Proton Parton Distribution Functions and LHC

- Motivation – LHC PDF needs
- PDF extraction procedure — Data, Theory and the Combination.
- Predictions for LHC
- Conclusions and Outlook.

DPG Dortmund 31 March 2006

LHC and precision physics

Two opposite opinions exist about the need of LHC for precision parton distribution functions (PDFs):

- Proton structure is absolutely vital for a *pp* machine, precise knowledge of PDFs + exact theoretical NNLO QCD calculations $\rightarrow$ exact predictions for LHC.
- LHC is a discovery machine operating in a messy hadronic environment. Experimental uncertainties would dominate anyhow.

As usual the truth is in between. A limited set of observables can be measured at LHC with high precision. For the rest PDFs are used for QCD predictions to estimate background – with moderate precision.

LHC Discovery channels

An (incomplete) list of channels to be searched at LHC

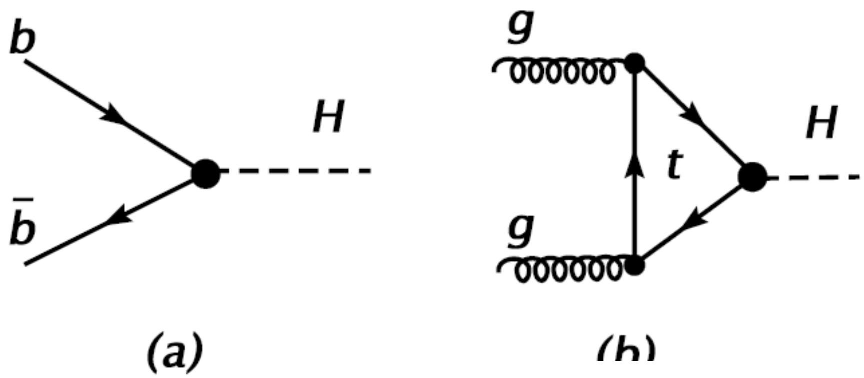
- Peak in invariant mass distribution $M_{\ell\ell}$ for high momentum leptons (Drell-Yan processes).
- Large missing $p_\perp$ — escaping LSP. Modeling of QCD background is the main issue.
- High $p_\perp$ jets — large systematic uncertainties.
- Higgs in various decay modes ($\gamma\gamma$, WW ...).

For experimental signatures with leptons in the final state few percent experimental counting precision is feasible (including $H \to \gamma\gamma$).

On the theory side, precision calculation are available for SM Drell-Yan processes.

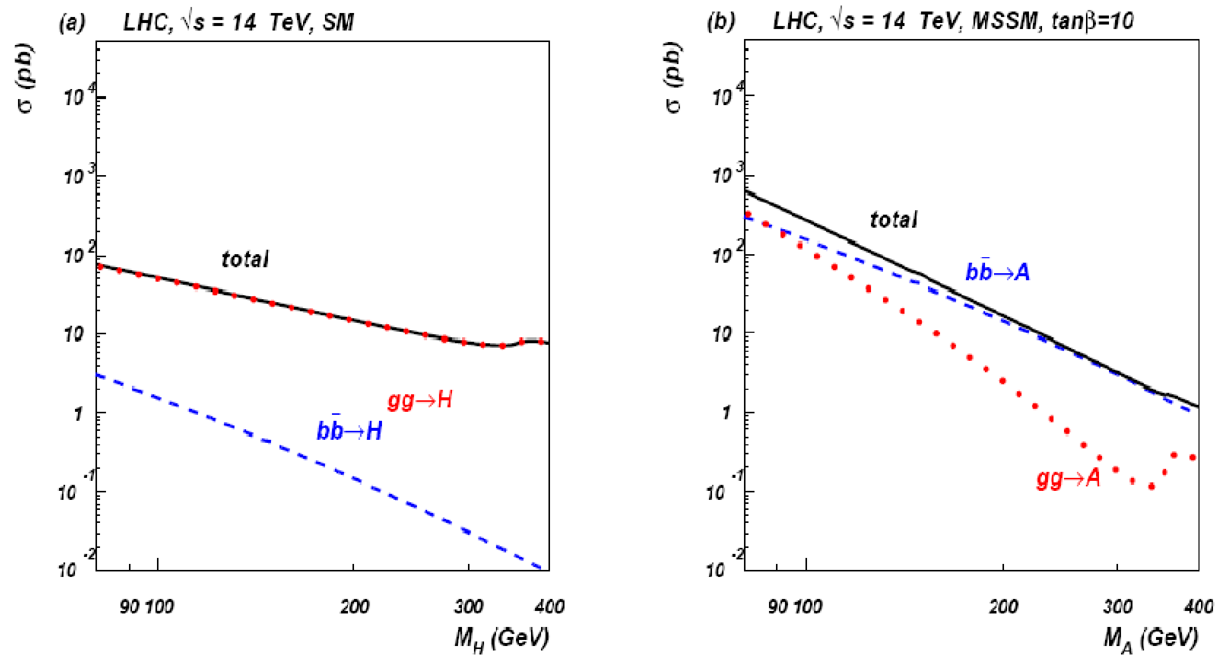
Higgs production cross section is calculable to $\sim 10\%$ in SM, very different for little Higgs models, $\sim 10\%$ different for MSSM.

Case study: Higgs production at LHC, SM vs MSSM



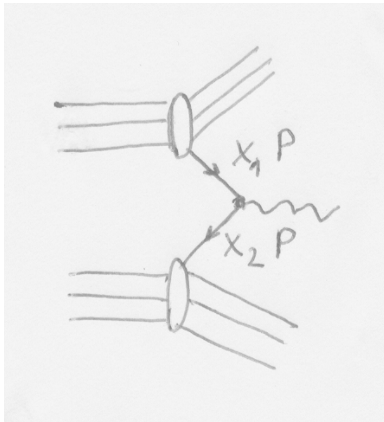
In SM, $b\bar{b} \to H$ is small vs $gg \to H$.
In MSSM, $b\bar{b} \to H$ can be enhanced by $\times \tan^2 \beta$

Even for MSSM with $\tan\beta = 10$, $b\bar{b} \to H$ dominates over gg production.



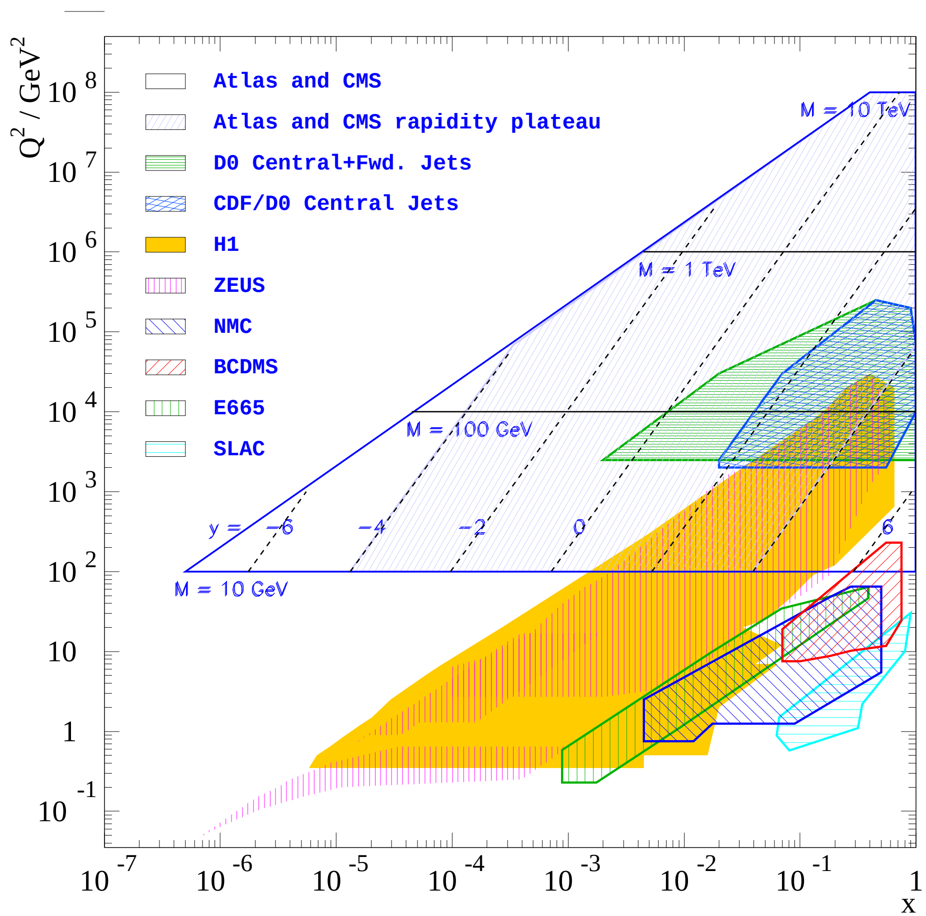
$\rightarrow$ production cross section measurement of Higgs is a key ingredient to disentangle new physics scenarios.

LHC kinematics



x_1, x_2 are momentum fractions. Factorization theorem states that cross section can be calculated using universal partons × short distance calculable partonic reaction.

$$x_{1,2} = \frac{M}{\sqrt{S}} \exp(\pm y)$$



Notation clash: y – rapidity (LHC) vs y – inelasticity (HERA, $Q^2 = Sxy$).

LHC and luminosity measurement

For a hadron collider, there is no good minimum bias QED (like $ep \to ep\gamma$) reference process which can be predicted/measured with high precision.

A better way is to normalize rate of a process of interest to a high $p_\perp$ process which is calculated/can be measured with high accuracy.

The best candidates for the reference is W, Z production:

- Have high rate even at low luminosity (few Hz)
- Calculated to **NNLO** (2% precision)
- Can be measured via leptonic decays up to 1-2% precision (Z in the central rapidity region).

