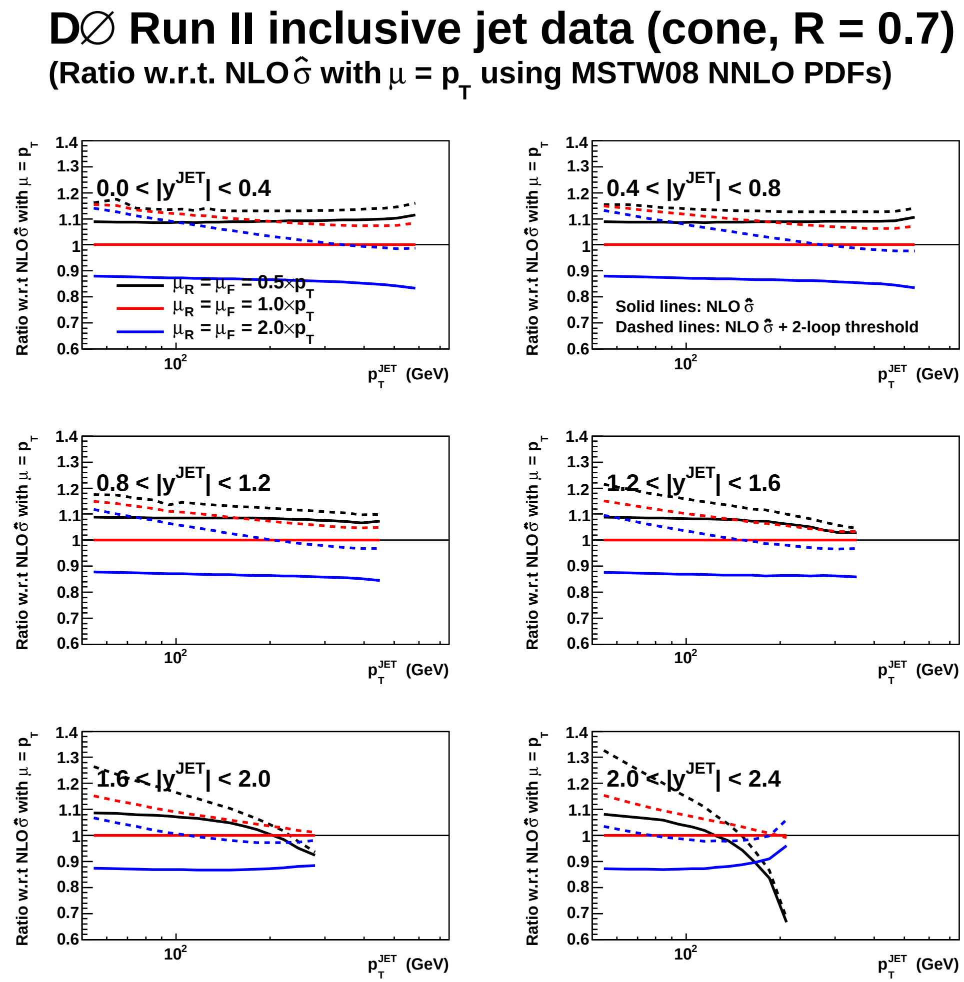
NNLO approx. jet corrections.

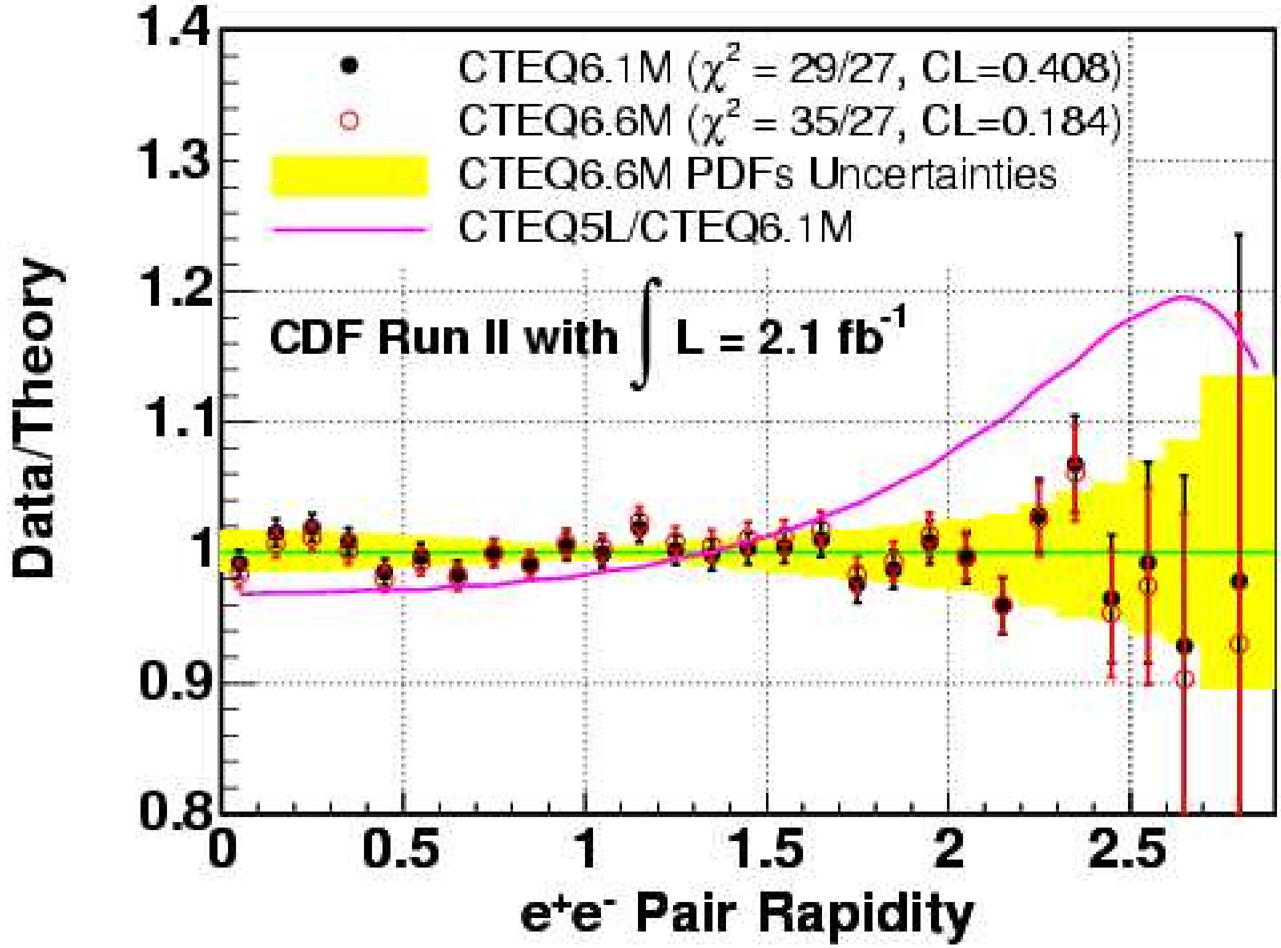
Shape of corrections as function of p_T at NLO and also at approx. NNLO in inclusive case.

NNLO uses threshold (Kidonakis and Owens) approx. for Tevatron jets.

NNLO approximation not large and aids stability – always worst at high-p_T i.e. high-x. Includes large $\ln(p_T/\mu)$ terms predicted by renormalisation group.

Similar conclusion from de Florian and Vogelsang.





Important point, CDF Z-rapidity data sets Tevatron normalisation in a fit.

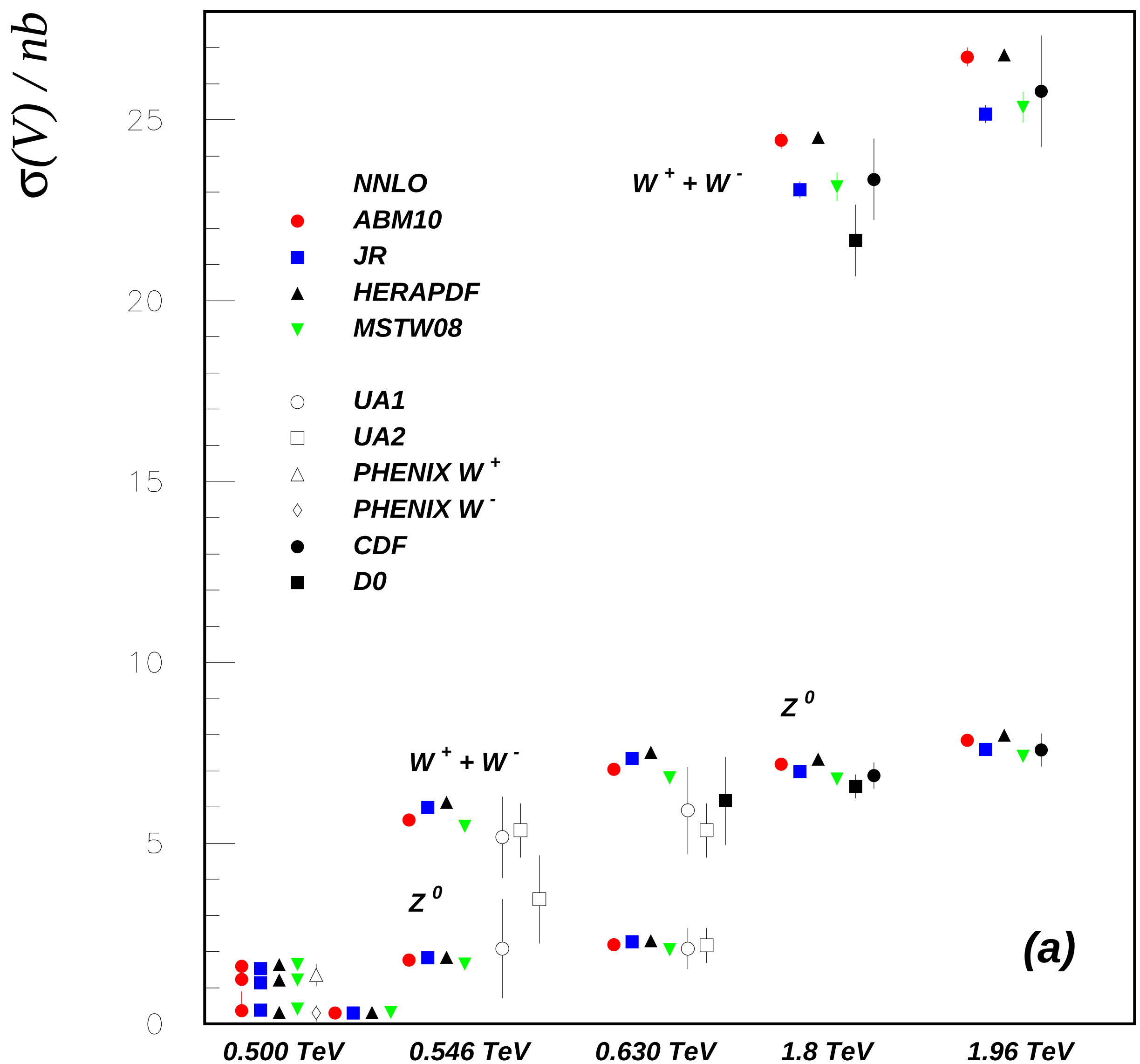
Only allows a few percent variation in normalisation.

Similarly total W, Z cross sections set the normalisation, i.e. theory and data must match.

Study of predictions published by Alekhin *et al*.

If anything theory too high – Tevatron normalisation should go up.

Note consistent normalisation difference between CDF and D0, the latter rather low.



NLO PDF (with NLO $\hat{\sigma}$)	$\mu = p_T/2$	$\mu = p_T$	$\mu = 2p_T$
MSTW08	**0.75** (0.30)	**0.68** (0.28)	0.91 (0.84)
CTEQ6.6	1.25 (0.14)	1.66 (0.20)	2.38 (0.84)
CT10	1.03 (0.13)	1.20 (0.19)	1.81 (0.84)
NNPDF2.1	**0.74** (0.29)	**0.82** (0.25)	1.23 (0.69)
HERAPDF1.0	2.43 (0.39)	3.26 (0.66)	4.03 (1.67)
HERAPDF1.5	2.26 (0.40)	3.05 (0.66)	3.80 (1.66)
ABKM09	1.62 (0.52)	2.21 (0.85)	3.26 (2.10)
GJR08	1.36 (0.23)	0.94 (0.13)	**0.79** (0.36)

NNLO PDF (with NLO+2-loop $\hat{\sigma}$)	$\mu = p_T/2$	$\mu = p_T$	$\mu = 2p_T$
MSTW08	1.39 (0.42)	**0.69** (0.44)	0.97 (0.48)
HERAPDF1.0, $\alpha_S(M_Z^2) = 0.1145$	2.64 (0.36)	2.15 (0.36)	2.20 (0.46)
HERAPDF1.0, $\alpha_S(M_Z^2) = 0.1176$	2.24 (0.35)	1.17 (0.32)	1.23 (0.31)
ABKM09	2.55 (0.82)	2.76 (0.89)	3.41 (1.17)
JR09	**0.75** (0.37)	1.26 (0.41)	2.21 (0.49)

Table 1: Values of $\chi^2/N_{\rm pts.}$ for the CDF Run II inclusive jet data using the k_T jet algorithm with $N_{\rm pts.} = 76$ and $N_{\rm corr.} = 17$, for different PDF sets and different scale choices At most a 1-σ shift in normalisation is allowed.

