# The global fit of PDFs including the low-$Q$ DIS data

(S.Alekhin, IHEP, Protvino)

Despite a long history of this studies we are still working on PDFs, in order to achieve accuracy of O(1%), necessary for precision physics at the LHC.

**Problems and pitfalls:**

- *The high order QCD corrections*

  The value of $\alpha_s$ is and the pQCD series converge rather slowly (LO: 20-30%, NLO: 5-10%, NNLO: 2-3%). Basically, the NNLO accuracy is necessary to achieve the PDFs precision of O(1%).

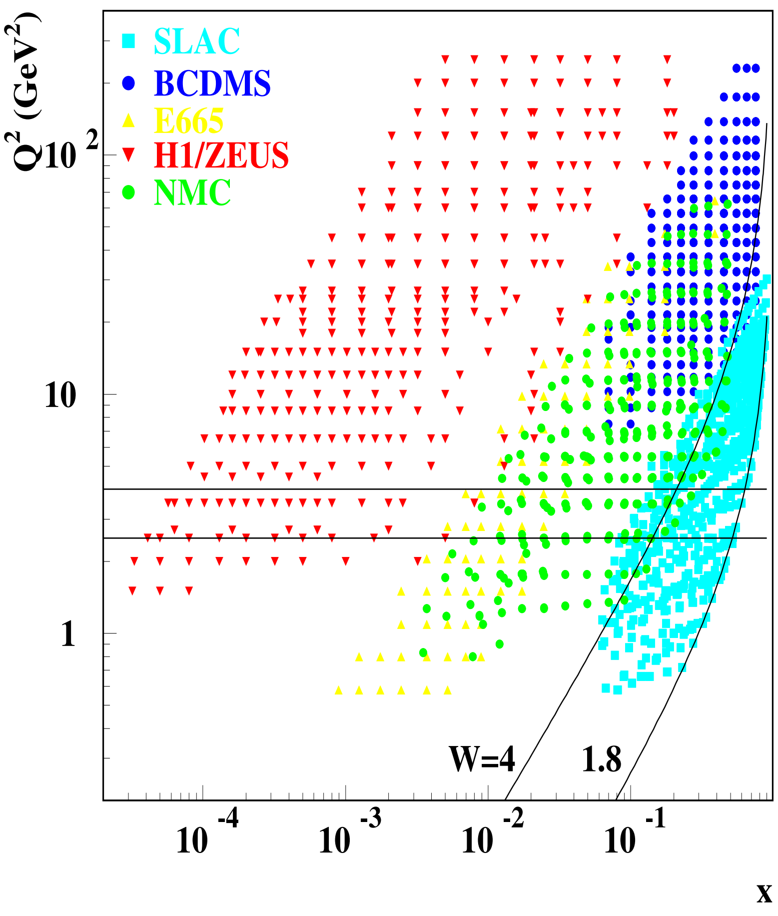
- *Inconsistency of the data coming from different experiments*

  Ideally, one single experiment should be used for the fit of PDFs (Dittmar-Pauss-Zurcher 97). In (current) practice, however, only combination of the data sets can provide constraints on the PDFs with the required accuracy.

- *Assessment of the experimental errors*

  For the existing data sets the systematic errors dominate.

# The global set of the charged-leptons DIS data




- Discrepancy between EMC and BCDMS was resolved in favour of the latter.
- Tension between H1 and ZEUS has been removed by tuning of the both data sets.
- Tension between the SLAC and BCDMS data.

(Whitlow 90)