The Bjorken x range for W, Z production measured at $|y| < 2.5$ is $0.0005 - 0.05$; for H production with $m_H \sim 140$ GeV is $0.001 - 0.02$

PDF determination

$$\frac{d^2\sigma^{NC}_{e^\mp p}}{dx dQ^2} = \frac{2\pi\alpha^2 Y_+}{xQ^4}\left(F_2 - \frac{y^2}{Y_+}F_L \pm \frac{Y_-}{Y_+}xF_3\right) \quad Y_\pm = 1 \pm (1-y)^2$$

Leading order relations:

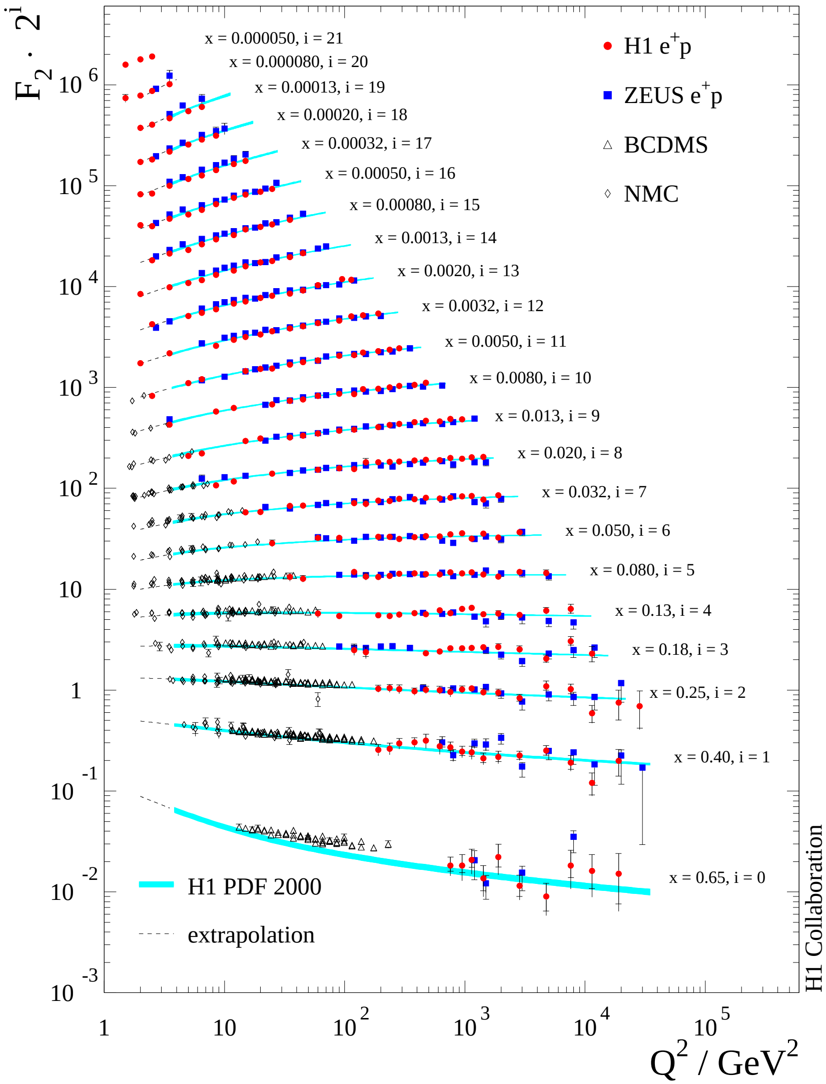
$$\boxed{\begin{array}{cc} F_2 & = x\sum e_q^2(q(x)+\bar{q}(x)) \\ xF_3 & = x\sum 2e_q a_q(q(x)-\bar{q}(x)) \\ \sigma^{CC}_{e^+p} & \sim x(\bar{u}+\bar{c}) + x(1-y)^2(d+s) \\ \sigma^{CC}_{e^-p} & \sim x(u+c) + x(1-y)^2(\bar{d}+\bar{s}) \\ pp \to (\ell\bar{\ell})X & \sim \sum x_1 x_2 q(x_1)\bar{q}(x_2) \end{array}}$$

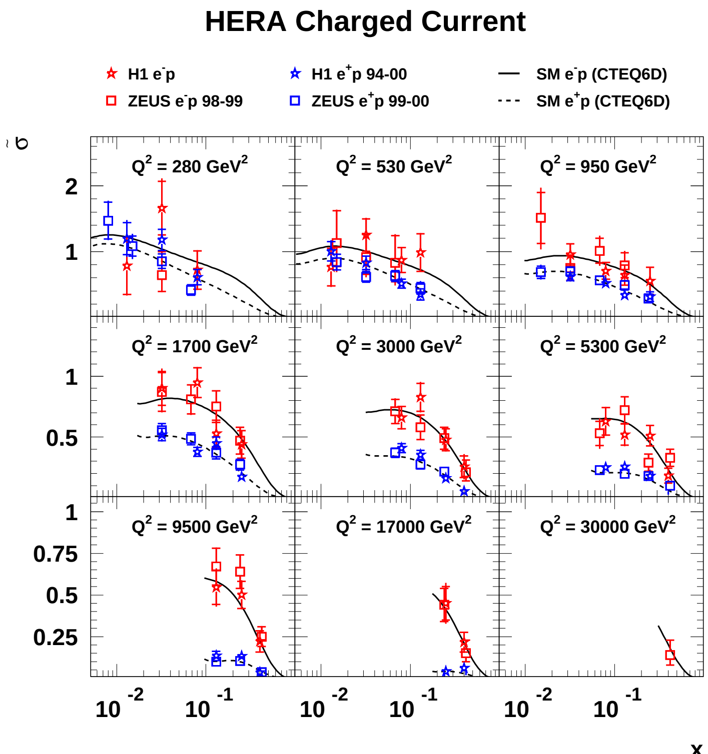
DIS ep and ed data allows to unfold individual quark flavors.

Drell-Yan data assists in determination of $\bar{q}$ content.

Gluon is determined from F_2 scaling violation and from pp jet cross section.

The Measured Cross Sections





HERA data allows to measure $xU = x(u + c)$, $xD = x(d + s)$, $x\bar{U} = x(\bar{u} + \bar{c})$, $x\bar{D} = x(\bar{d} + \bar{s})$, and xg in a single experiment.

Status of the theory

Factorization theorem: Observables are convolution of universal parton densities and calculable process dependent coefficient functions.

Coupled DGLAP evolution equation for $\Sigma = \sum q + \bar{q}$ and g:

$$\frac{d}{d \log Q^2} \begin{pmatrix} \Sigma \\ g \end{pmatrix} = \frac{\alpha_S}{2\pi} \int_x^1 \frac{dz}{z} \begin{pmatrix} P_{qq} & P_{qg} \\ P_{gq} & P_{gg} \end{pmatrix} \begin{pmatrix} \Sigma \\ g \end{pmatrix}$$

LO,NLO,NNLO calculations exist for DIS and Drell-Yan data. Recent calculation of NNLO splitting functions involved > 10000 diagrams, took a decade to complete.

Starting from NLO, definitions of parton densities are factorization scheme dependent. In DIS scheme $F_2 = x \sum e_q^2 (q + \bar{q})$ is preserved, i.e. parton densities directly related to physical observables.

Non-conventional QCD evolution

A long standing issue for small x (and similarly for $x \sim 1$) are possible modifications of the standard DGLAP evolution due to large $\log 1/x$ terms. In addition, high gluon density may lead to saturation effects, adding non-linear terms to the evolution.

For the starting scale Q_0^2, small x effects are absorbed in the input quark parton densities, together with long distance effects $\rightarrow$ small-x effects can be only observed in the modification of the evolution. The QCD evolution at small x is driven by the gluon density which is a free parameter of DGLAP at the starting scale $\rightarrow$ at first order small-x effects in the evolution are absorbed in the gluon density.

This is <u>bad news</u> for theorists interested to find new QCD dynamics using inclusive data but <u>good news</u> for F_2-like predictions (W, Z-production cross section at LHC). The argument <u>may not</u> hold for gluon dominated observables (F_L, H-production cross section at LHC).

In this talk, only standard DGLAP

Treatment of experimental data

$$\chi^2(\{p\},\{\alpha\}) = \sum_i \frac{\left[F_2(p) - \left(F_2^i + \sum_j \frac{\partial F_2^i}{\partial \alpha_j}\alpha_j\right)\right]^2}{\sigma_{F_2}^2} + \sum_j \frac{\alpha_j^2}{\sigma_{\alpha_j}^2}.$$

p — are parameters used to describe PDFs,
α_j — systematic uncertainty source j
$F_2(p)$ — theoretical prediction for F_2
F_2 — central value of the measured F_2
σ_{F_2} — statistical uncertainty of F_2
σ_{α_j} — uncertainty on systematic source j

Systematic uncertainties, including absolute normalizations of the data sets, are varied together with the PDF parameters. In some fits (CTEQ, H1) systematics is fitted, in others (MRST,ZEUS) central values are fixed. All fits float absolute normalizations.

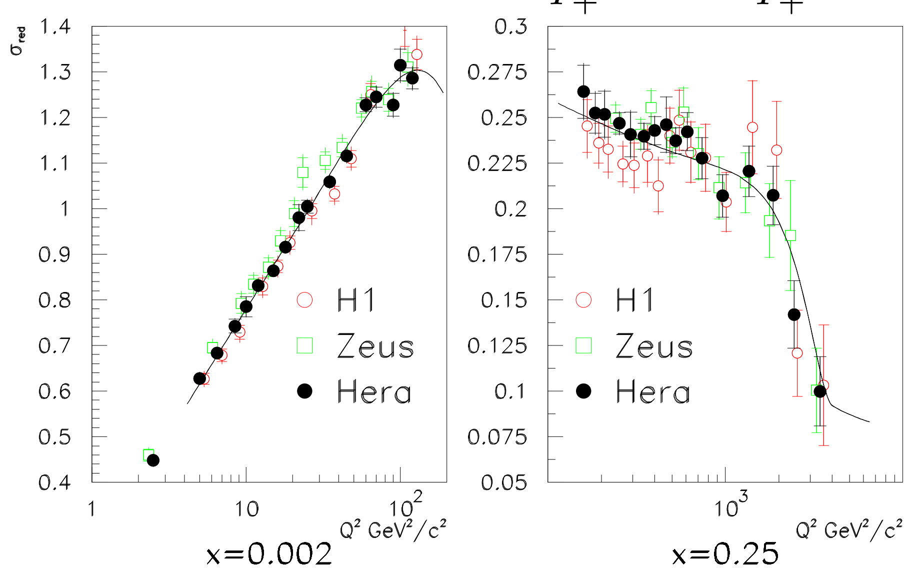
Combination of Experimental Data

Before fitting to theory one can combine data in a generalized averaging procedure. Achieved by fitting χ^2 vs F_2.

– Number of the fit parameters is equal to number of x, Q^2 points — large matrix inversion.

+ Simple quadratic dependence χ^2 — unique and simple solution.

$$\sigma_{red}(e^+p) = F_2 - \frac{y^2}{Y_+}F_L - \frac{Y_-}{Y_+}xF_3$$



Average of H1 and Zeus data: model independent check of the consistency, $\chi^2/ndf = 534/601$. Experiments cross calibrate each other → systematic errors reduced.

Extraction of Parton Densities

Parameterization of PDFs at starting scale (CTEQ case):

$$xf(x,Q_0^2) = Ax^B(1-x)^C e^{Dx}(1+Ex)^F$$

For simplest minimal form $D = E = F = 0$.