NLO PDF (with NLO $\hat{\sigma}$)	$\mu = p_T/2$	$\mu = p_T$	$\mu = 2p_T$
MSTW08	0.75 (+0.32)	0.68 (−0.88)	0.63 (−*2.69*)
CTEQ6.6	1.03 (−*2.47*)	1.04 (−**3.49**)	0.99 (−**4.75**)
CT10	0.99 (−*1.64*)	0.92 (−*2.69*)	0.86 (−**4.10**)
NNPDF2.1	0.74 (−0.33)	0.79 (−*1.60*)	0.80 (−**3.12**)
HERAPDF1.0	1.52 (−**4.07**)	1.57 (−**5.21**)	1.43 (−**6.22**)
HERAPDF1.5	1.48 (−**3.85**)	1.52 (−**5.00**)	1.39 (−**6.03**)
ABKM09	1.03 (−**3.49**)	1.01 (−**4.53**)	1.05 (−**5.80**)
GJR08	1.14 (+*2.47*)	0.93 (+*1.25*)	0.79 (−0.50)

NNLO PDF (with NLO+2-loop $\hat{\sigma}$)	$\mu = p_T/2$	$\mu = p_T$	$\mu = 2p_T$
MSTW08	1.39 (+0.35)	0.69 (−0.45)	0.97 (−*1.30*)
HERAPDF1.0, $\alpha_S(M_Z^2) = 0.1145$	2.37 (−*2.65*)	1.48 (−**3.64**)	1.29 (−**4.12**)
HERAPDF1.0, $\alpha_S(M_Z^2) = 0.1176$	2.24 (−0.48)	1.13 (−*1.60*)	1.09 (−*2.23*)
ABKM09	1.53 (−**4.27**)	1.23 (−**5.05**)	1.44 (−**5.65**)
JR09	0.75 (+0.13)	1.26 (−0.61)	2.20 (−*1.22*)

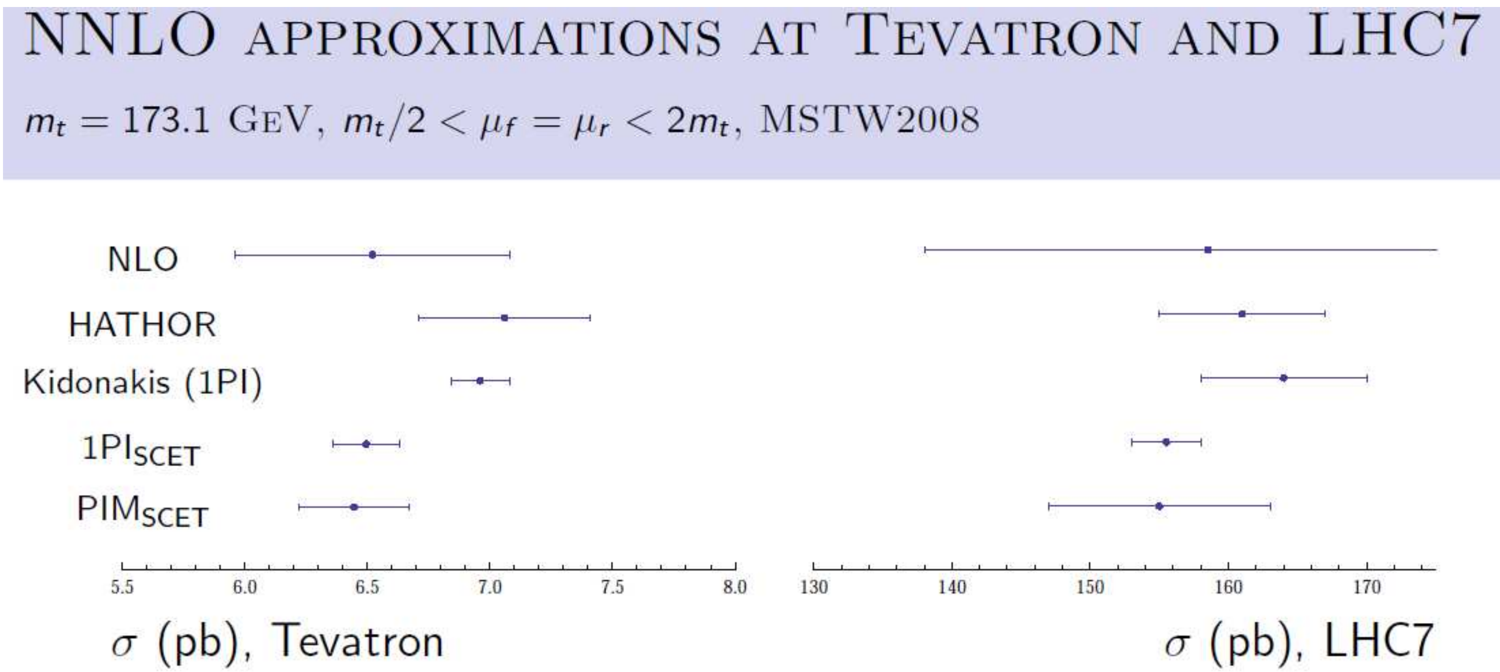
Table 2: Values of $\chi^2/N_{\rm pts.}$ for the CDF Run II inclusive jet data using the k_T jet algorithm No restriction is imposed on the shift in normalisation and the optimal value of "$-r_{\rm lumi.}$" is shown in brackets.

Top-antitop Cross-section

Inclusive cross-section known approximately to NNLO

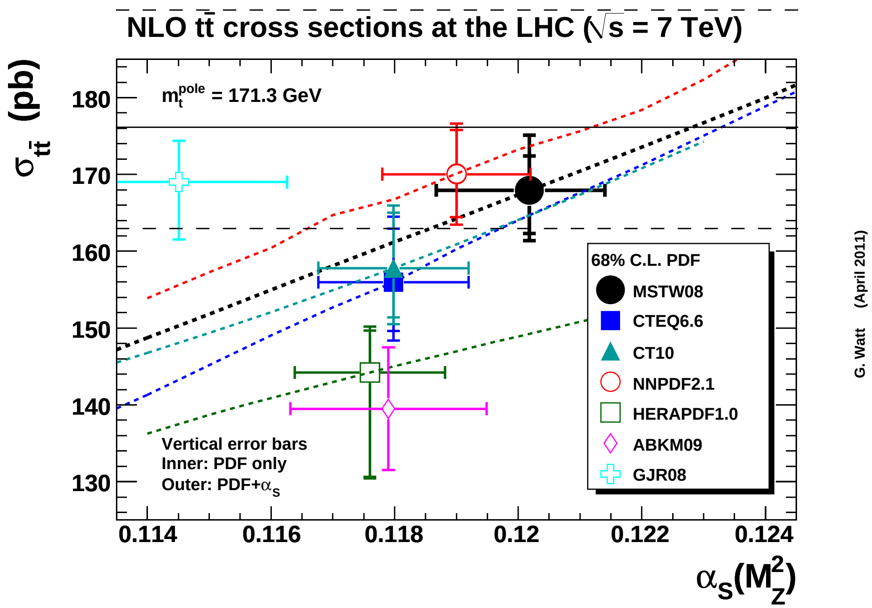
Intrinsic theory uncertainty not very large – see. e.g. talk by Pecjak at EPS 2011.

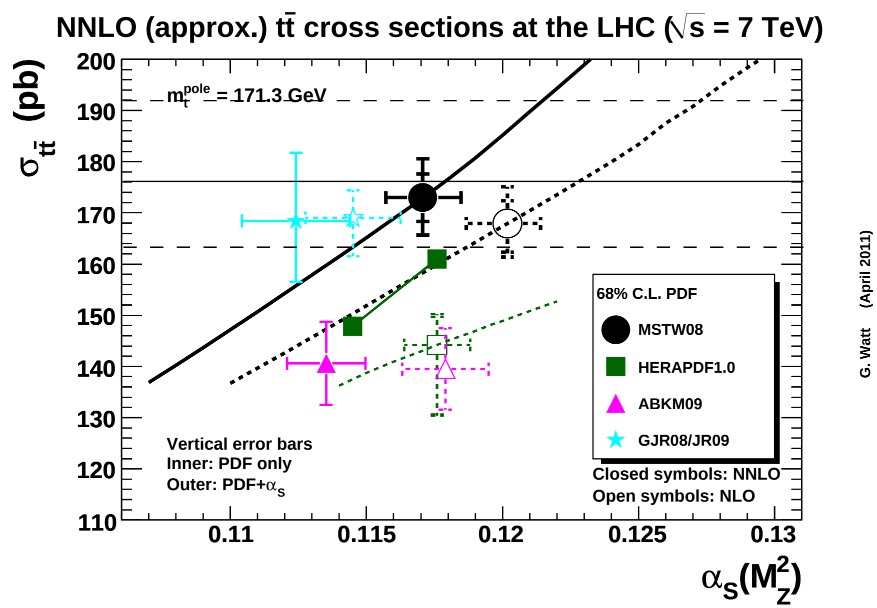
Error bars contain scale dependence and PDF uncertainty at 90%CL.



Data getting precise. Main uncertainty in choice of PDFs, not in individual uncertainty but choice of set.

Plots by G. Watt – modified by RST



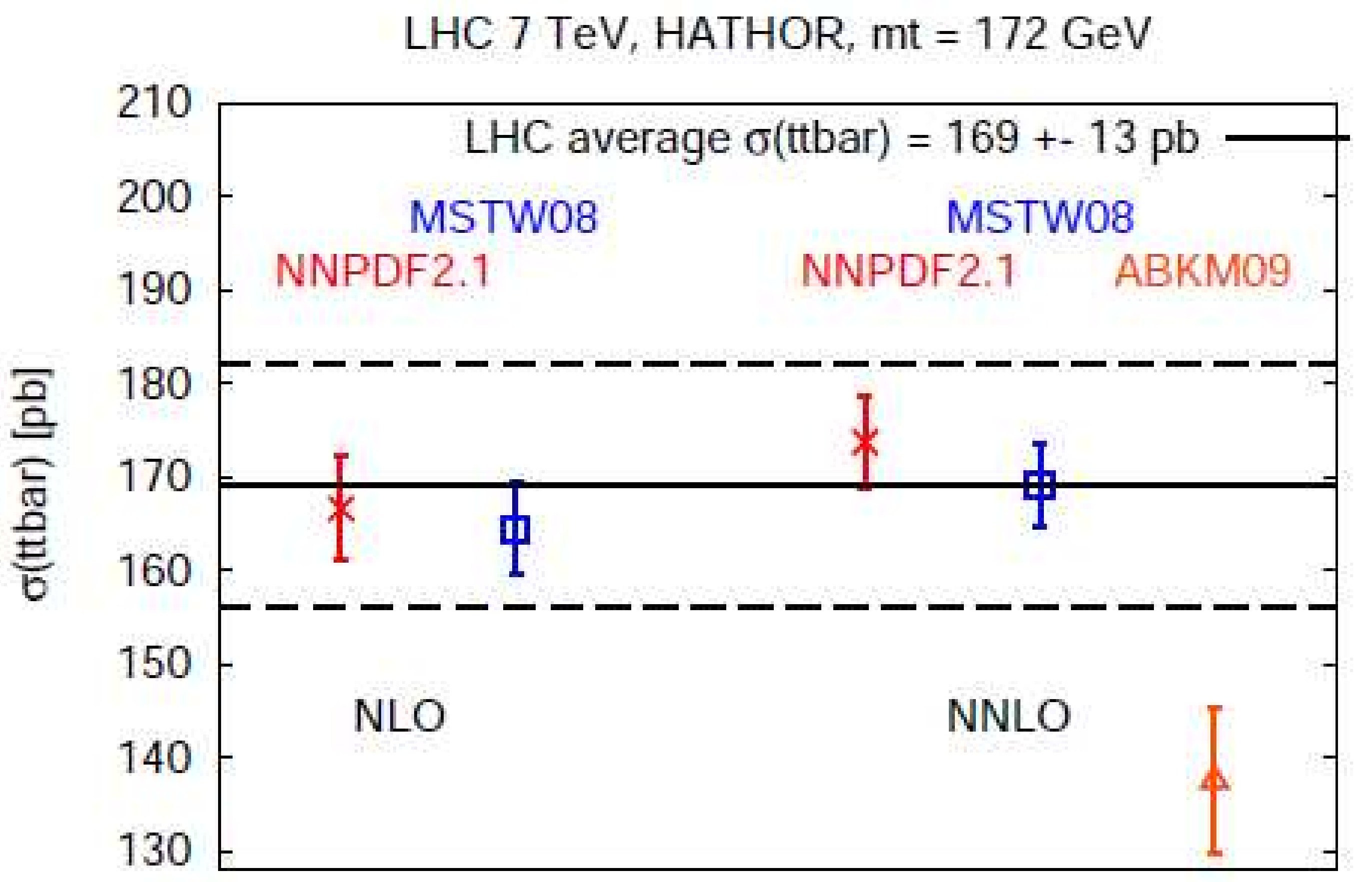


Differences between groups significant at NLO, and at NNLO.

Approx NNLO using HATHOR - (Aliev *et al*), includes scale-dependent parts and large threshold corrections at NNLO. Hence some theoretical uncertainty, but NNLO corrections not large at LHC. See lower NNLO α_S improves stability.

m_t settled at about 172-3GeV? Lowers these predictions by $5 - 10\mathrm{pb}$.

Top cross-section measurement potential discriminator of PDF sets, and correlated to Higgs predictions. For example, ATLAS preliminary combined $\sigma_{t\bar{t}} = 176^{+16}_{-13}\mathrm{pb}$.



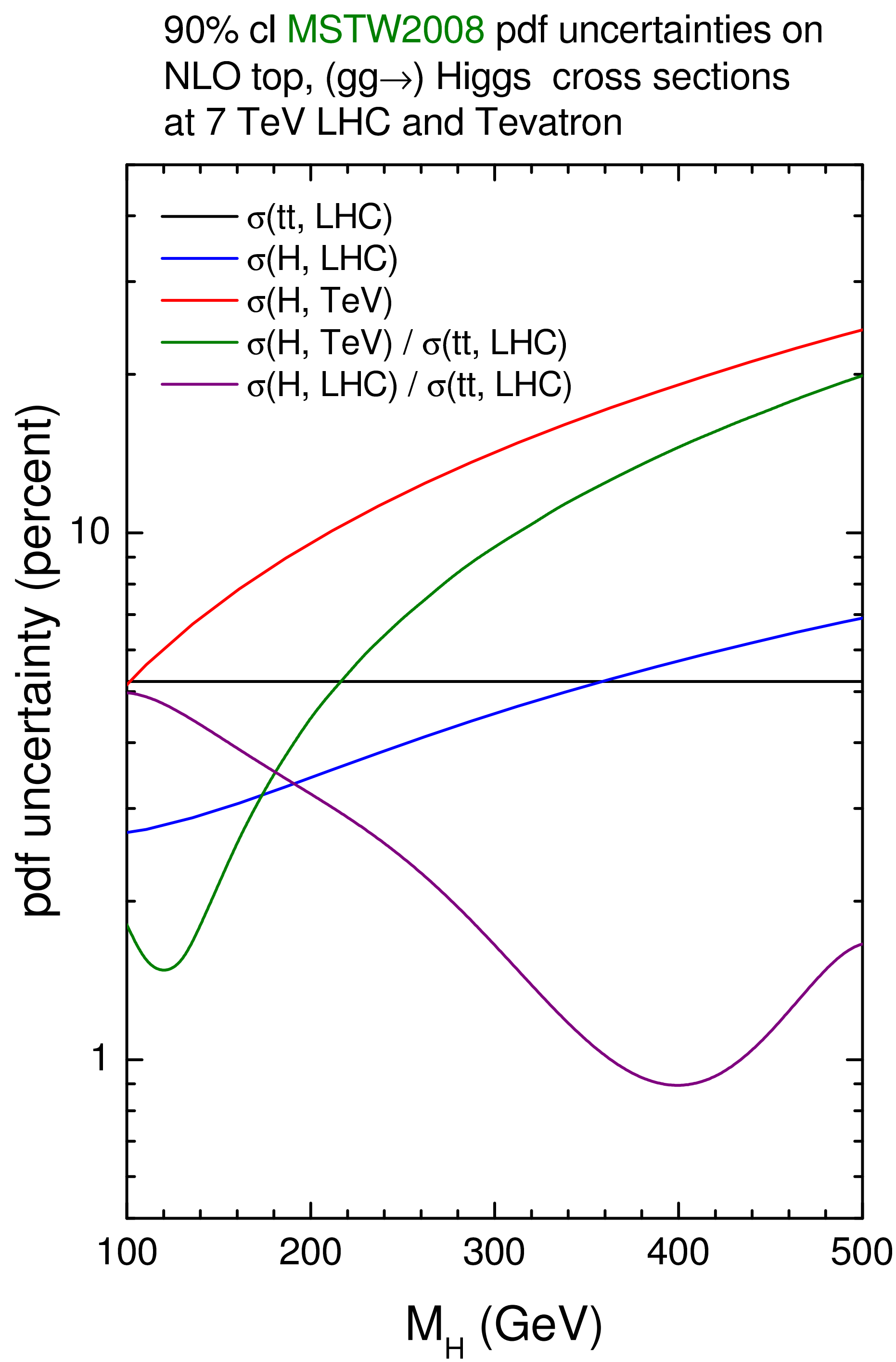
NNPDF NNLO prediction slightly bigger than MSTW, but use $\alpha_S = 0.119$ – not preferred value? General very good agreement

Uncertainty in $t\bar{t}$, Higgs via gluon fusion and ratios. PDF only uncertainty, but α_S uncertainty cancels in ratios.