5.4 $F_2$ Extraction and Modeling

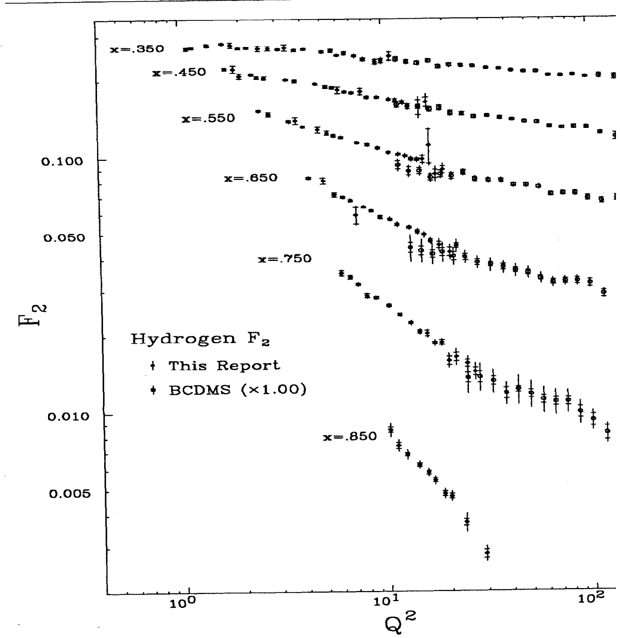



Figure 5.14/continued: Comparison of SLAC and BCDMS hydrogen $F_2$

128 Global Reanalysis of the SLAC Data

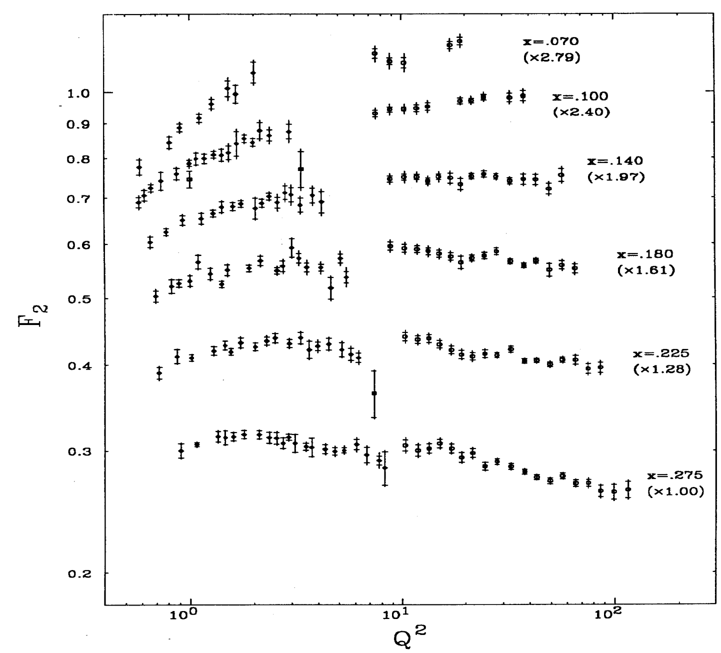



Figure 5.14. Shown is a comparison of the SLAC and BCDMS hydrogen $F_2$ results. A relative normalization of 1.000 is assumed. See also the caption to Figure 5.9. Figure continues on next page.

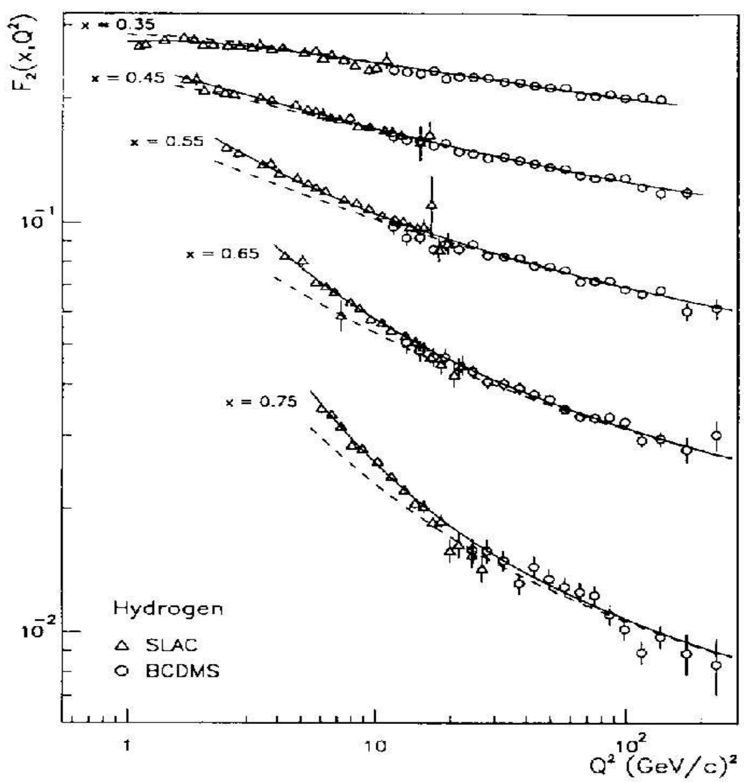
The SLAC and BCDMS data on $F_2$ are not perfectly consistent, particularly at $x \sim 0.2$, out of the region of overlap.