Momentum sum rule: $\int x\,(\Sigma(x) + g(x))\,dx = 1$,

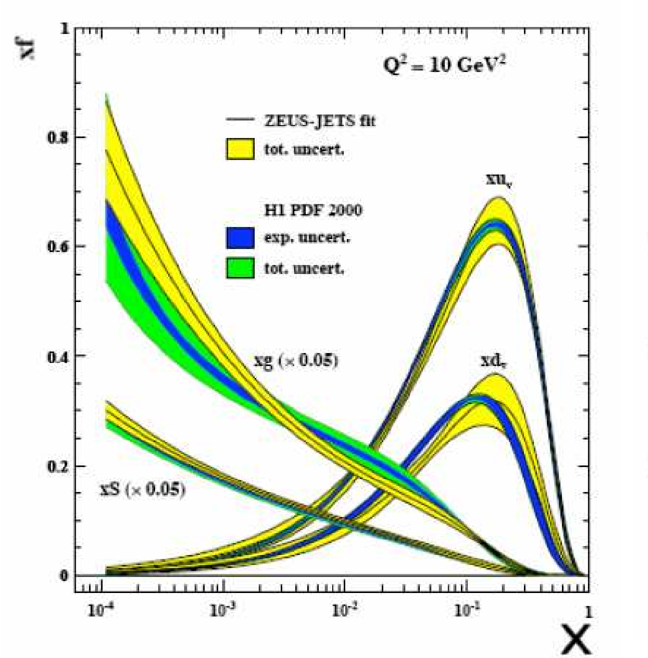
Assumptions: $s = \bar{s} = 0.2(\bar{u} + \bar{d})$, $B_u = B_{\bar{u}} = B_d = B_{\bar{d}}$

Recent global fits follow different approaches in details:

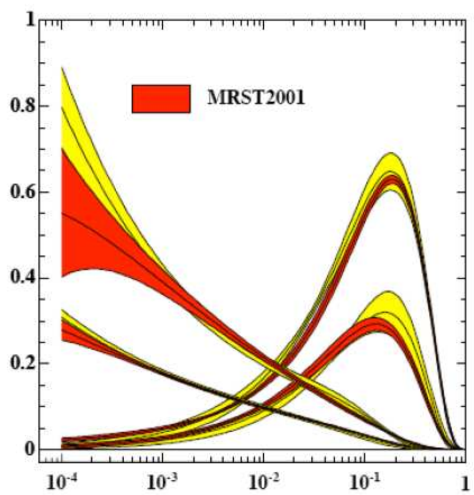
Alekhin – NNLO DIS and DY
MRST and ZEUS – float data normalizations only
CTEQ and H1 – float data normalizations and syst. sources
H1 – minimum set of parameters for PDFs to get good χ^2
CTEQ – maximum set of parameters still getting stable fit

$\rightarrow$ an estimate of the uncertainty should use errors provided by different groups but also compare central values.

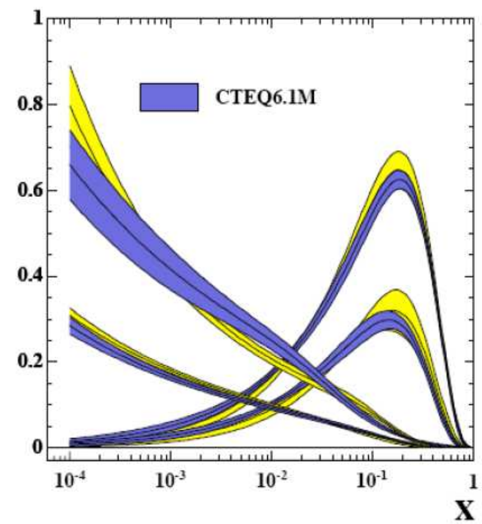
PDFs extracted by various groups



Zeus and H1 PDFs using their own data. Agree within the uncertainties.

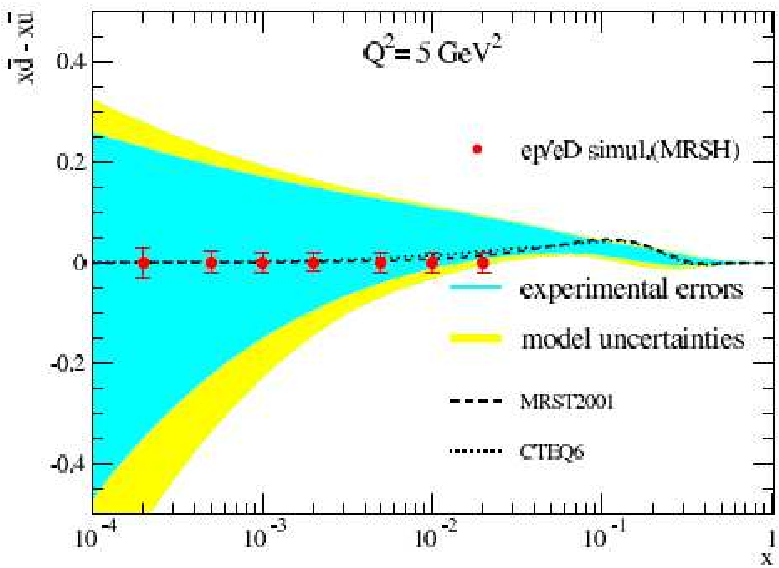


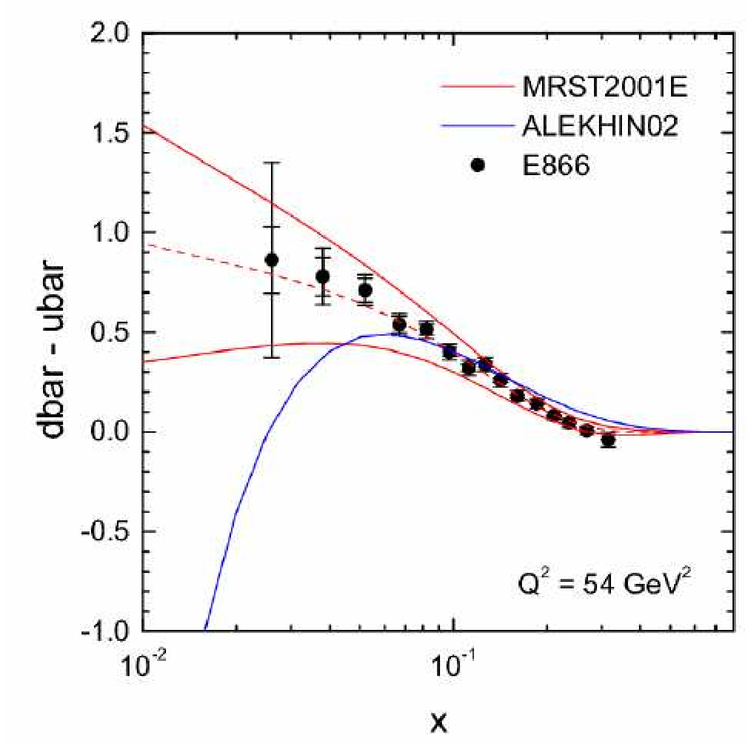
Zeus gluon agrees better with MRST, H1 — with CTEQ. (similar data treatment ?).



Effect of PDF constraints

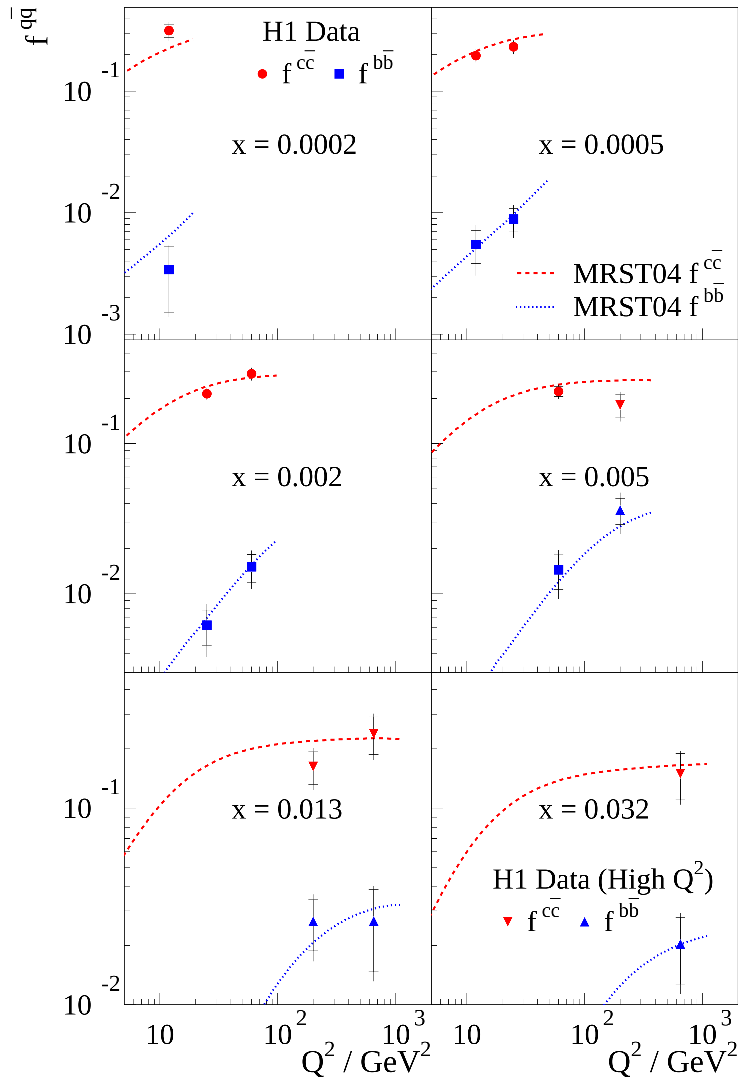
Warning: relaxing low x constraints leads to large variations of PDFs





Could have been measured with *ed* run at HERA. Can be studied using Drell-Yan at LHC. For now assume conventional low x shape.

Heavy flavors



Sea decomposition can be studied using structure function data with tagged leading particle, using for example c, b long life times.