Very strong correlation of top with Higgs production for $m_H \sim 400\text{GeV}$ at the LHC.

Similar correlation for $m_H \sim 400 \times 1.96/7 \sim 130\text{GeV}$ at the Tevatron.

Particularly important at the moment.



PDFs for LO Monte Carlo generators.

Often (sometimes) need to use generators which calculate only at LO in QCD.

LO matrix elements $+$ LO PDFs often very inaccurate in normalisation and general shape.

Using NLO PDFS suggested – sometimes better, sometimes even worse (particularly small x, important for underlying event etc).

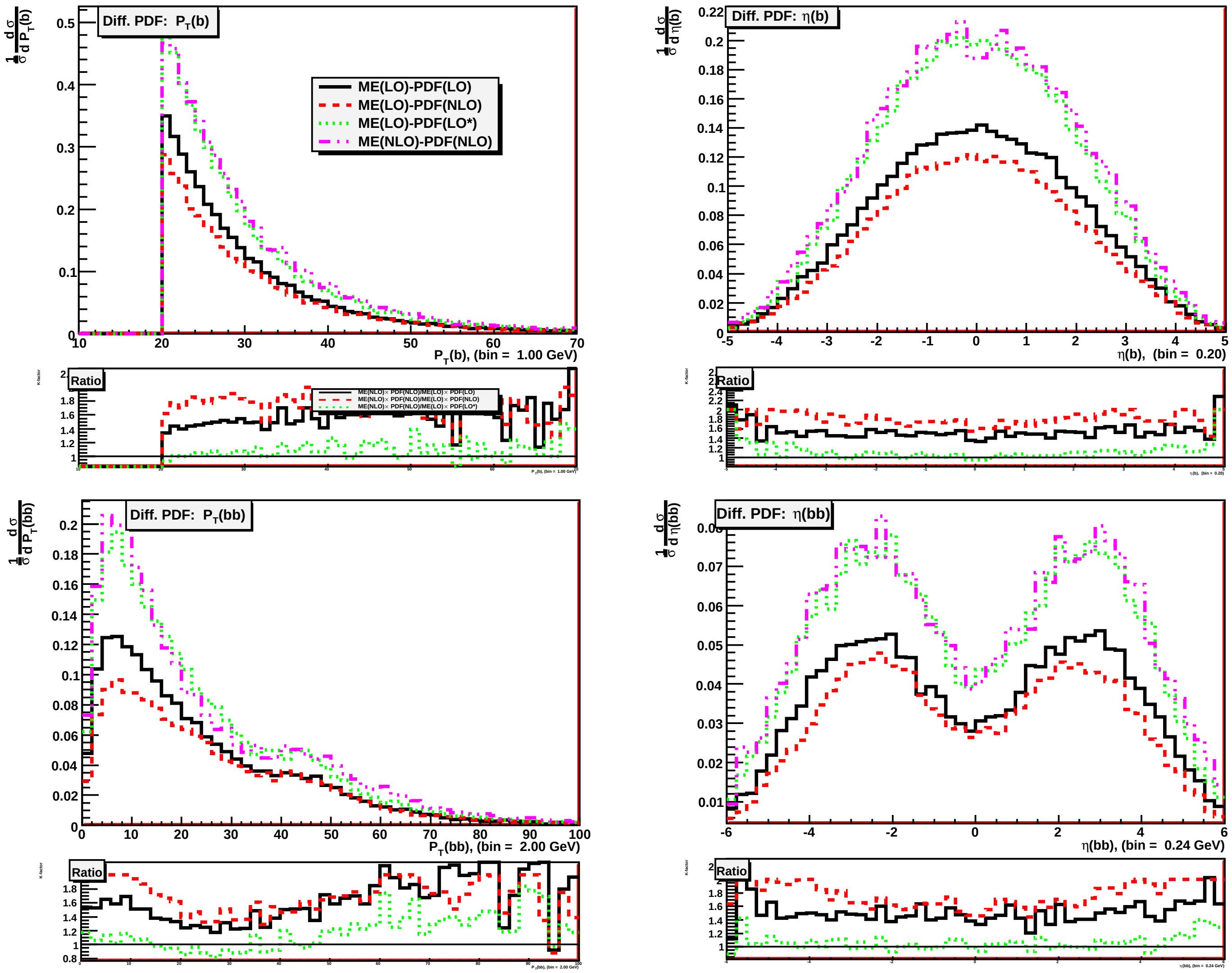
Leads to introduction of new type of LO* PDF.

NLO corrections to total cross-section usually positive $\rightarrow$ LO PDFs bigger by allowing momentum violation in global fits, using NLO α_S, fit LHC pseudo-data ......

Can also make evolution more "Monte Carlo like", e.g. change of scale in coupling.

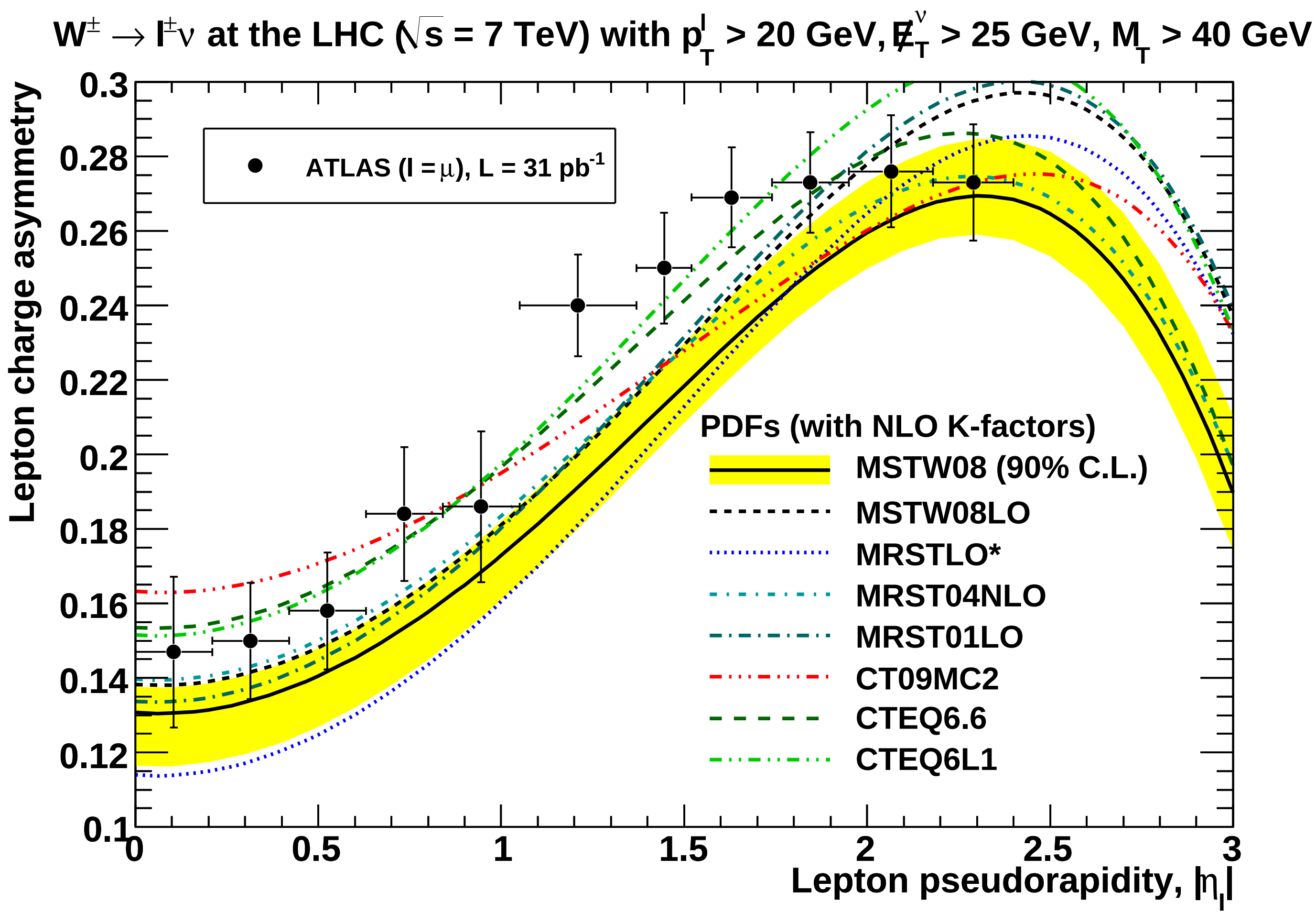
LO*, LO** PDFs from MRST and sets from CTEQ and very recently NNPDF

Example, look at e.g. distributions for single b and $b\bar{b}$ pair (Shertsnev, RT).

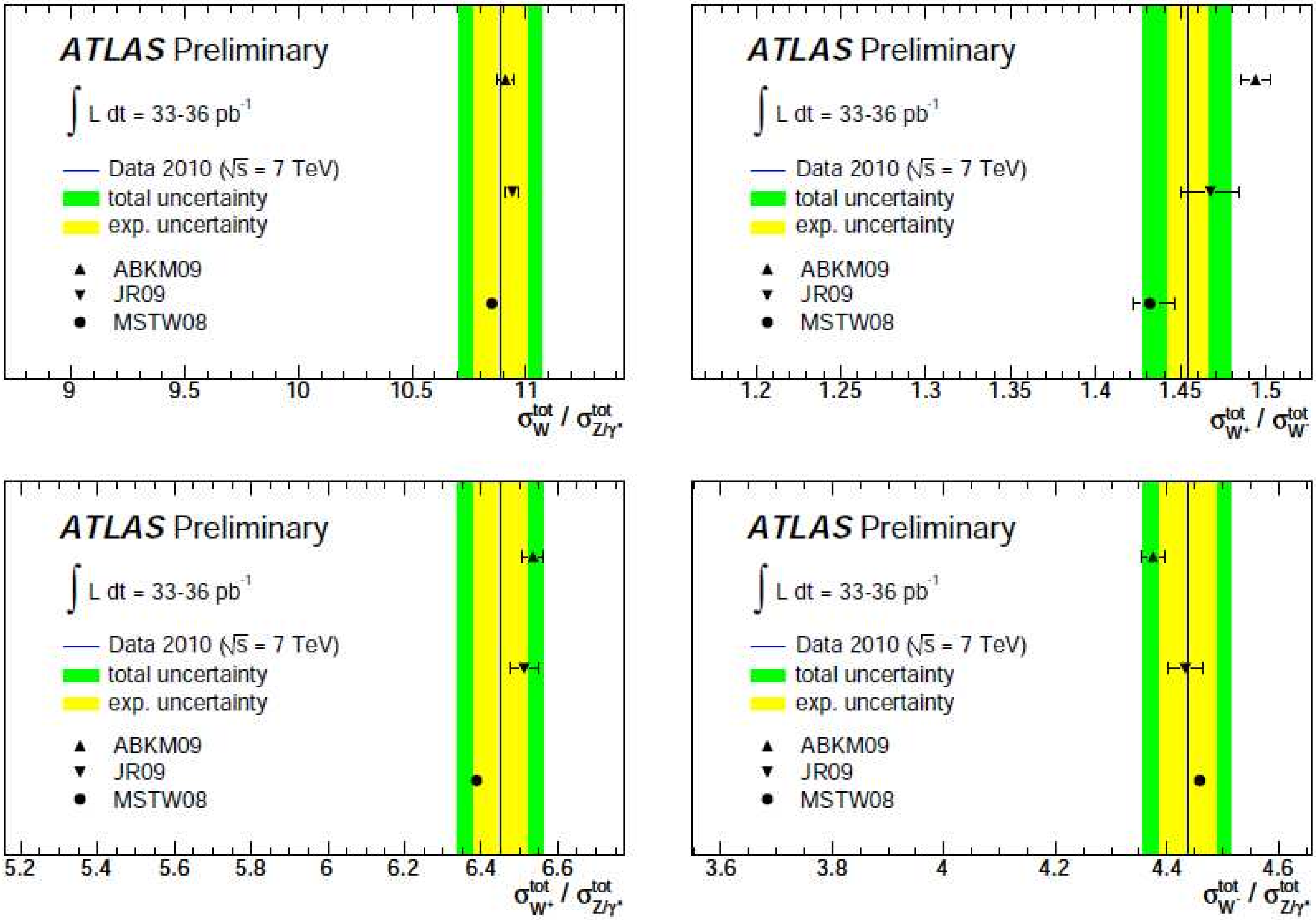


Results using LO* partons clearly best in normalization. NLO worst and problems with shape at low scales (i.e. small x).

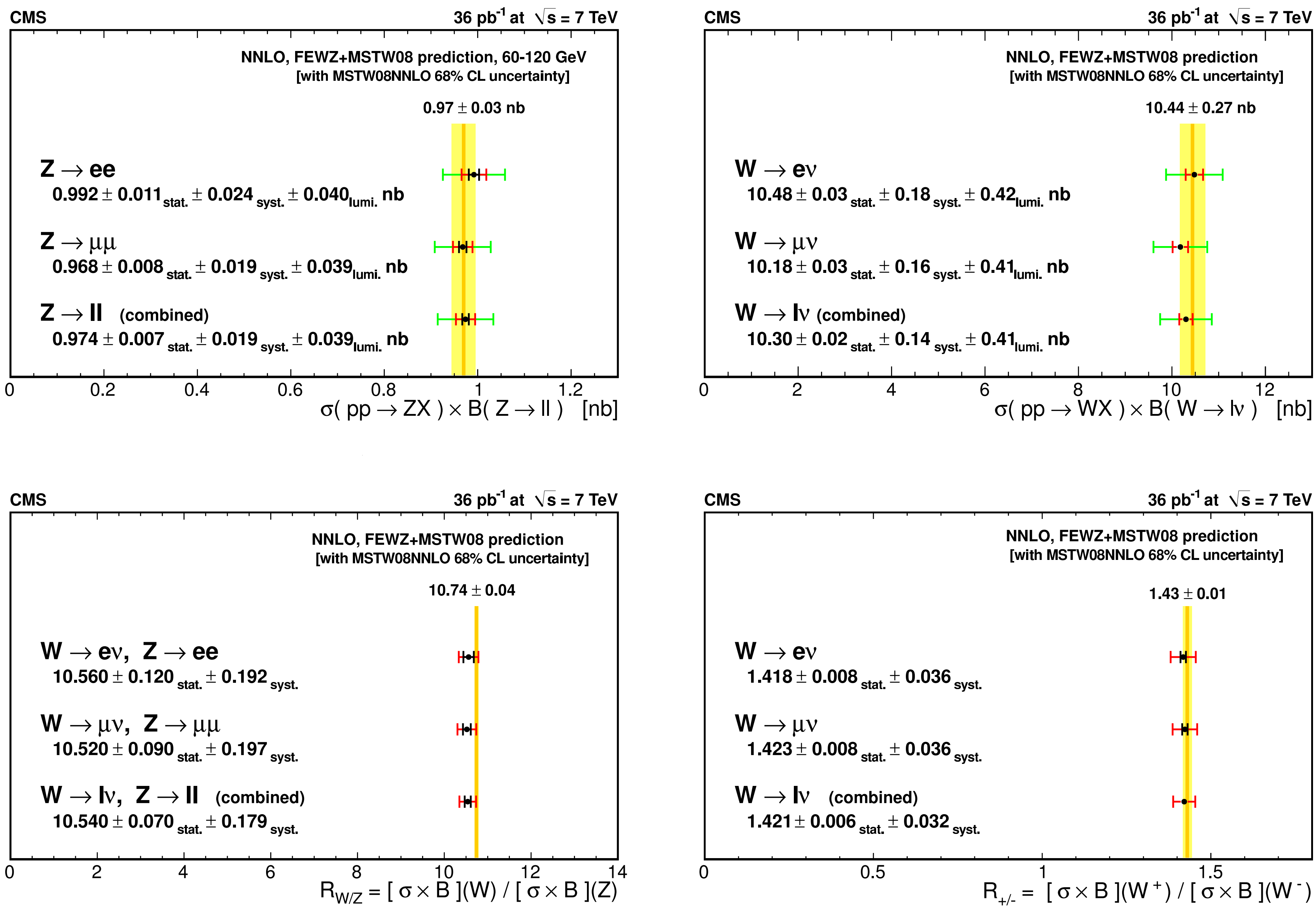
Acceptance Corrections – Watt



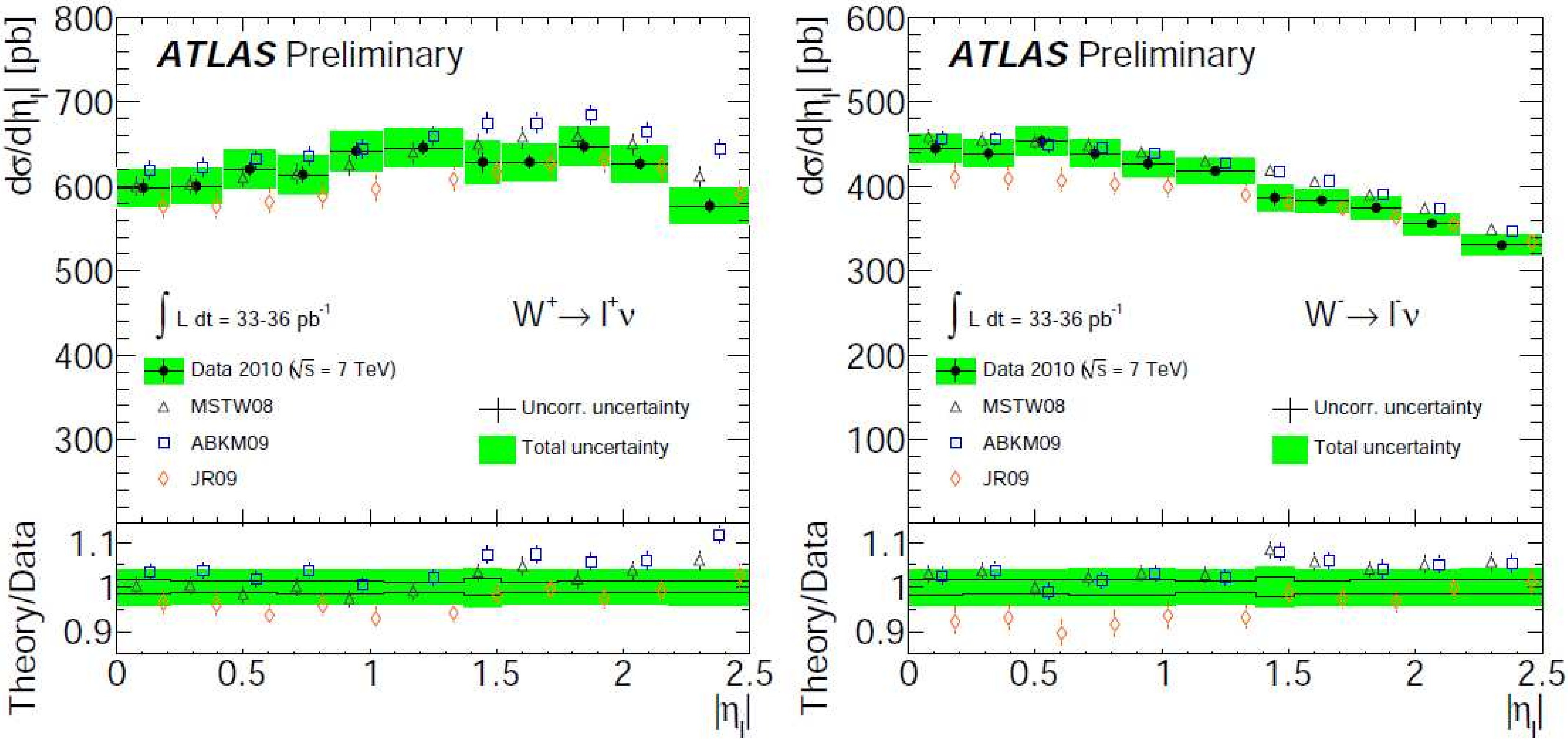
Need to be careful with precision quantities relying on flavour decomposition (Watt), especially if NLO corrections are available.



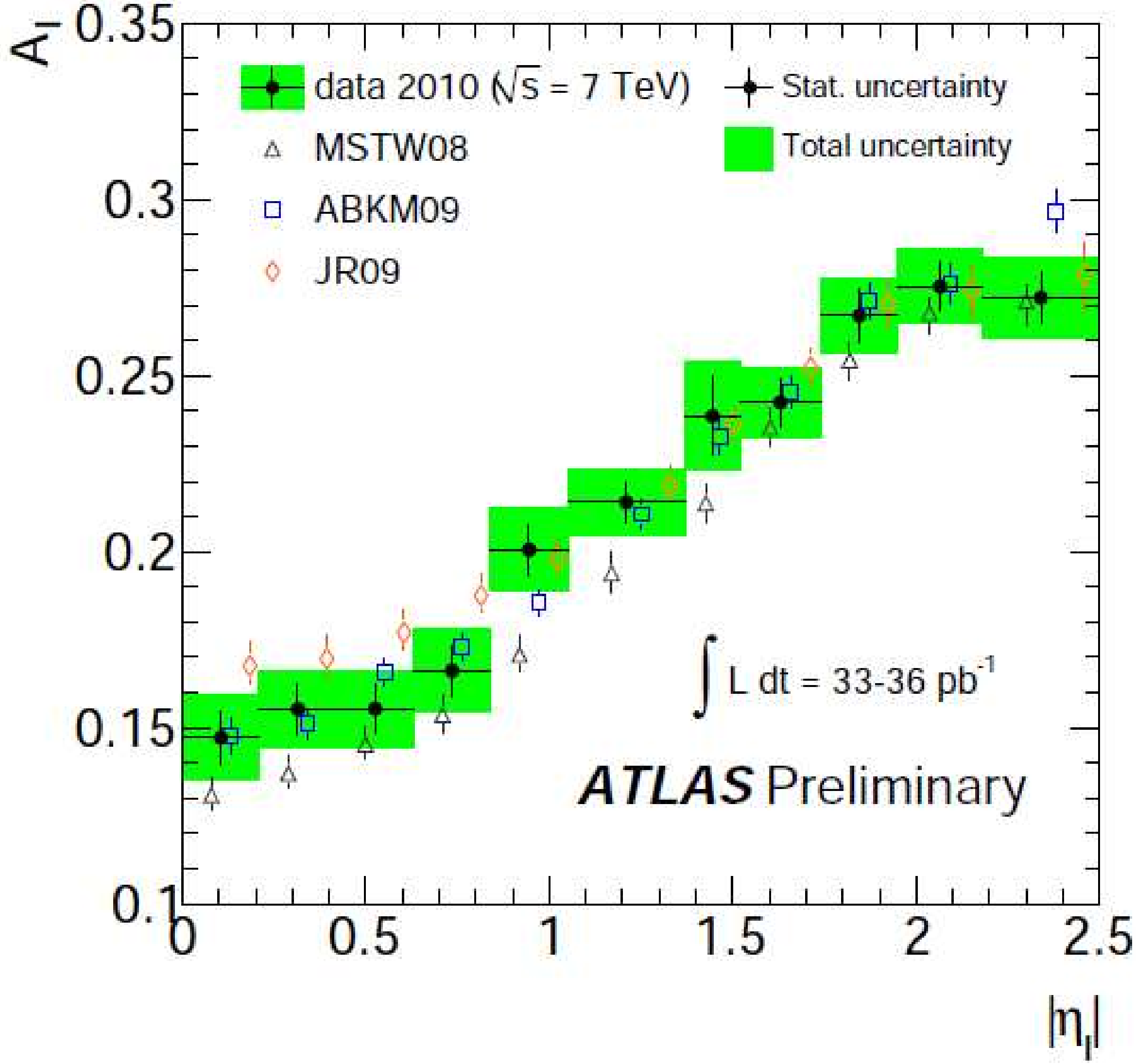
Example, earlier versions of ATLAS results implied some slightly different in ratios.



CMS results very similar.



Differential data on rapidity is becoming very constraining – on both shapes and on normalisations of predictions.



Clearly some of this information lost in ratios and asymmetries.

Ideally want individual distributions, with full correlations.

Inclusion of LHC data and reweighting

Reweighting
Adding new information without refitting

- The $N_{\rm rep}$ reps of a NNPDF fit give the probability density in the space of PDFs
- Expectation values for observables are Monte Carlo integrals
- One can study the effect of adding new data in the fit without refitting

$$\langle \mathcal{F}[f_i(x)] \rangle^{\rm UW} = \int [\mathcal{D}f_i]\, \mathcal{F}[f_i(x)]\mathcal{P}[f_i(x)] = \frac{1}{N_{\rm rep}} \sum_{k=1}^{N_{\rm rep}} \mathcal{F}[f_i^{(k)}(x)]$$

$$\langle \mathcal{F}[f_i(x)] \rangle^{\rm RW} = \int [\mathcal{D}f_i]\, \mathcal{F}[f_i(x)]\mathcal{P}_{\rm new}[f_i(x)] = \frac{1}{N_{\rm rep}} \sum_{k=1}^{N_{\rm rep}} w_k\, \mathcal{F}[f_i^{(k)}(x)]$$

- According to **Bayes Theorem** we have

$$\mathcal{P}_{\rm new}(\{f\}) = \mathcal{N}_\chi \mathcal{P}(\chi^2|\{f\})\mathcal{P}_{\rm init}(\{f\}), \quad \mathcal{P}(\chi^2|\{f\}) = [\chi^2(y,\{f\})]^{\frac{n_{dat}-1}{2}} e^{-\frac{\chi^2(y,\{f\})}{2}}$$

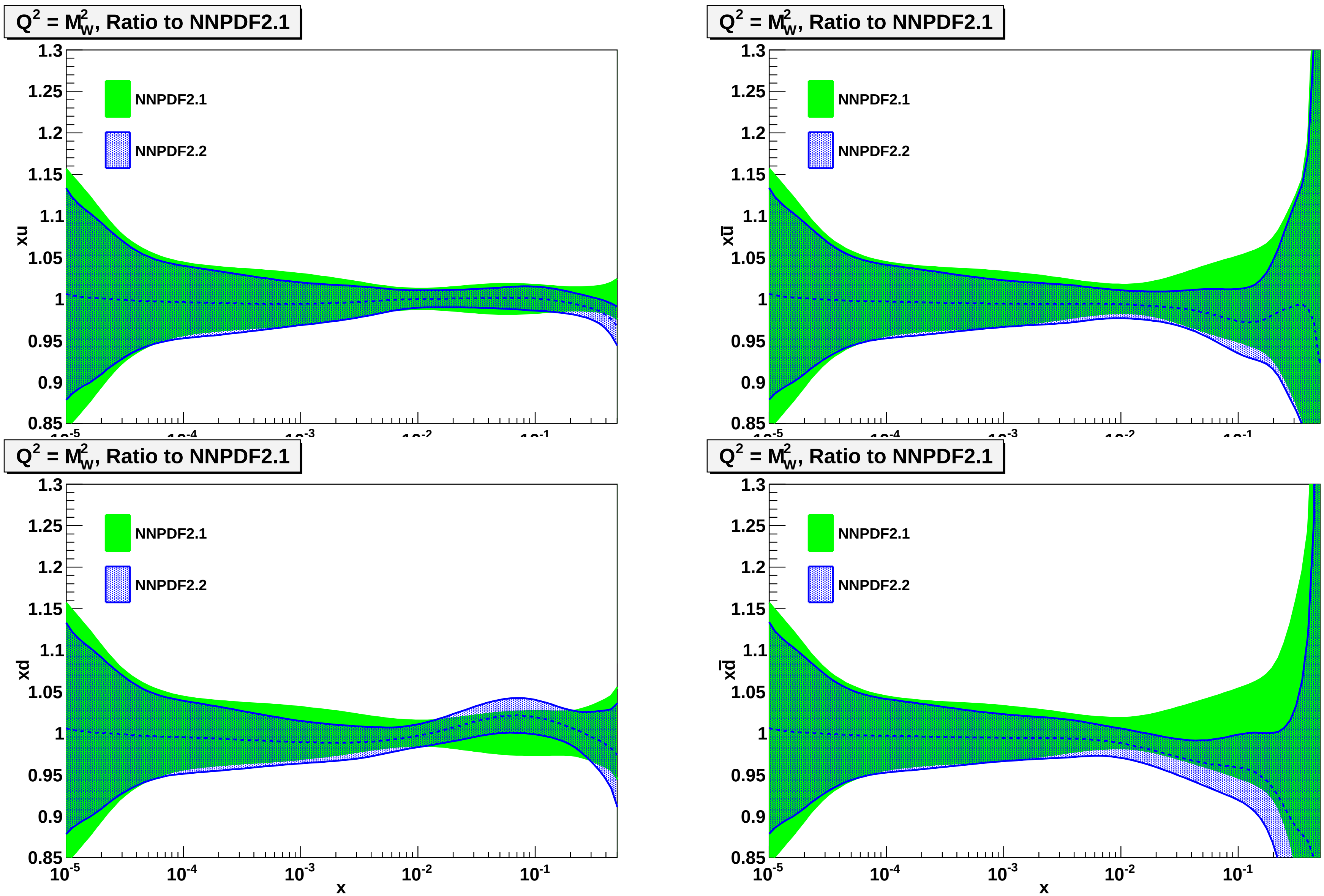
- **Averages over the sample** are now **weighted sums**

$$\langle \mathcal{F}[f_i(x,Q^2)] \rangle = \sum_{k=1}^{N_{rep}} w_k \mathcal{F}\left(f_i^{(net)(k)}(x,Q^2) \right)$$

where the **weights** are

$$w_k = \frac{[\chi^2(y,f_k)]^{\frac{n_{dat}-1}{2}} e^{-\frac{\chi^2(y,f_k)}{2}}}{\sum_{i=1}^{N_{\rm rep}} [\chi^2(y,f_i)]^{\frac{n_{dat}-1}{2}} e^{-\frac{\chi^2(y,f_i)}{2}}}$$

NNPDF have included the asymmetry data using reweighting, of PDFs and then unweighting (checking consistency with fitting directly).