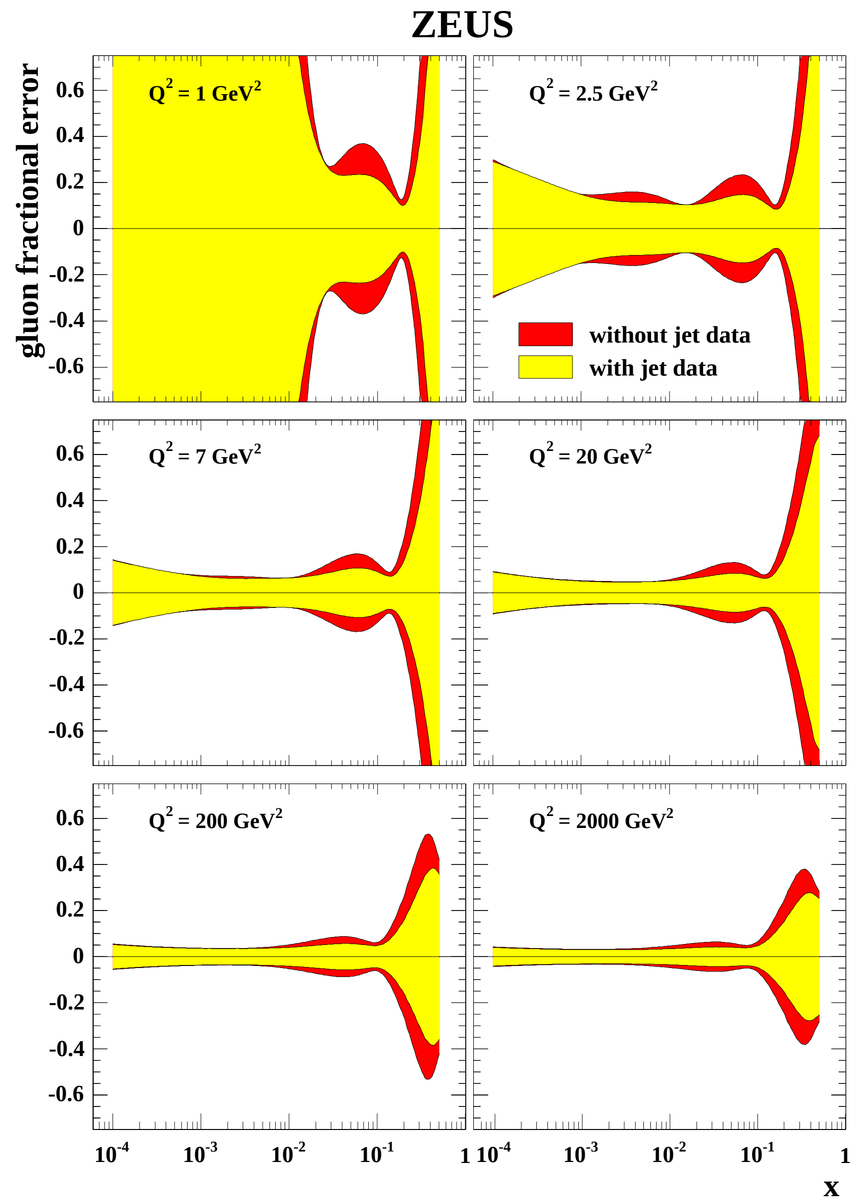
These measurements will benefit from larger HERA-II statistics. ZEUS joints H1, both detectors are now equipped with silicon vertex detectors.

Reducing Gluon uncertainties – high x

Global PDF fits (CTEQ, MRST) include Tevatron jet data to extract $g(x)$ for high x. The data from Tevatron suffers from large systematic uncertainties (jet energy scale).

Recently Zeus collaboration included their jet data in the QCD fit.

Note that NNLO theoretical predictions are needed to make full use of the data. Pure NNLO fits (e.g. Alekhin) don't use jets data for now.

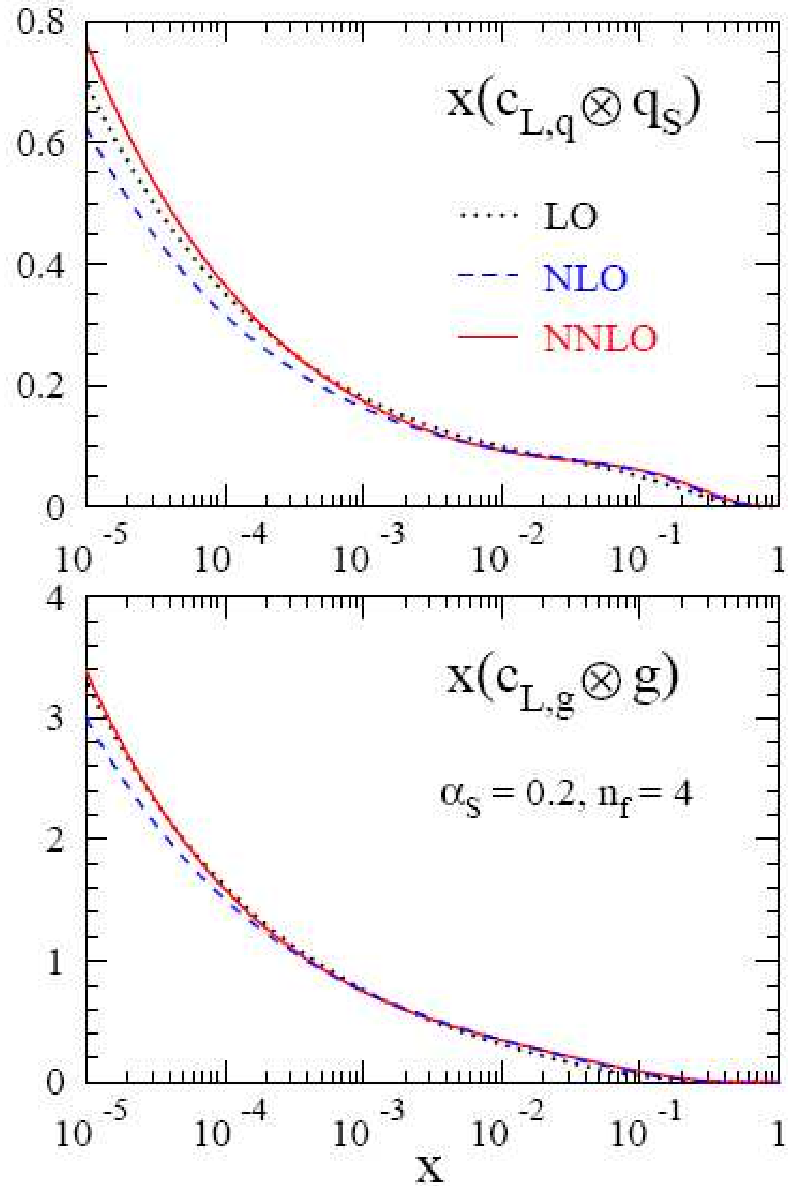


Low x Gluon and F_L predictions

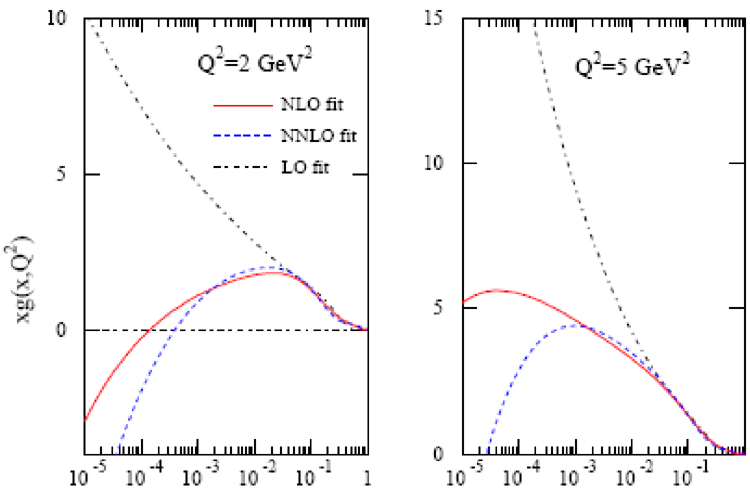
In QPM, $F_L = 0$
To lowest order of DGLAP at low x, F_L can be solved approximately.

$$xg(x) \approx \frac{8.3}{\alpha_S} F_L(0.4x)$$

Recently NNLO coefficient functions has been calculated.
Rather large variation from NLO to NNLO, but the same fixed PDFs are used for the plot.

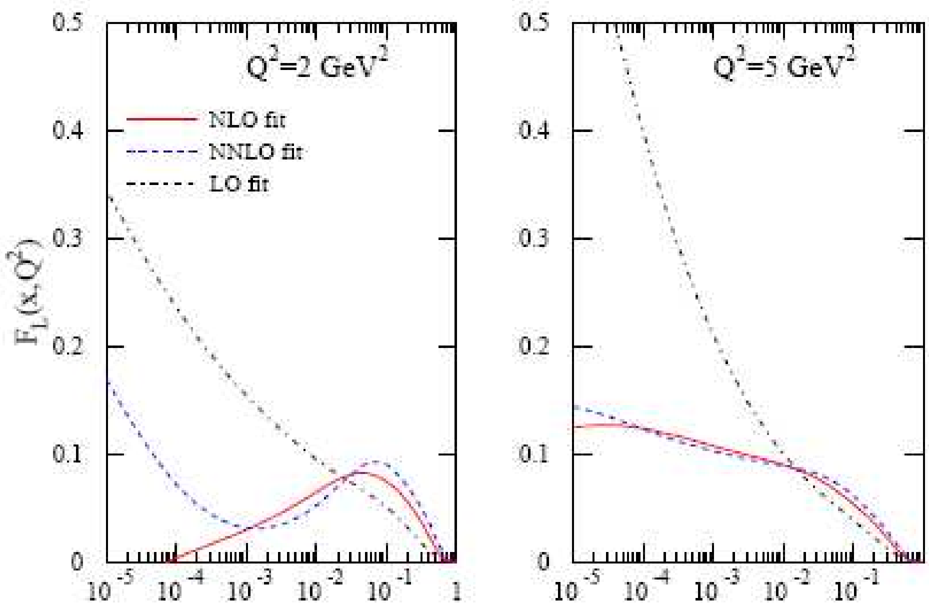


F_L predictions II



Gluon obtained from LO,NLO and NNLO MRST fits. For $Q^2 = 2$ GeV^2, negative gluon.

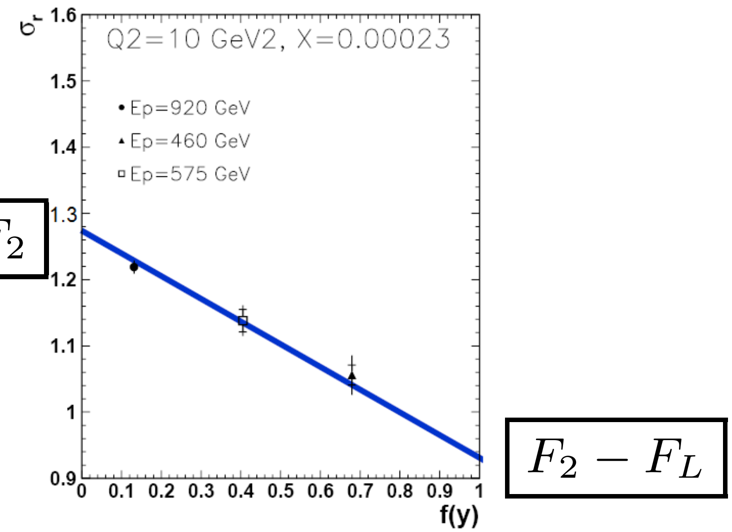
F_L calculated using LO,NLO,NNLO gluon. For $Q^2 = 5$ GeV^2, perturbative convergence.

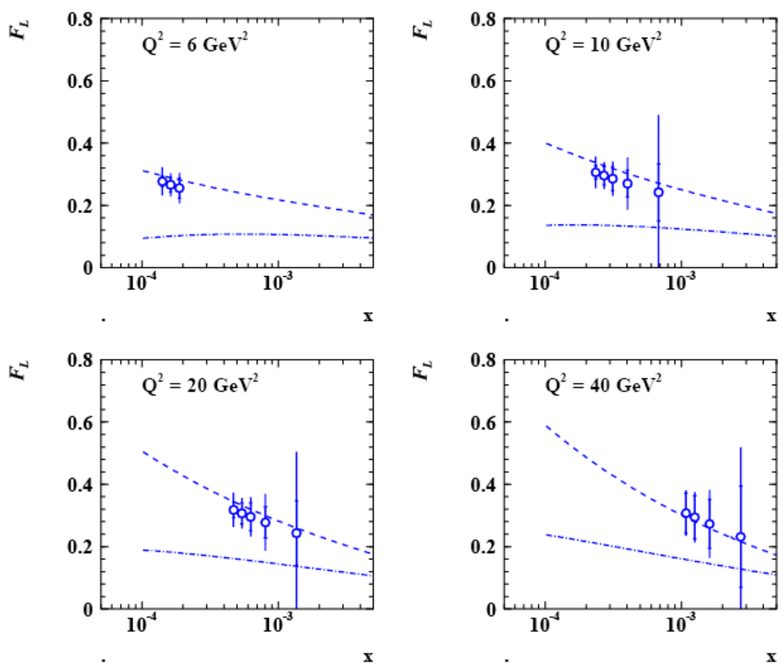


F_L measurement

$$\sigma_r(x,Q^2) = F_2(x,Q^2) - f(y)F_L(x,Q^2)$$

Measure σ_r at the same Q^2, x for different beam energies

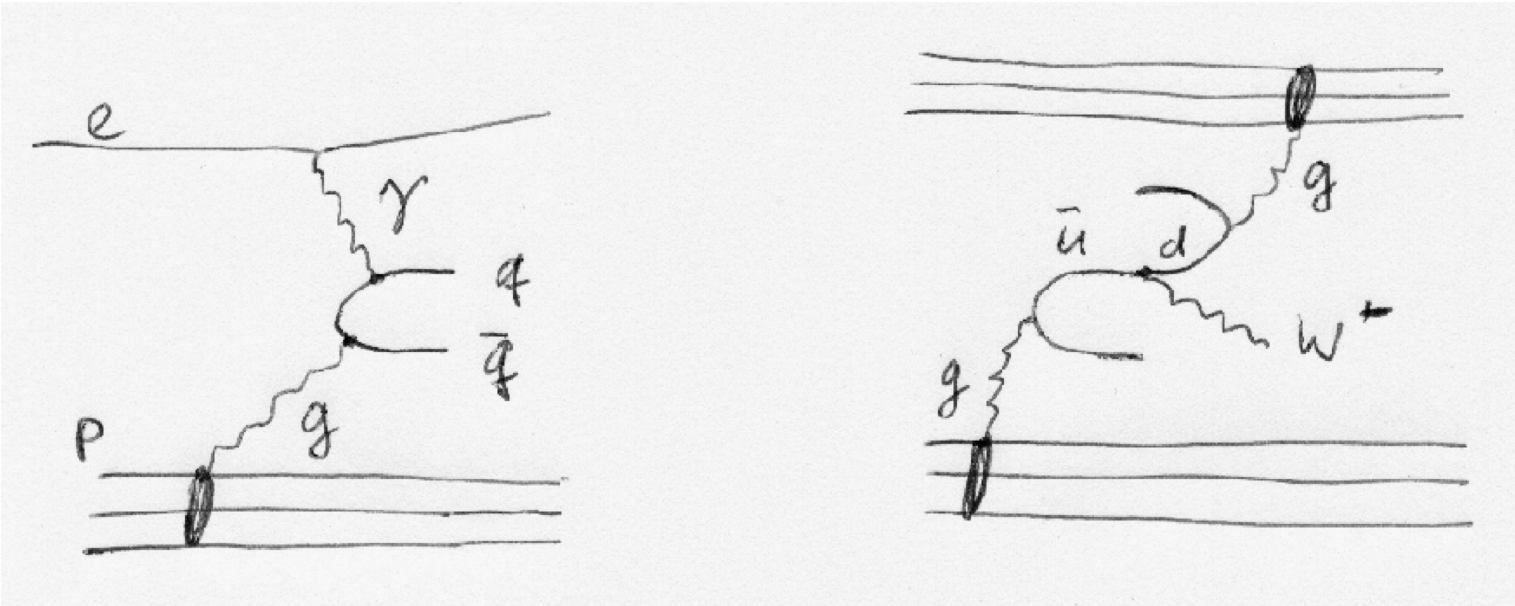




Measurement based on 10 pb^{-1} run at $E_p = 460$ GeV allows to distinguish between different PDF fits (MRST vs CTEQ).

Both H1 and ZEUS expressed interest in this measument

W, Z production at LHC

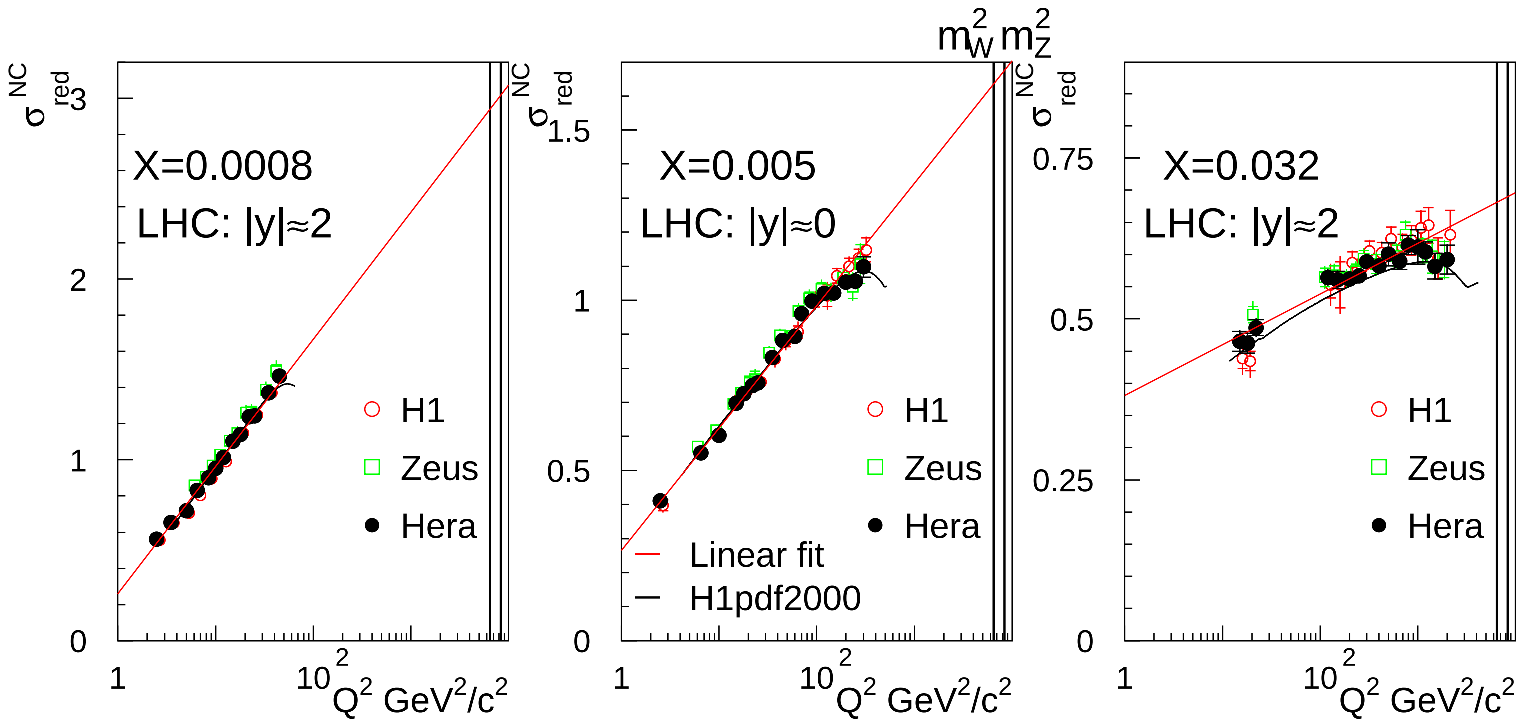


At low x similar boson-gluon fusion leading diagram for DIS and Drell-Yan.

One needs relation from $F_2 \sim 4/9U(x) + 1/9D(x)$ to

$$
\begin{array}{ll}
|V_{ud}|^2\bar{u}(x_1)d(x_2) + |V_{cs}|^2\bar{c}(x_1)s(x_2) & \text{for } W^- \\
|V_{ud}|^2\bar{d}(x_1)u(x_2) + |V_{cs}|^2\bar{s}(x_1)c(x_2) & \text{for } W^+ \\
\bar{u}(x_1)u(x_2) + \bar{u}(x_2)u(x_1) + \ldots & \text{for } Z
\end{array}
$$

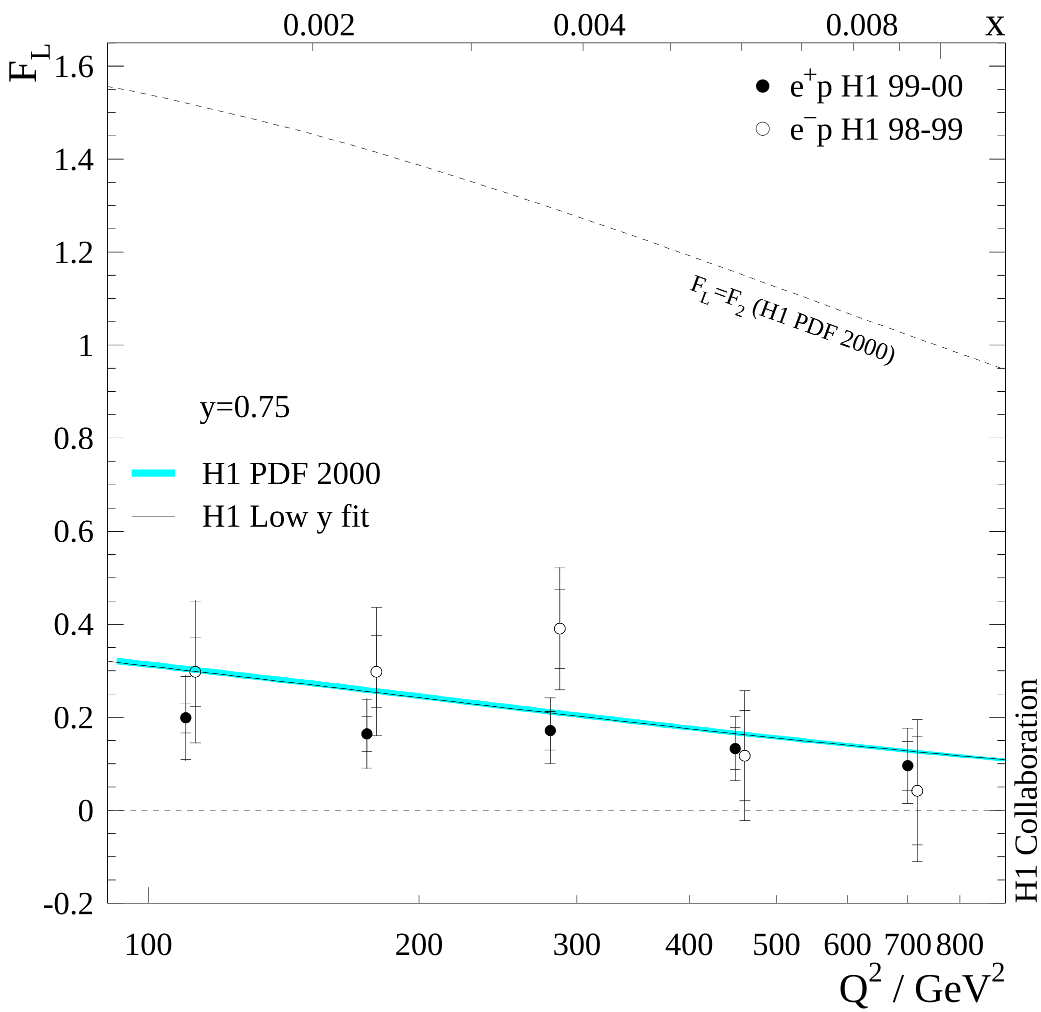
F_2 extrapolation to W, Z mass



HERA data covers complete central rapidity range of LHC for W, Z production. "Leading order" predictions can be read directly from HERA data + linear extrapolation.