The (fairly small) effect of the inclusion of new asymmetry data.

Details from single charged-lepton cross sections and asymmetries – Stirling

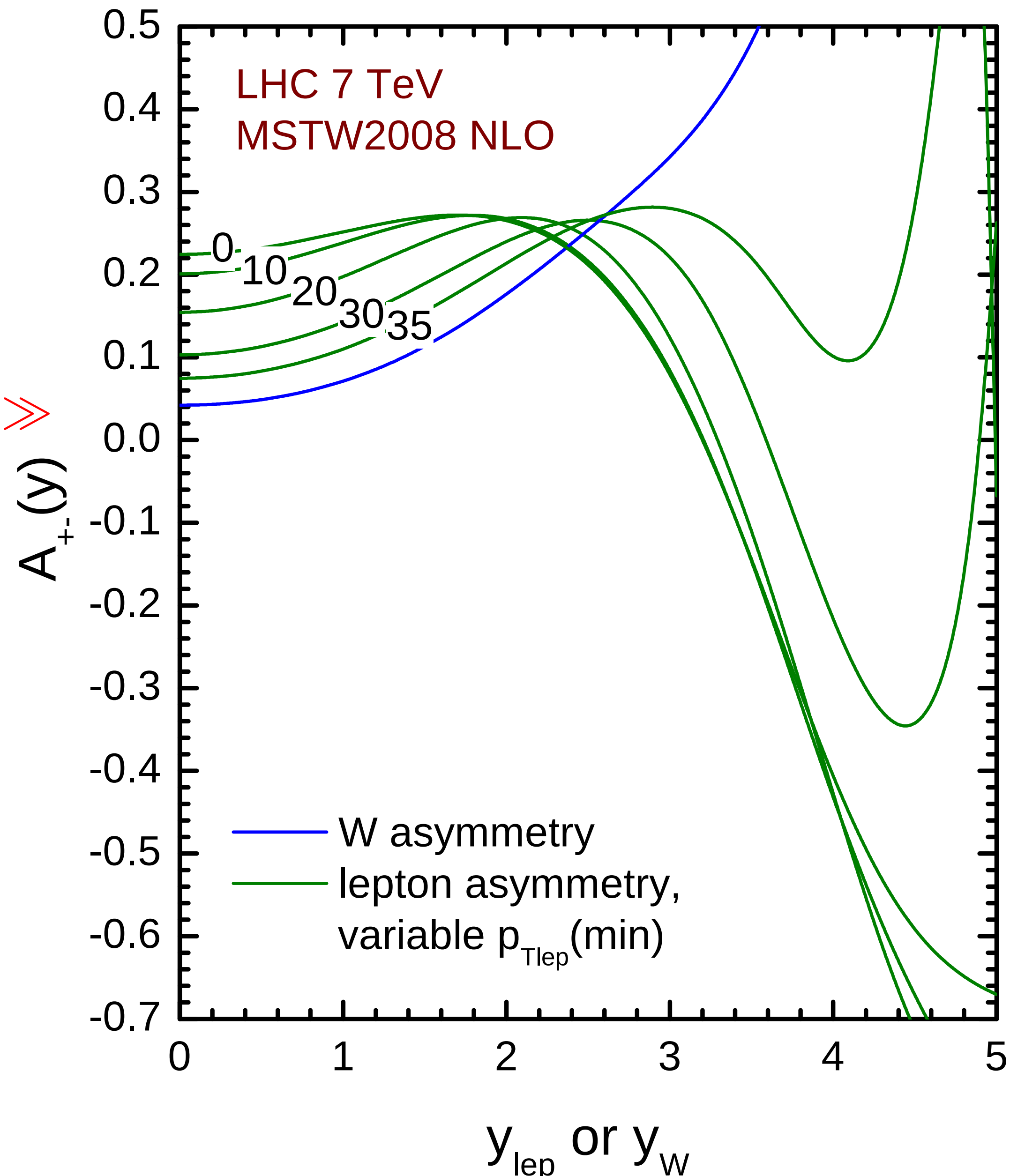
for low p_T main boost from W decay to leptons.

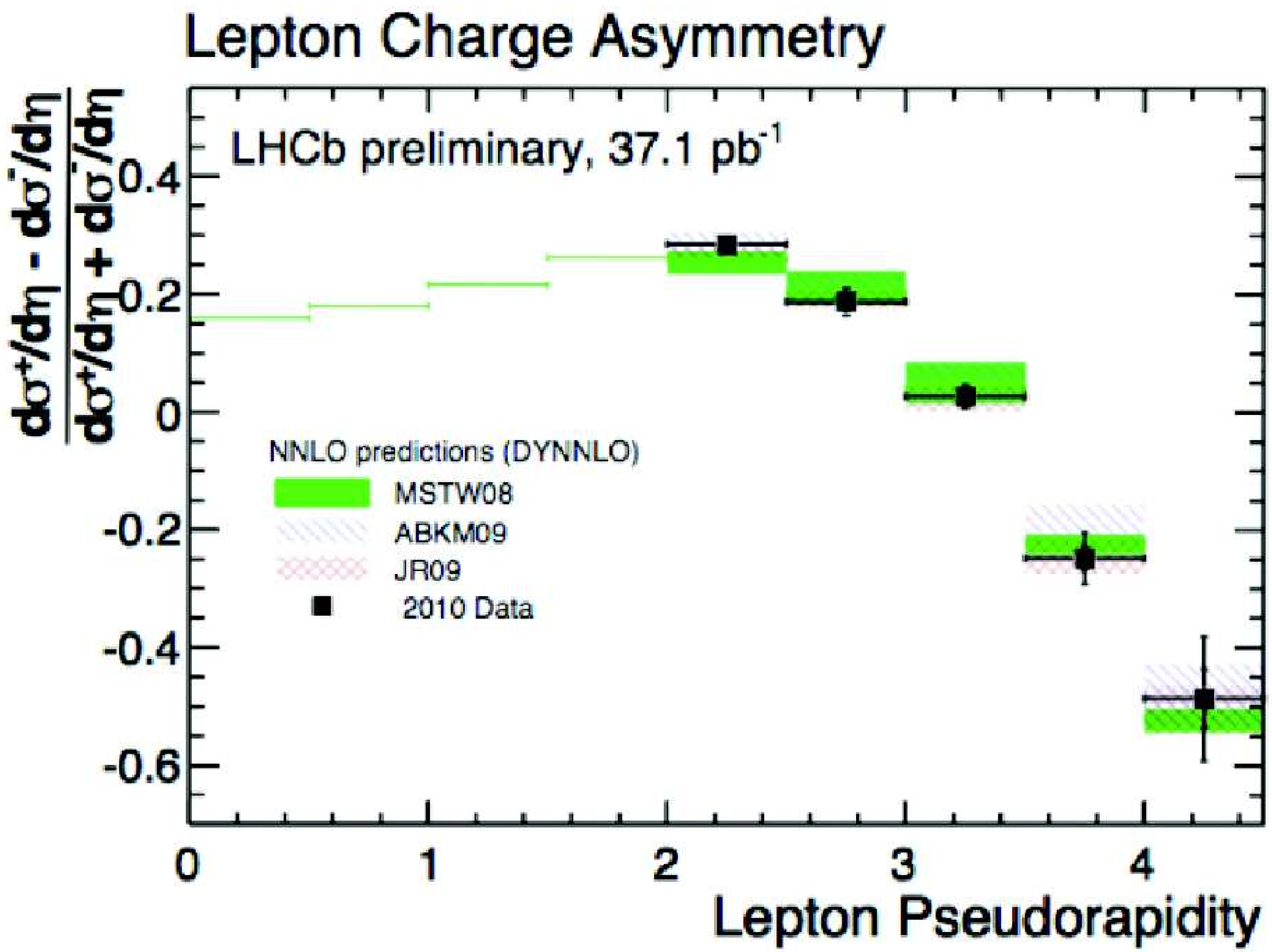
Dip towards -1 for lower p_T cuts from preferential forward production from $d_V(x_1)\bar{u}(x_2)$ due to axial vector nature of coupling.

Eventual turn-up when/if $u_V(x_1)\bar{d}(x_2) \gg d_V(x_1)\bar{u}(x_2)$

The larger the lepton p_T the earlier (in terms of increasing y_ℓ) this will happen, and for $p_T \to m_W/2$ there is no $V \pm A$ dominance at all.

So asymmetry at large y_ℓ in terms of p_T tells us about d/u at large x.





LHCb (with $p_T(\mathrm{min}) = 20\mathrm{GeV}$ already testing dip.

With higher $p_T(\mathrm{min})$ could potentially see upturn.

Conclusions

One can determine the parton distributions and predict cross-sections at the LHC, and the fit quality using NLO or NNLO QCD is fairly good. Nearly full range of NNLO PDFs now. Comparison between different PDF sets at NLO and NNLO very similar.

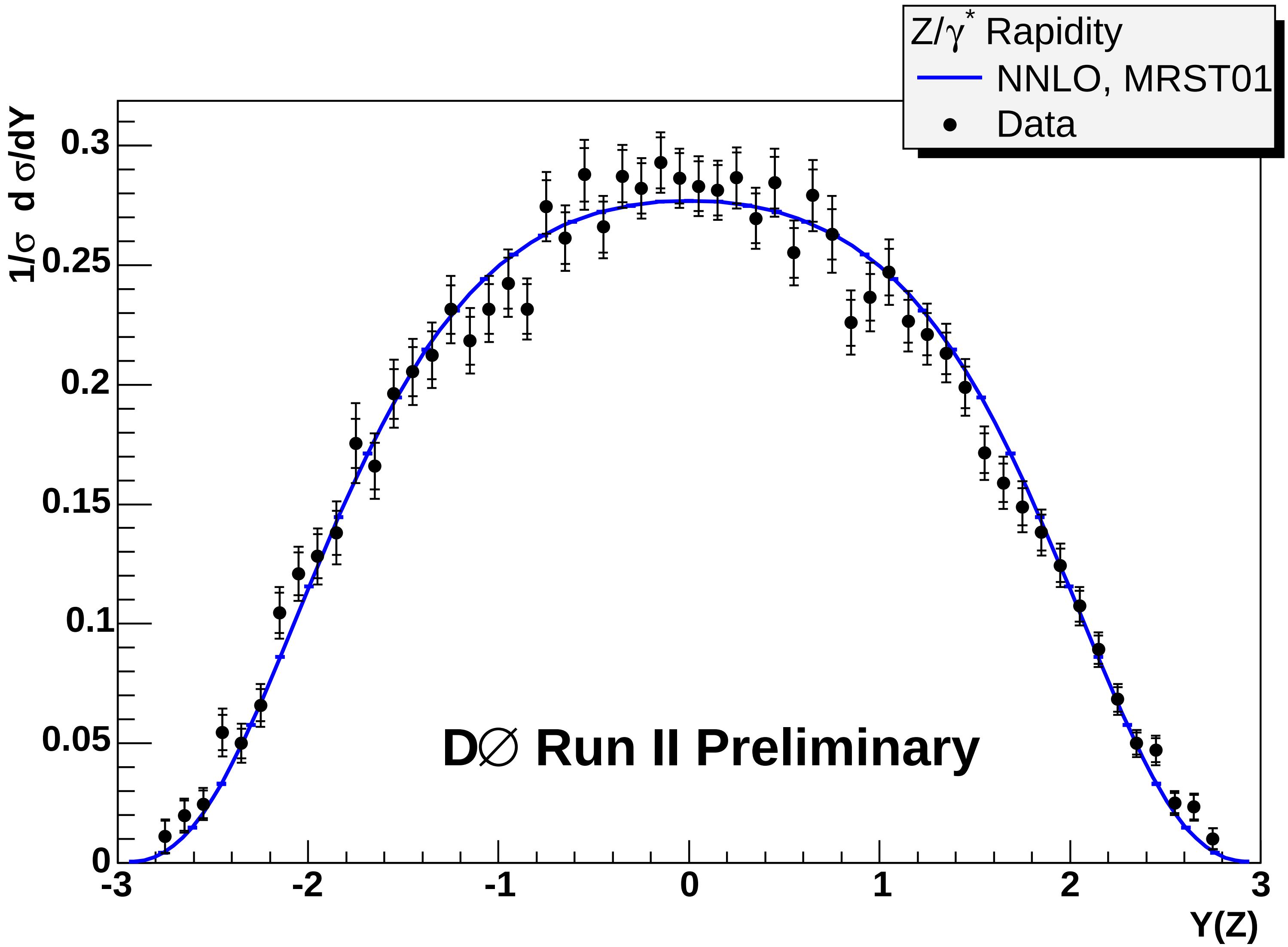
Various ways of looking at $experimental$ uncertainties. Uncertainties $\sim 1-5\%$ for most LHC quantities. Ratios, e.g. W^+/W^- tight constraint on partons, but don't want to lose information when taking ratios.

Effects from input assumptions e.g. selection of data fitted, cuts and input parameterisation can shift central values of predictions significantly. Also affect size of uncertainties. Want balance between freedom and sensible constraints.

Data from the LHC just starting to have some effect on improving the precision of PDFs. Might start to discriminate between PDFs first.

Extraction of PDFs from existing data and use for LHC far from a straightforward procedure. Lots of issues to consider for real precision. Relatively few cases where Standard Model discrepancies will not require some significant input from PDF physics to determine real significance.

Excellent predictive power – comparison of MRST prediction for Z rapidity distribution with preliminary data.



Interplay of LHC and pdfs/QCD

Make predictions for all processes, both SM and BSM, as accurately as possible given current experimental input and theoretical accuracy.

Check against well-understood processes, e.g. central rapidity W, Z production (luminosity monitor), lowish-E_T jets, ...