Experimental part of PDF uncertainties comes from absolute F_2 normalization and the slope, $dF_2/d\log Q^2$ (gluon). Turn down of σ_{red}^{NC} for highest Q^2 ($\rightarrow$ highest y) is due to F_L.

Consistency check: H1 F_L determination at high Q^2



Determination of F_L as
$$F_L = \frac{Y_+}{y^2}\left(F_2^{fit} - \sigma_r\right)$$

Important consistency check of gluon determined from F_2 scaling violation vs X-section decrease at high y.

Still large statistical uncertainties, to be improved with HERA-II

W, Z production predictions to NNLO

$$+16S_{1,-2,-2}+\frac{103}{3}S_{1,-2,1}-2S_{1,-2,2}-36S_{1,1,-3}+56S_{1,1,-2,1}+8S_{1,1,3}-\frac{109}{9}S_{1,2}-4S_{1,2,-2}+\frac{43}{3}S_{1,3}-8S_{1,3,1}-11S_{1,4}+\frac{11}{3}S_{2,2}$$

$$+21S_{2,-3}-30S_{2,-2,1}-4S_{2,1,-2}-5S_{2,3}-S_{4,1}+\frac{31}{6}S_{2,-2}-\frac{67}{9}S_{2,1}\Big]+(\mathbf{N}_--\mathbf{N}_+)\Big[9S_2\zeta_3+2S_{2,-3}+4S_{2,-2,1}-12S_{2,1,-2}-2S_{2,3}$$

$$+13S_{4,1}+\frac{1}{2}S_{2,-2}+\frac{11}{2}S_4-\frac{33}{2}S_{3,1}+\frac{59}{9}S_3+\frac{127}{6}S_{2,1}-\frac{1153}{72}S_2\Big]+\mathbf{N}_+\Big[8S_{3,-2}+\frac{4}{3}S_{3,1}-2S_{3,2}+14S_5+\frac{23}{6}S_4+\frac{73}{3}S_3+\frac{151}{24}S_2\Big]\Big)$$

$$+16{C_F}^2n_f\Big(\frac{23}{16}-\frac{3}{2}\zeta_3+\frac{4}{3}S_{-3,1}-\frac{59}{36}S_2+\frac{4}{3}S_{-4}-\frac{20}{9}S_{-3}+\frac{20}{9}S_1-\frac{8}{3}S_{1,-2}-\frac{8}{3}S_{1,1}-\frac{4}{3}S_{1,2}+\mathbf{N}_+\Big[\frac{25}{9}S_3-\frac{4}{3}S_{3,1}-\frac{1}{3}S_4\Big]$$

$$-(\mathbf{N}_+-1)\Big[\frac{67}{36}S_2-\frac{4}{3}S_{2,1}+\frac{4}{3}S_3\Big]+(\mathbf{N}_-+\mathbf{N}_+)\Big[S_1\zeta_3-\frac{325}{144}S_1-\frac{2}{3}S_{1,-3}+\frac{32}{9}S_{1,-2}-\frac{4}{3}S_{1,-2,1}+\frac{4}{3}S_{1,1}+\frac{16}{9}S_{1,2}-\frac{4}{3}S_{1,3}$$

$$+\frac{11}{18}S_2-\frac{2}{3}S_{2,-2}+\frac{10}{9}S_{2,1}+\frac{1}{2}S_4-\frac{2}{3}S_{2,2}-\frac{8}{9}S_3\Big]\Big)+16{C_F}^3\Big(12S_{-5}-\frac{29}{32}-\frac{15}{2}\zeta_3+9S_{-4}-24S_{-4,1}-4S_{-3,-2}+6S_{-3,1}$$

$$-4S_{-3,2}+3S_{-2}+25S_3-12S_{-2}\zeta_3-12S_{-2,-3}+24S_{-2,1,-2}-52S_{1,-3}+4S_{1,-2}+48S_{1,-2,1}-4S_{3,-2}+\frac{67}{2}S_2-17S_4$$

$$+(\mathbf{N}_-+\mathbf{N}_+-2)\Big[6S_1\zeta_3-\frac{31}{8}S_1+35S_{1,1}-12S_{1,1,-2}+S_{1,2}+10S_{2,-2}+S_{2,1}+2S_{2,2}-2S_{3,1}-3S_5\Big]+(\mathbf{N}_-+\mathbf{N}_+)\Big[23S_{1,-3}$$

$$-22S_{1,-4}+32S_{1,-3,1}-2S_{1,-2}-8S_{1,-2,-2}-30S_{1,-2,1}-6S_{1,3}+4S_{1,-2,2}+40S_{1,1,-3}-48S_{1,1,-2,1}+8S_{1,2,-2}+4S_{1,2,2}+8S_{1,3,1}$$

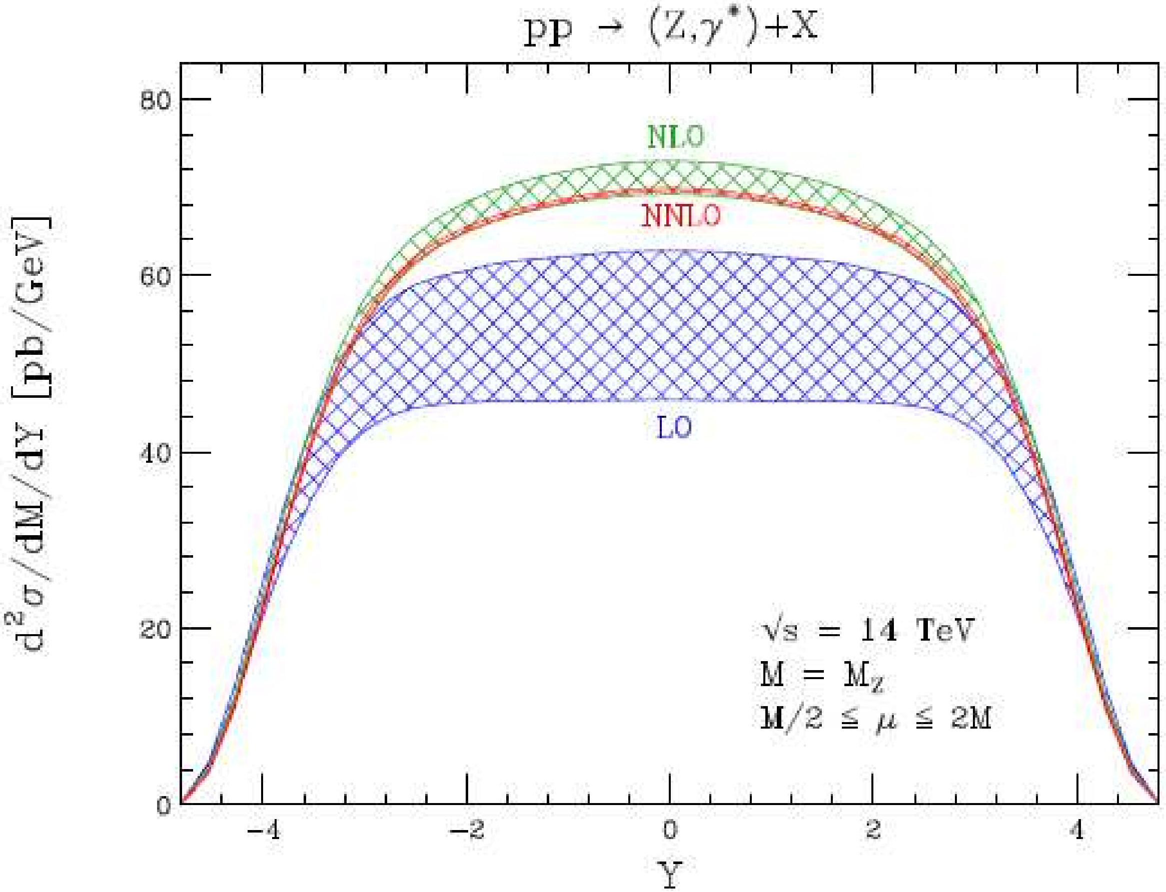
$$+4S_{1,4}+28S_{2,-2,1}+4S_{2,1,2}+4S_{2,2,1}+4S_{3,1,1}-4S_{3,2}+8S_{2,1,-2}-26S_{2,-3}-2S_{2,3}-4S_{3,-2}-3S_{2,-2}-3S_{2,2}+\frac{3}{2}S_4\Big]$$

$$+(\mathbf{N}_--\mathbf{N}_+)\Big[12S_{2,1,-2}-6S_2\zeta_3-2S_{2,-3}+3S_{2,3}+2S_{3,-2}-\frac{81}{4}S_{2,1}+14S_{3,1}-5S_{2,-2}-\frac{1}{2}S_{2,2}+\frac{15}{8}S_2+\frac{1}{2}S_3-13S_{4,1}+4S_5\Big]$$

$$+\mathbf{N}_+\Big[14S_4-\frac{265}{8}S_2-\frac{87}{4}S_3-4S_{4,1}-4S_5\Big]\Big)$$

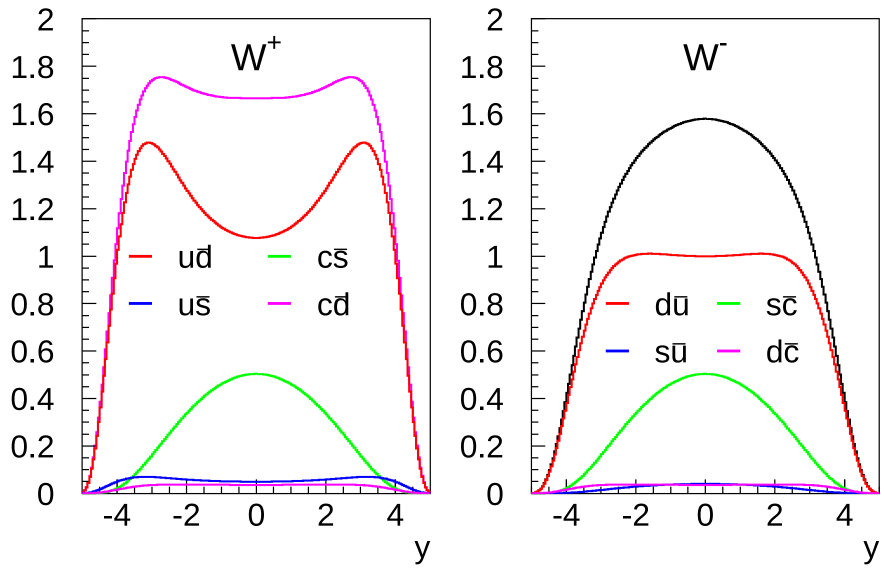
(part of anomalous dimension calculation)