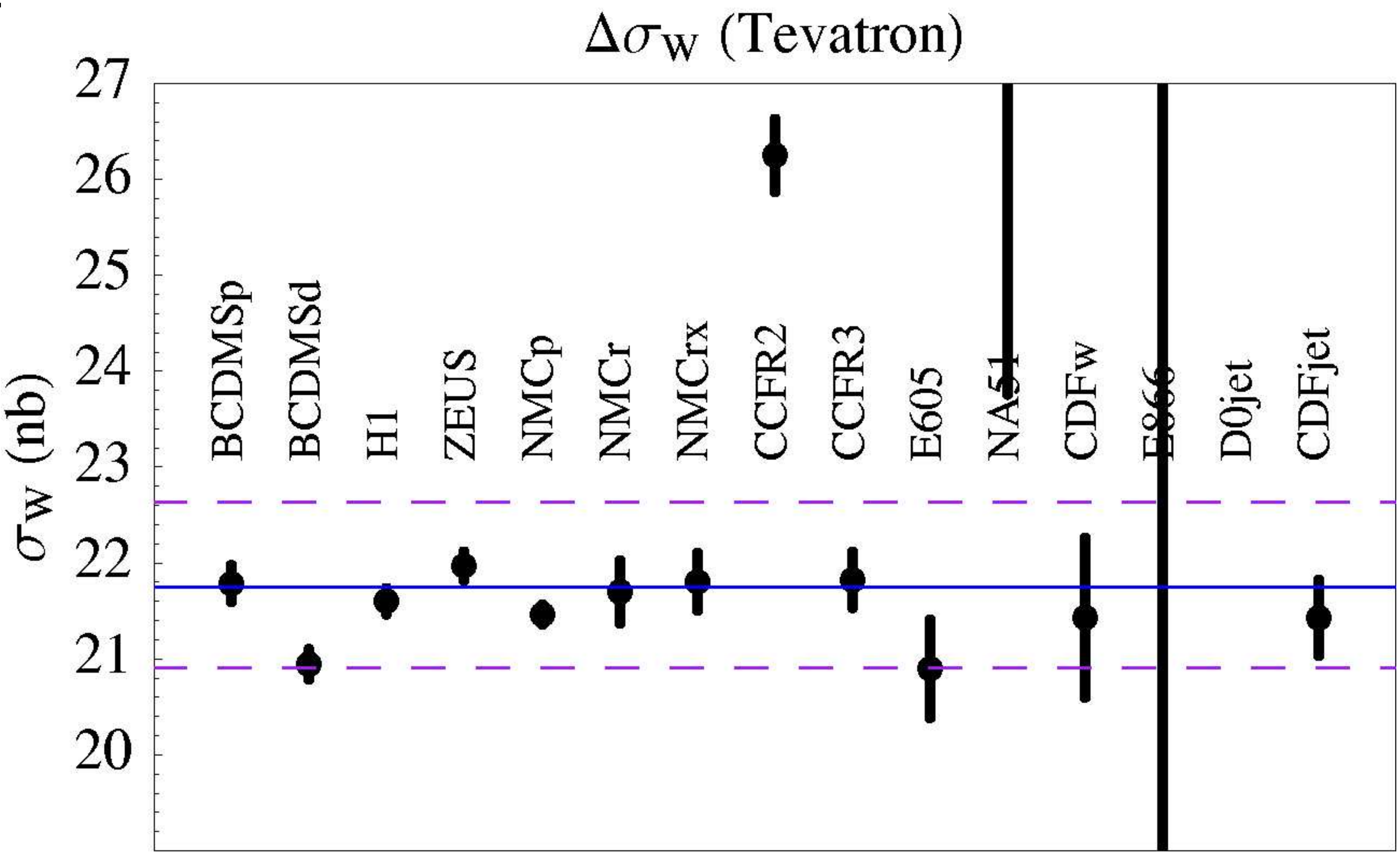
Compare with predictions with more uncertainty and lower confidence, e.g. high-E_T jets, high rapidity bosons or heavy quarks .....

Improve uncertainty on parton distributions by improved constraints, and check understanding of theoretical uncertainties, and determine where NNLO, electroweak corrections, resummations $etc.$ needed.

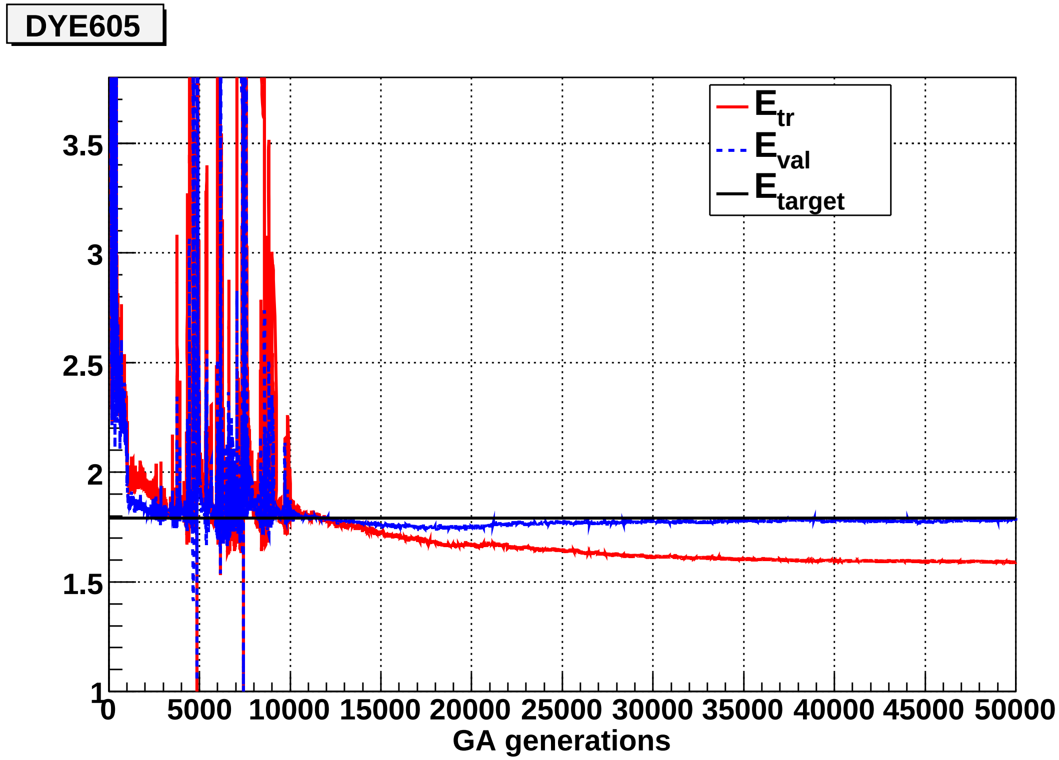
Make improved predictions for both background and signals with improved partons and surrounding theory.

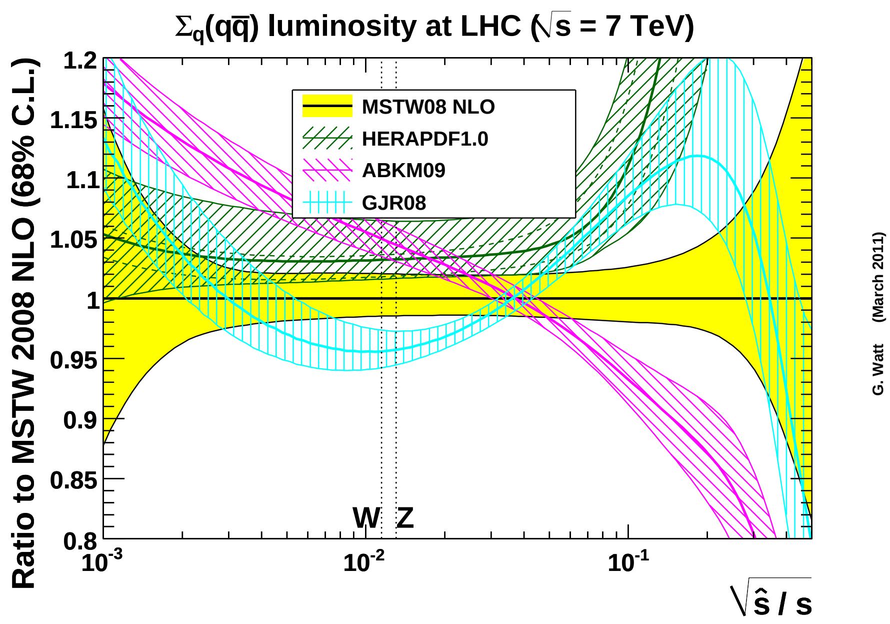
Spot new physics from deviations in these predictions. As a nice by-product improve our understanding of the strong sector of the Standard Model considerably.

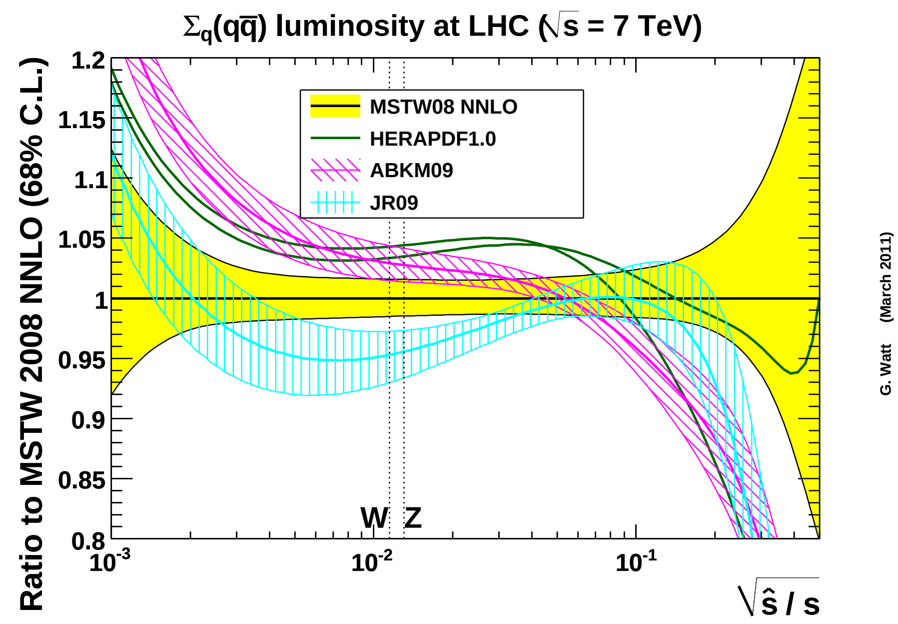
The inappropriateness of using $\Delta\chi^2 = 1$ when including a large number of sometimes conflicting data sets is shown by examining the best value of σ_W and its uncertainty using $\Delta\chi^2 = 1$ for individual data sets as obtained by CTEQ using Lagrange Multiplier technique.



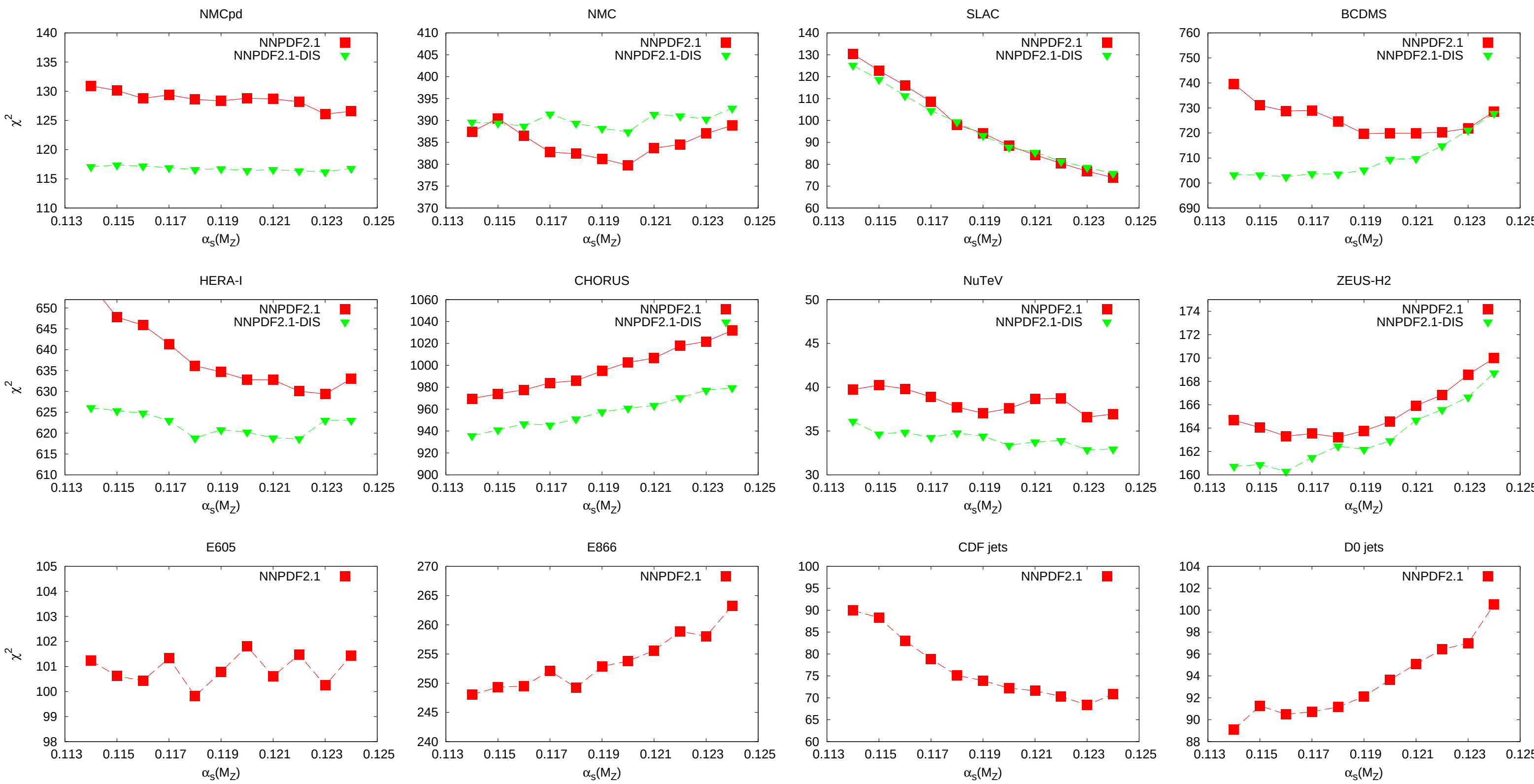
Difficult to know when fit to validation set has started increasing significantly for some sets.



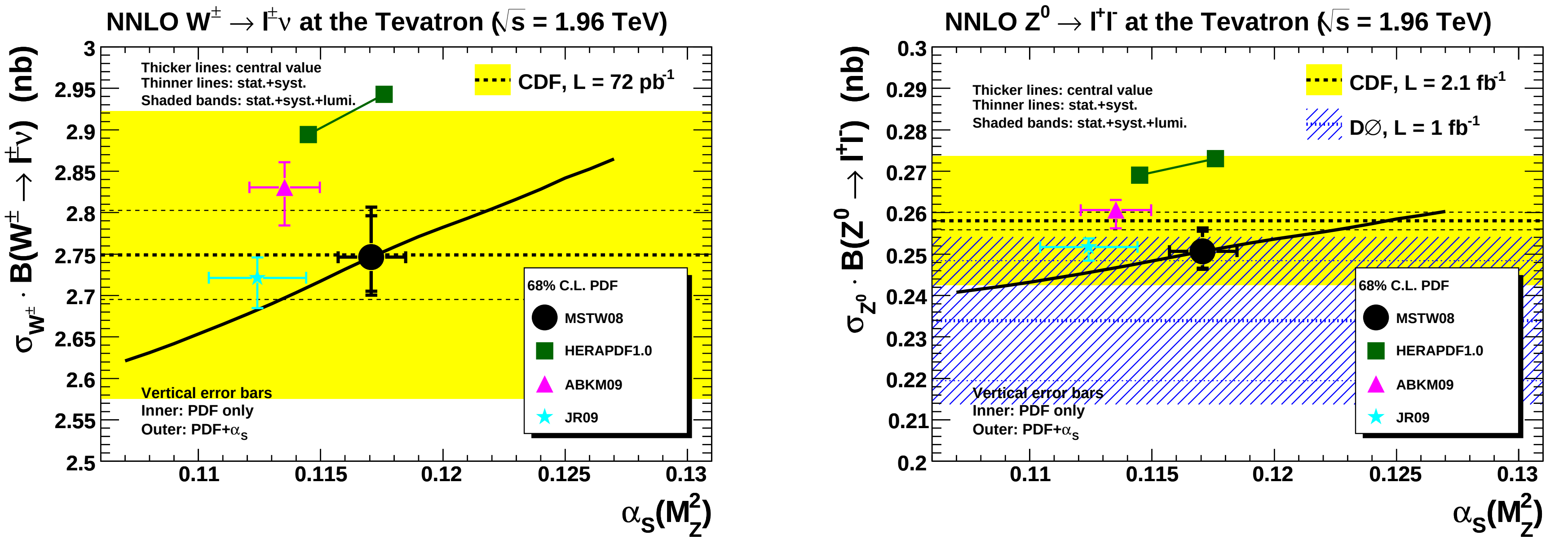




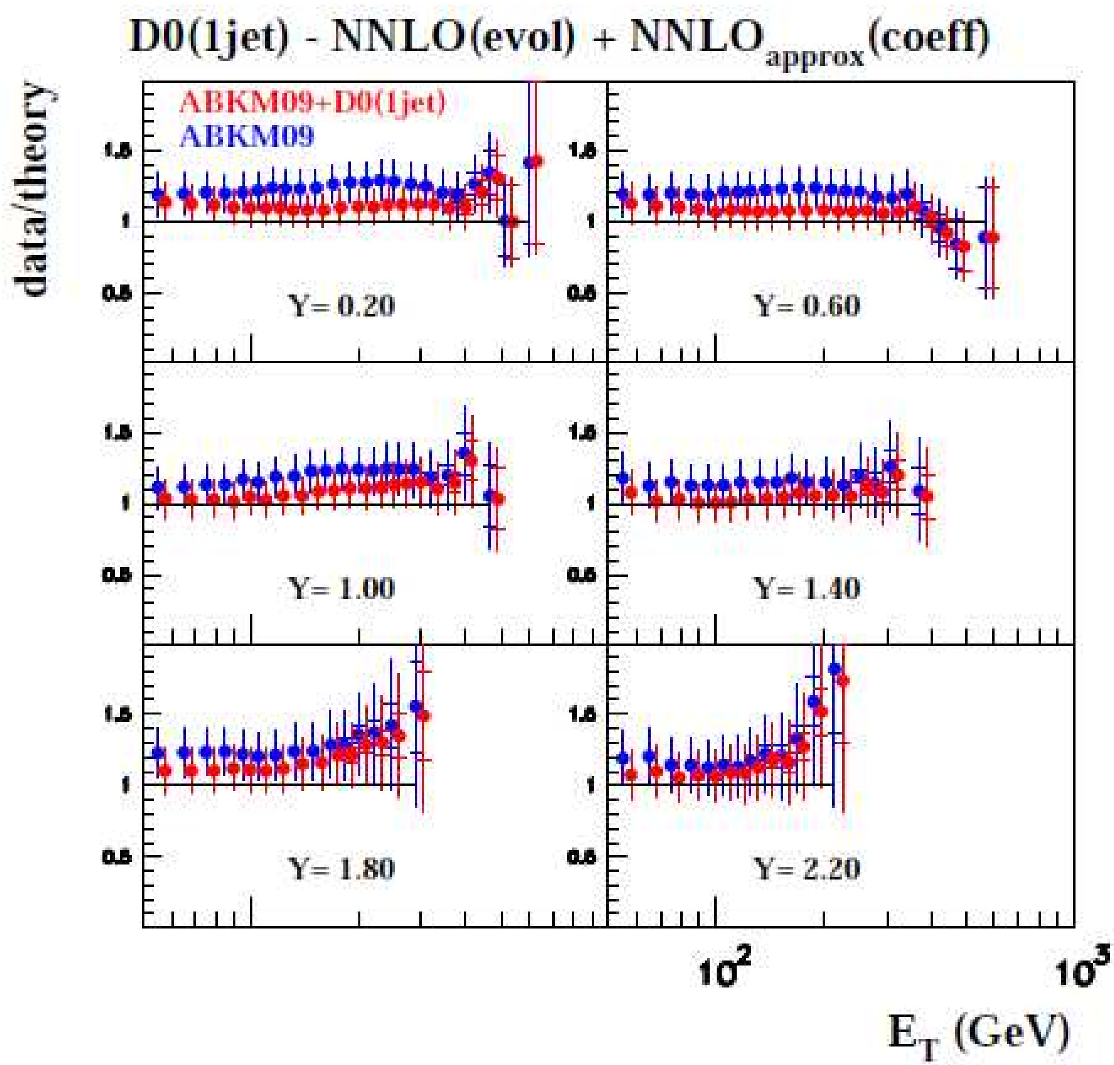
Not all luminosity differences the same at NLO as at NNLO, e.g. HERAPDF $q\bar{q}$.


NMCpd
NMC
SLAC
BCDMS
HERA-I
CHORUS
NuTeV
ZEUS-H2
E605
E866
CDF jets
D0 jets
NNPDF2.1
NNPDF2.1-DIS
χ^2
$\alpha_s(M_Z)$

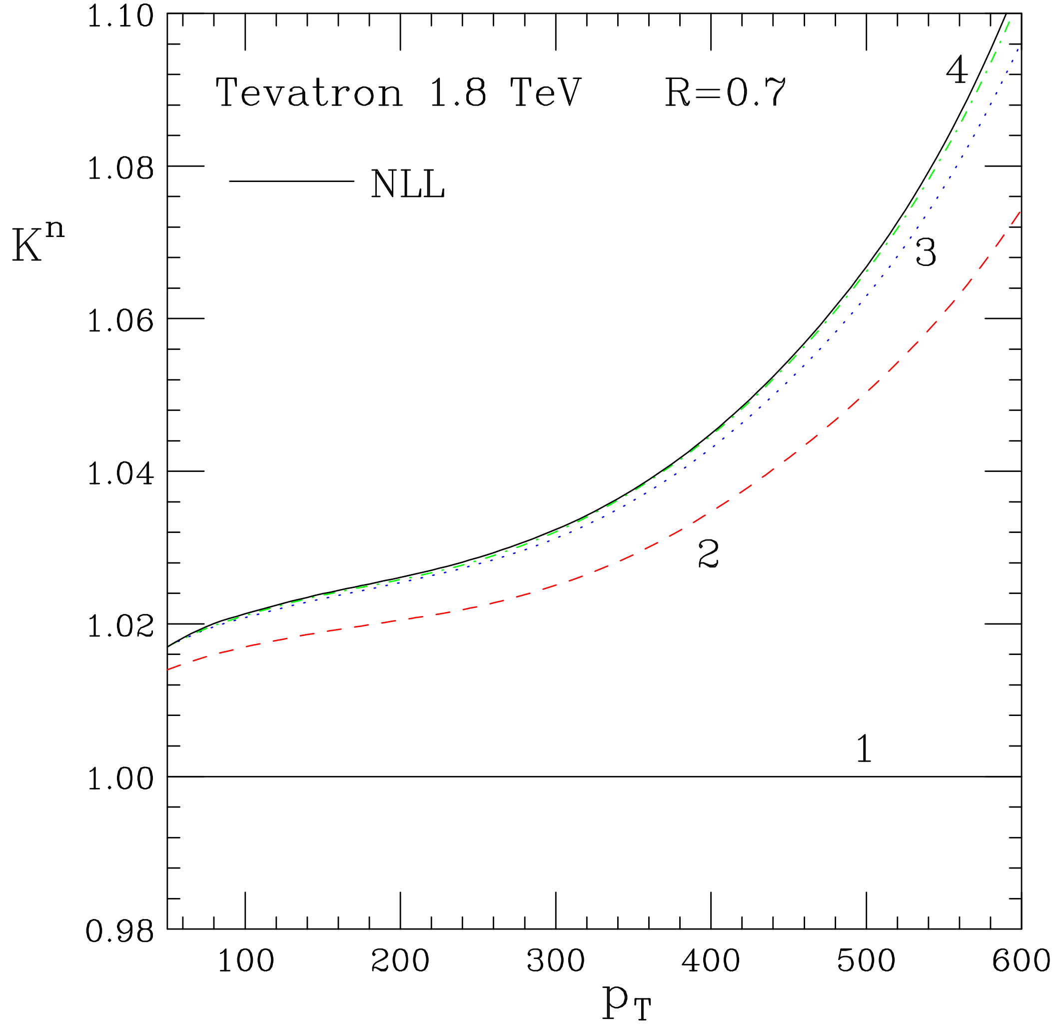
Different PDF predictions for W and Z cross sections at the Tevatron compared to data.



ABKM below data, and to a lesser extent still generally low after fit.



de Florian and Vogelsang result for inclusive jet K-factor for $d\sigma/dp_T$ at order α_S^{2+n} compared to NLO.

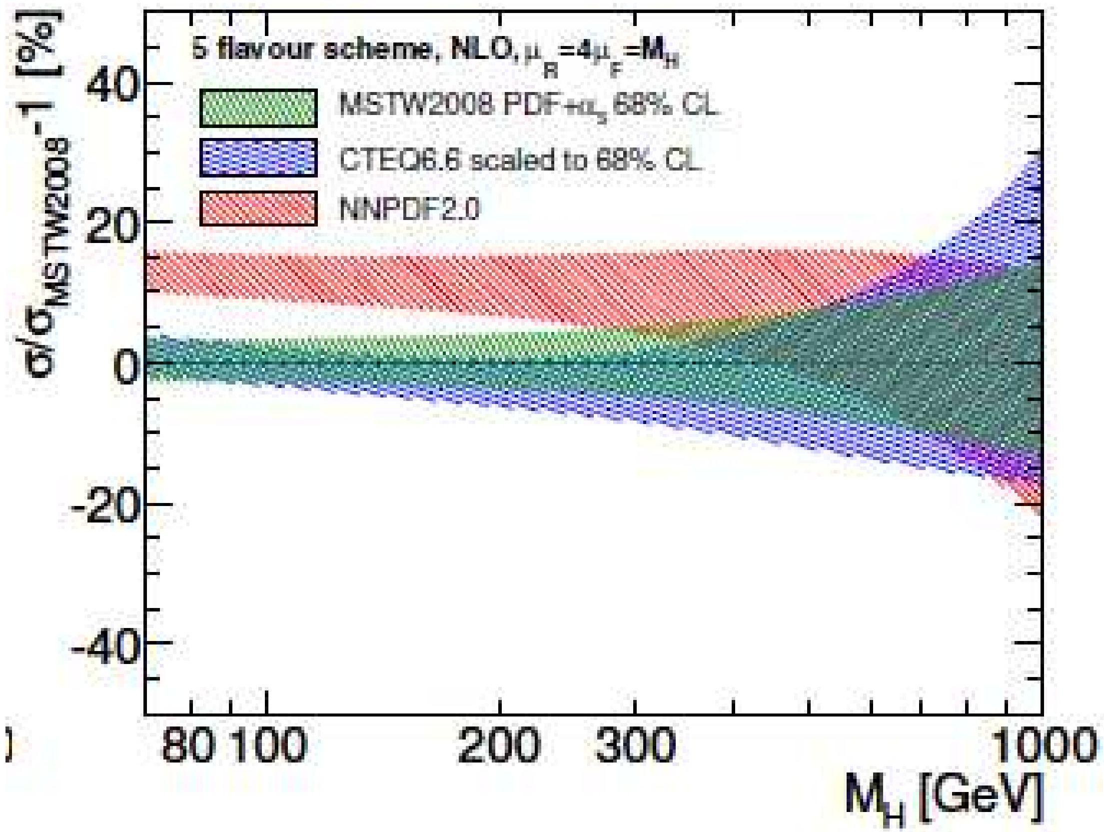


Sometimes the reason for cross section differences is unexpected.

Warsinsky at recent Higgs-LHC working group meeting.

m_b values bring CTEQ and MSTW together but exaggerate NNPDF difference.

Couplings have assumed common mass value.



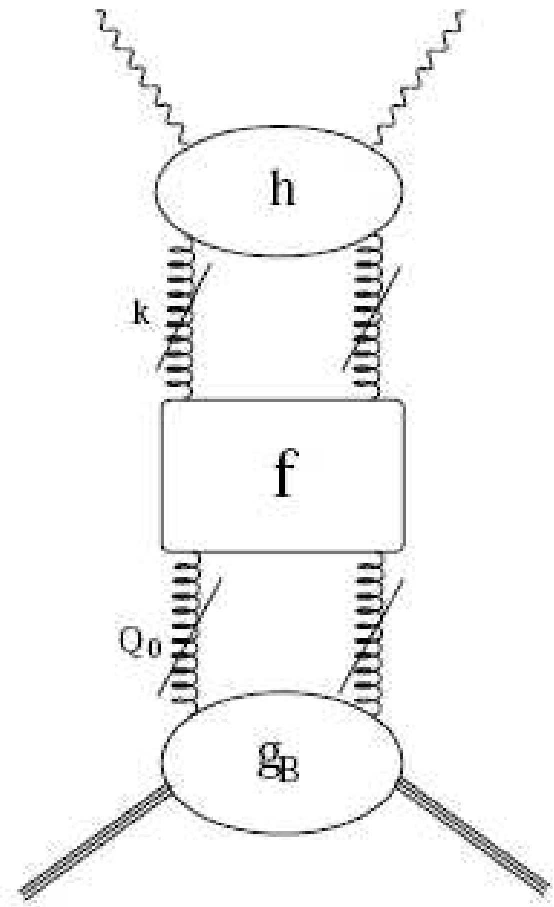
Small-x Theory

Reason for this instability – at each order in α_S each splitting function and coefficient function obtains an extra power of $\ln(1/x)$ (some accidental zeros in P_{gg}), i.e. $P_{ij}(x,\alpha_s(Q^2)), C_i^P(x,\alpha_s(Q^2)) \sim \alpha_s^m(Q^2)\ln^{m-1}(1/x)$.

BFKL equation for high-energy limit

$$f(k^2,x) = f_I(Q_0^2) + \int_x^1 \frac{dx'}{x'}\bar{\alpha}_S \int_0^\infty \frac{dq^2}{q^2} K(q^2,k^2) f(q^2,x),$$

where $f(k^2,x)$ is the unintegrated gluon distribution $g(x,Q^2) = \int_0^{Q^2} (dk^2/k^2) f(x,k^2)$, and $K(q^2,k^2)$ is a calculated kernel known to NLO.



Physical structure functions obtained from

$$\sigma(Q^2,x) = \int (dk^2/k^2)\, h(k^2/Q^2) f(k^2,x)$$

where $h(k^2/Q^2)$ is a calculable impact factor.

The global fits usually assume that this is unimportant in practice, and proceed regardless.

Fits work well at small x, but could improve.