W, Z predictions theoretical stability



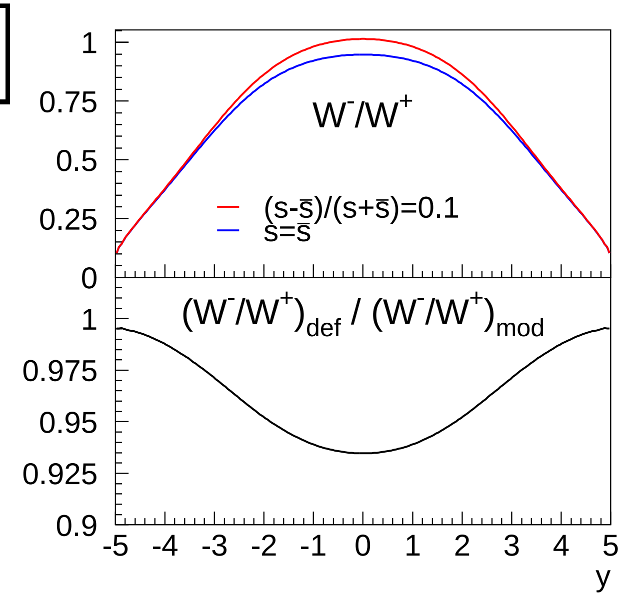
NNLO calculations agree well with NLO estimate → perturbative convergence.
NNLO is within 2% stable vs scale variation

W production flavor decomposition



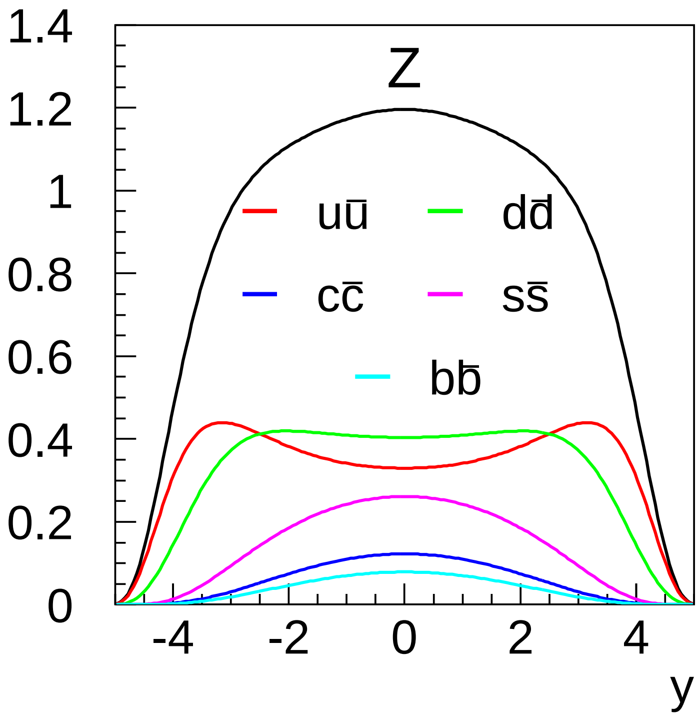
u, d, s, c quarks contribute to $W^{\pm}$ production. Cabibbo enhanced sea $c\bar{s}$, $s\bar{c}$ by $\sim$ 25%, Cabibbo sea suppressed by $\sim$ 5%.

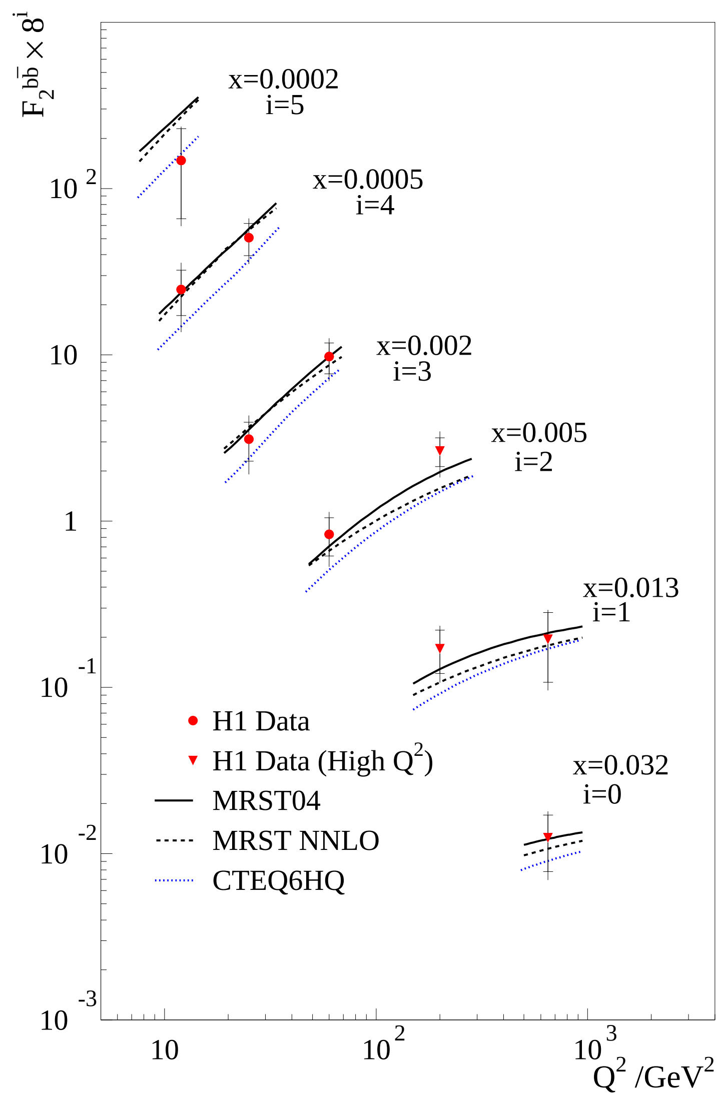
W via subthreshold $t\bar{b}$ as for low Q^2 F_2^{cc} ?



Recall that F_2 controls only $4/9U(x)+1/9D(x)$, change in sea symmetry assumption leads to significant W^-/W^+ rate change while F_2 is not modified.

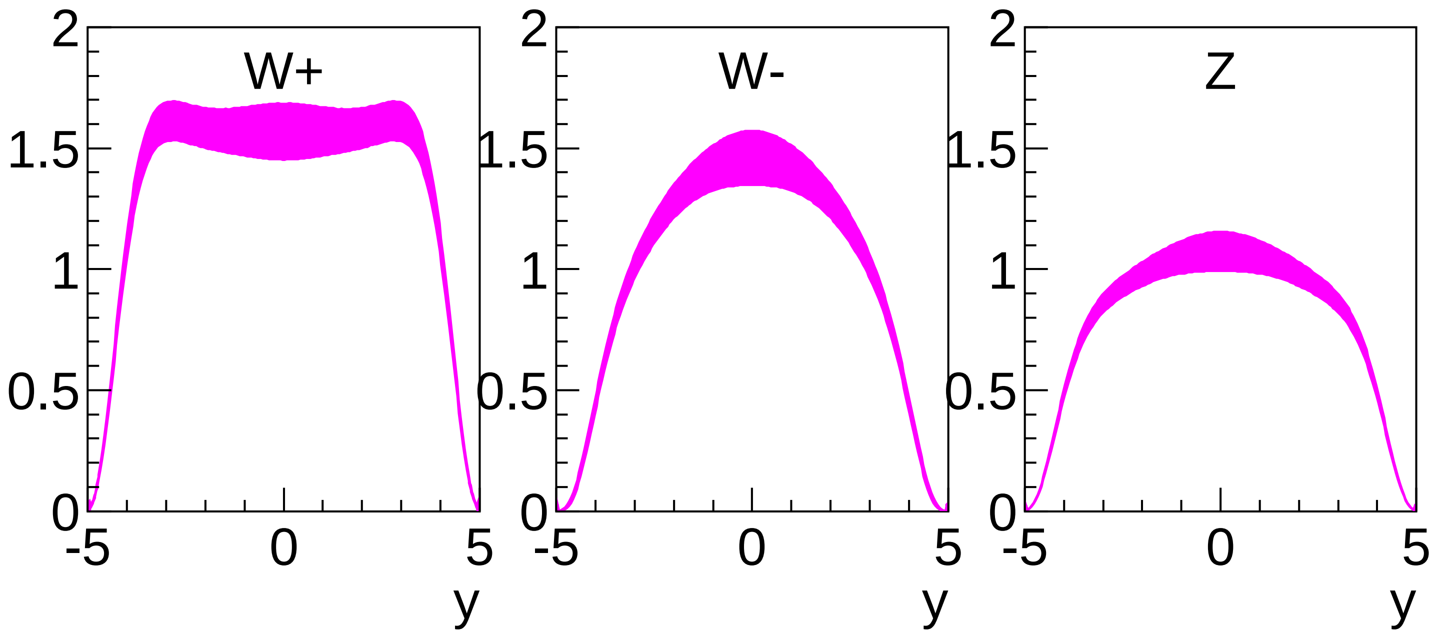
Z production flavor decomposition





Larger coupling to Z vs γ makes $b\bar{b}$ contribution more important for Z production vs inclusive F_2. Still, F_2^{bb} is measured at HERA.