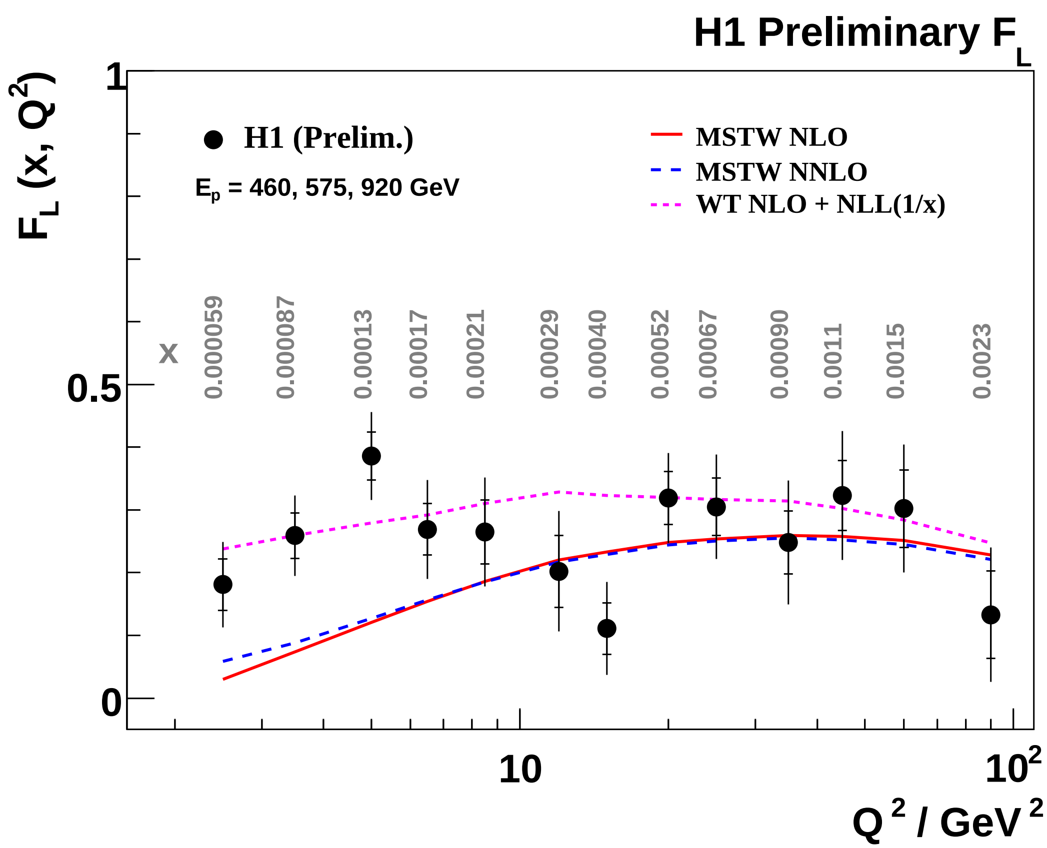
Small-x Theory

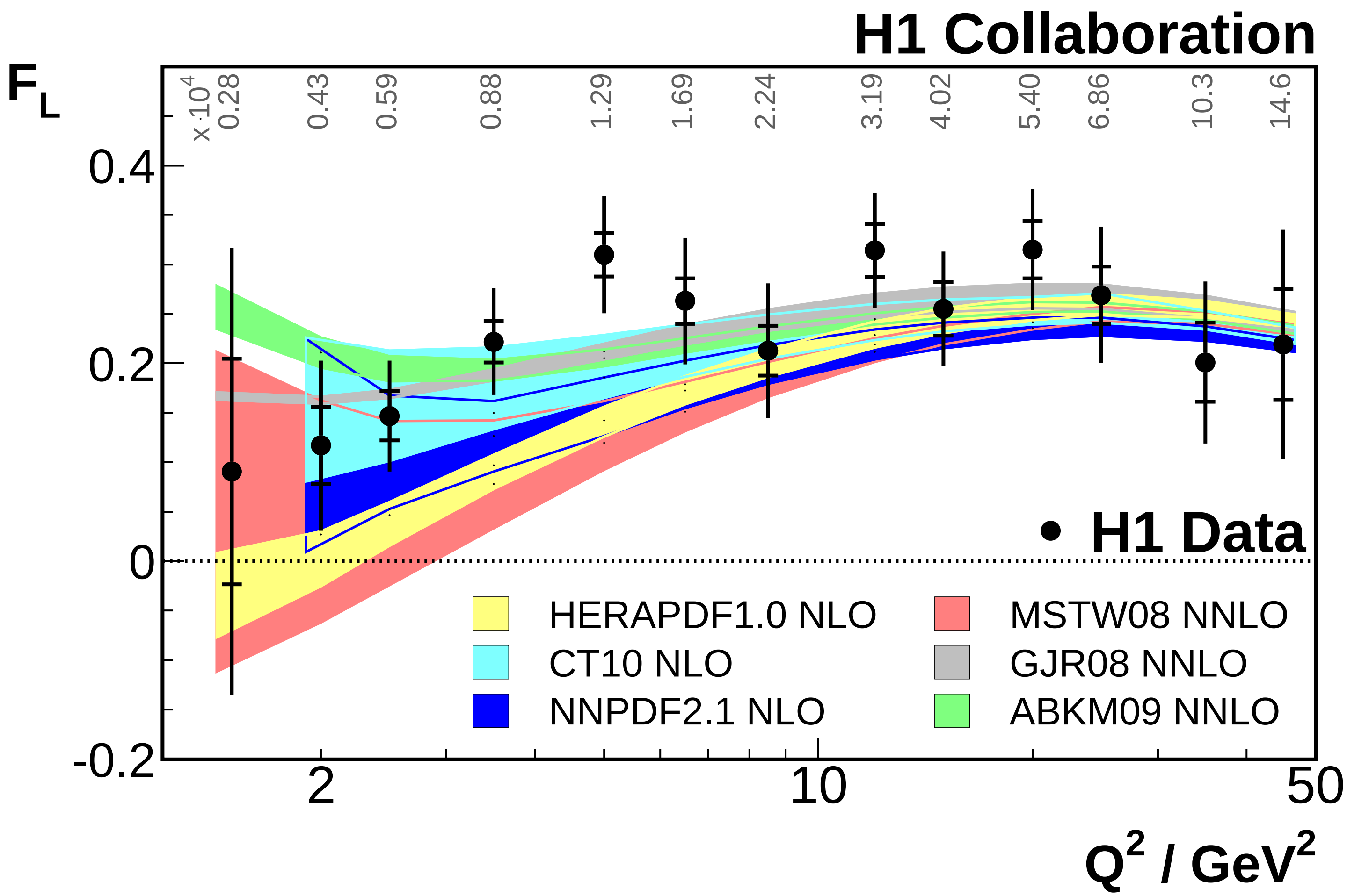
At each order in α_S each splitting function and coefficient function obtains an extra power of $\ln(1/x)$ (some accidental zeros in P_{gg}), i.e. $P_{ij}(x, \alpha_s(Q^2)), C_i^P(x, \alpha_s(Q^2)) \sim \alpha_s^m(Q^2) \ln^{m-1}(1/x)$.

Summed using BFKL equation (and a lot of work – Altarelli-Ball-Forte, Ciafaloni-Colferai-Salam-Stasto and White-RT)

Comparison to H1 prelim data on $F_L(x, Q^2)$ at low Q^2, only within White-RT approach, suggests resummations may be important.

Could possibly give a few percent effect on Higgs cross sections.





However, quite a large PDF uncertainty (in general) and even larger spread, at fixed order (though differences in definition of order).

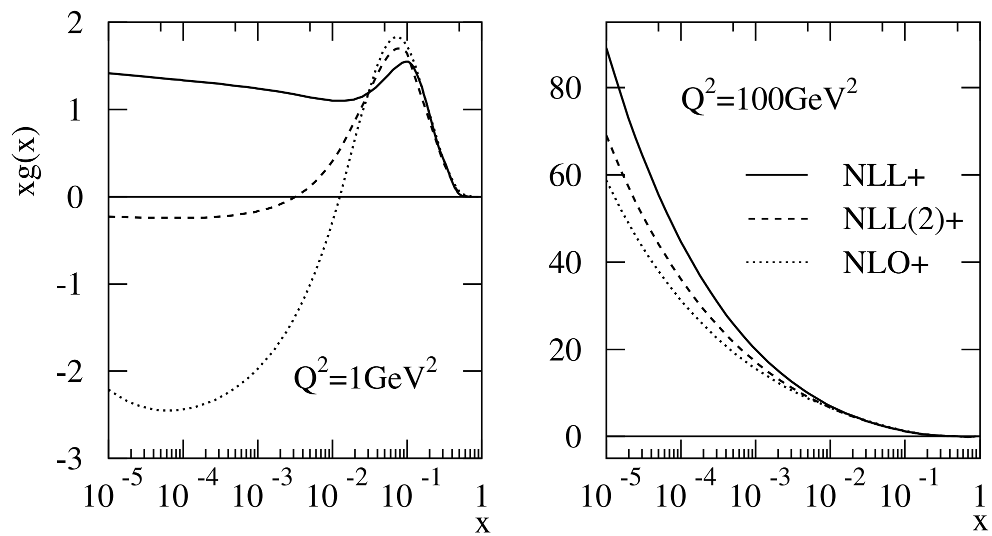
Good recent progress in incorporating $\ln(1/x)$ resummation Altarelli-Ball-Forte, Ciafaloni-Colferai-Salam-Stasto and White-RT.

Include running coupling effects and variety (depending on group) of other corrections

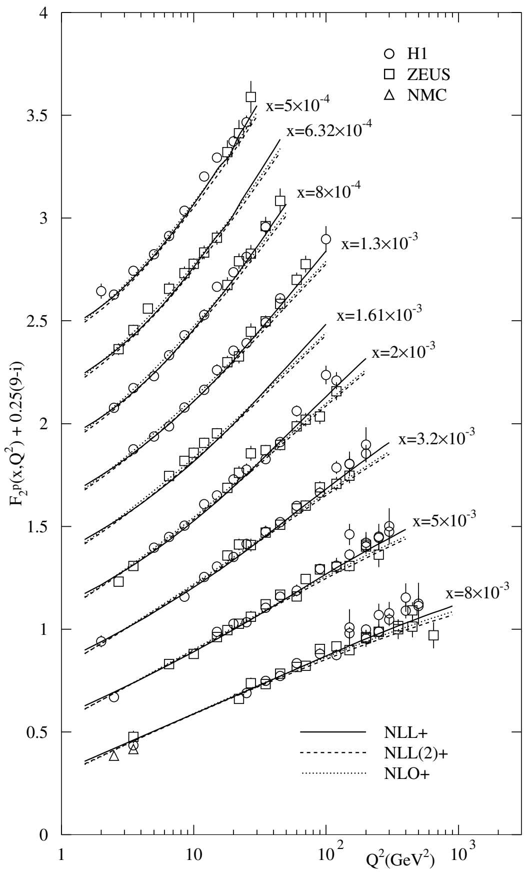
By 2008 very similar results coming from the competing procedures, despite some differences in technique.

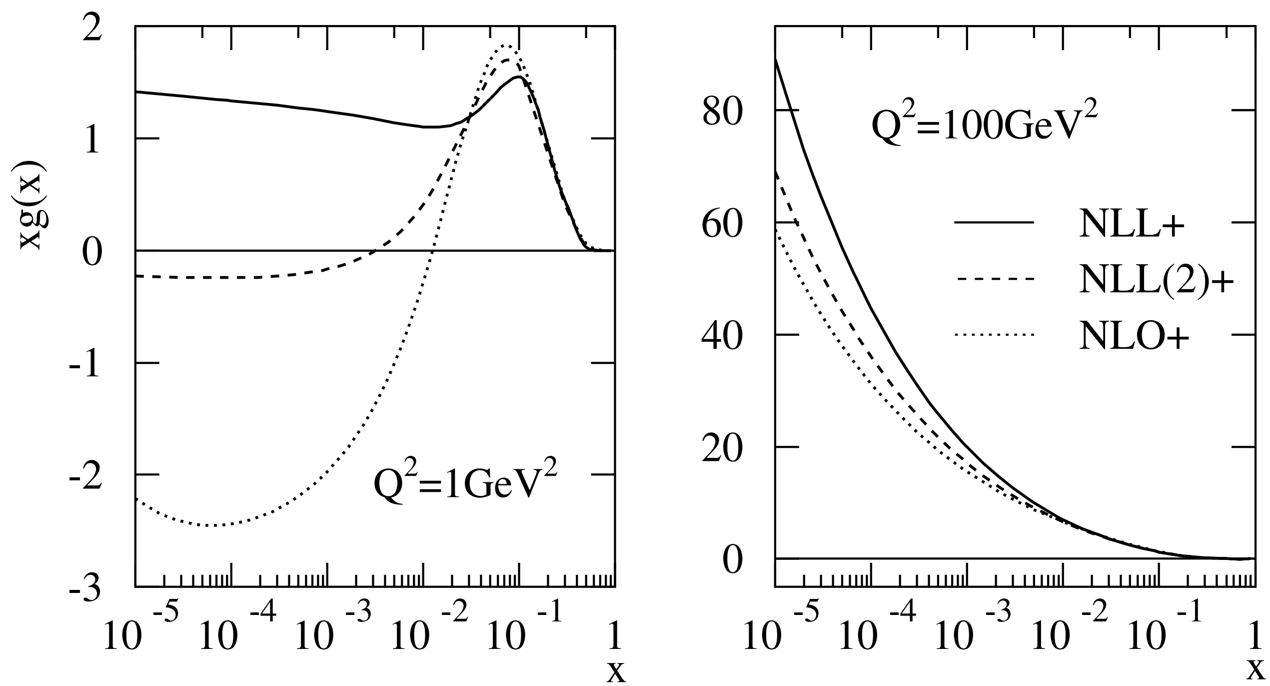
Full set of coefficient functions still to come in some cases, but splitting functions comparable.

Note, in all cases NLO corrections lead to dip in functions below fixed order values until slower growth (running coupling effect) at very small x.



A fit to data with NLO plus NLO resummation, with heavy quarks included (White,RT) performed.





$\rightarrow$ moderate improvement in fit to HERA data within global fit, and change in extracted gluon (more like quarks at low Q^2).

Together with indications from Drell Yan resummation calculations (Marzani, Ball) few percent effect quite possible.

Low Q^2.

Perform fits with the known NNNLO large $\ln(1-x)$ terms included explicitly.

Also parameterize higher twist contributions by

$$F_i^{\mathrm{HT}}(x,Q^2) = F_i^{\mathrm{LT}}(x,Q^2)\left(1+\frac{D_i(x)}{Q^2}\right)$$

where i spans bins of x.

No evidence for any higher twist except at low W^2.

x	LO	NLO	NNLO	NNNLO
0–0.0005	-0.07	-0.02	-0.02	-0.03
0.0005–0.005	-0.03	-0.01	0.03	0.03
0.005–0.01	-0.13	-0.09	-0.04	-0.03
0.01–0.06	-0.09	-0.08	-0.04	-0.03
0.06–0.1	-0.02	0.02	0.03	0.04
0.1–0.2	-0.07	-0.03	-0.00	0.01
0.2–0.3	-0.11	-0.09	-0.04	0.00
0.3–0.4	-0.06	-0.13	-0.06	-0.01
0.4–0.5	0.22	0.01	0.07	0.11
0.5–0.6	0.85	0.40	0.41	0.39
0.6–0.7	2.6	1.7	1.6	1.4
0.7–0.8	7.3	5.5	5.1	4.4
0.8–0.9	20.2	16.7	16.1	13.4

Table 3: The values of the higher-twist coefficients D_i, in the chosen bins of x, extracted from the LO, NLO, NNLO and NNNLO (NNLO with the approximate NNNLO non-singlet quark coefficient function) global fits.