$W^{\pm}, Z$ pdf uncertainties

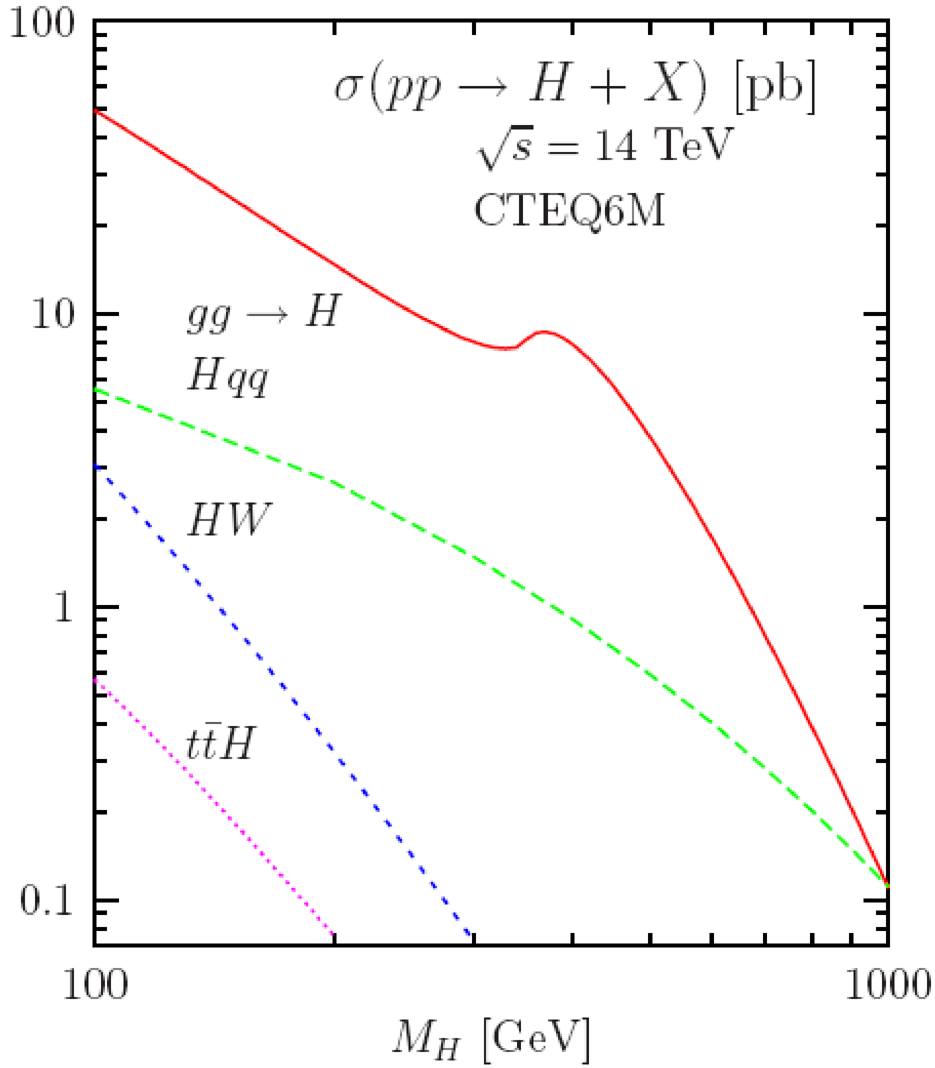


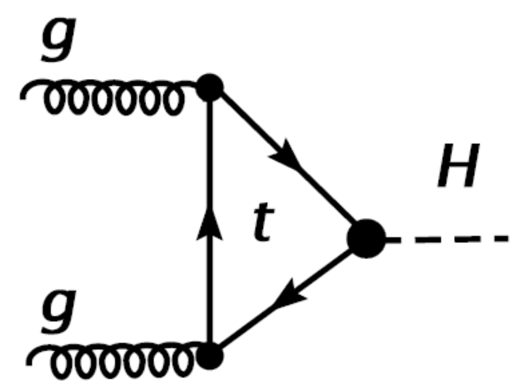
Current PDF uncertainties are $\sim 7\%$ (CTEQ). Without HERA data: $> 20\%$.

7% is large compared to 2% theory and ultimately $\sim 2\%$ experimental precision at LHC.

$\rightarrow$ need for more precise HERA data for $0.001 < x < 0.03$ kinematic domain.

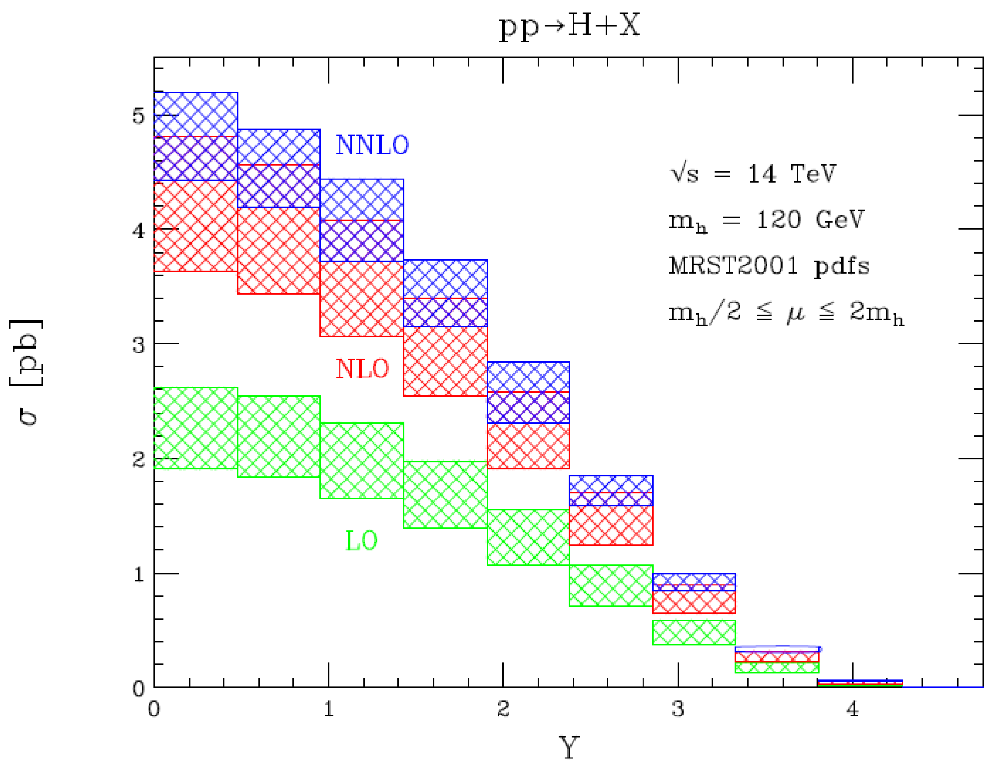
SM Higgs production





In SM, for light Higgs boson the dominant production mechanism is $gg \to H$.

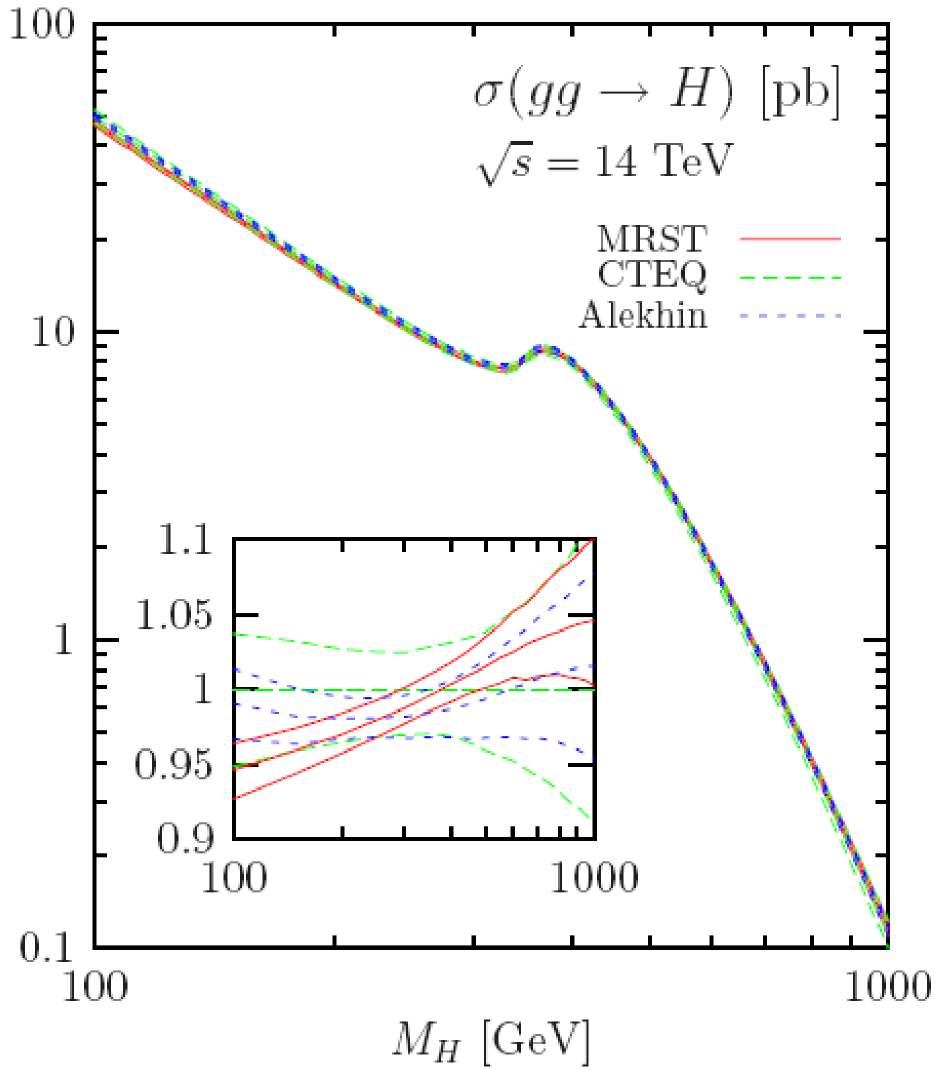
Higgs production NNLO predictions



Higgs production via $gg \to H$ receives significant positive corrections for higher orders.

Still relatively large renormalization scale uncertainty ($\sim 10\%$) even at NNLO. Larger uncertainties for H vs W, Z are similar to larger uncertainties of F_L vs F_2, – larger sensitivity to gluon.

Higgs PDF uncertainties



PDF uncertainties for $gg \to H$ are about $\sim 8\%$ taking into account difference between central values for different PDFs.

HERA-II for LHC: What next?

LHC needs improvement of PDF determination for W, Z production and better understanding of gluon for Higgs production.

- $\times 5$ increase in statistics expected at HERA-II should allow to reduce uncertainties for F_2 at $Q^2 > 1000$ GeV^2 by about factor of 2.
- For $Q^2 < 100$ GeV^2, $2 - 3\%$ uncertainty is already achieved at HERA-I. Detailed studies of systematic uncertainties show that the errors could be reduced to $1 - 1.5\%$ level.
- Direct measurement of F_L with a low energy run would provide an important constraint for gluon at low x; theory self consistency check in a complicated low x, Q^2 domain.
- Indirect determination of $F_L \sim F_2^{QCD} - \sigma_{red}$ at high Q^2 will improve with increased statistics; provides gluon universality check for Q^2, x values closer to LHC energy.

Still a lot of work at HERA but expected $\times 2$ improvement for W, Z production prediction and better control over theory.

Conclusions

- Precision PDFs are important for key measurements at LHC, in particular for the Higgs production cross section.
- Precision cross section measurements at LHC require measurements of SM W and Z production as a luminosity monitor.
- HERA measurement of the inclusive cross sections are vital for PDFs needed for LHC. Current precision on W,Z production cross section is about 7%.
- Further improvements in PDFs are expected. They should come from higher precision inclusive cross section data, from new measurements of gluon (F_L), from complete NNLO fits, from more rigorous treatment of the data systematic uncertainties.