This problem was somehow resolved by Milsztajn-Virchaux, similarly to the way used in the H1/ZEUS tuning: The systematics of the BCDMS experiment was fitted in combined analysis of the SLAC/BCDMS data.

- the biggest BCDMS systematic errors were combined into one "main source" and others were combined in quadrature. The data were shifted by $\sim 1.3$ of the "main" systematic error.
- the NLO QCD corrections ($\sim \ln(Q)$) were taken into account.
- the power terms ($\sim 1/Q$) are necessary at low $Q$: the target-mass corrections by Georgi-Politzer-Nachtmann and the dynamical twist-4 terms with the coefficients $h(x)$.

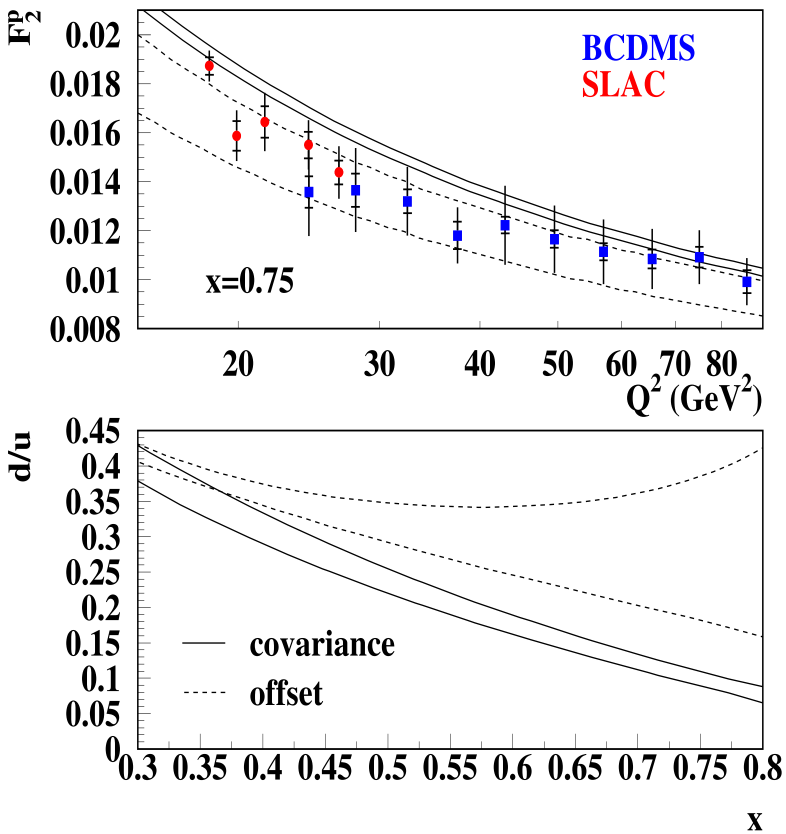
$$F_2 = F_2^{\mathrm{NLO}} \otimes \left[\mathbf{1} + \frac{M_N^2}{Q^2} C^{\mathrm{TMC}}\right] \cdot \left[1 + \frac{h(x)}{Q^2}\right]$$

(Virchaux-Milsztajn 91)




Due to the shifts consistency of the data is improved ($\chi^2/NDP = 599/738$), however obtained value of $\alpha_s(M_Z) = 0.113$, smaller than in the LEP determinations ($\sim 0.120$). *The need to add new light states (gluino?) to the strong coupling constant evolution.*

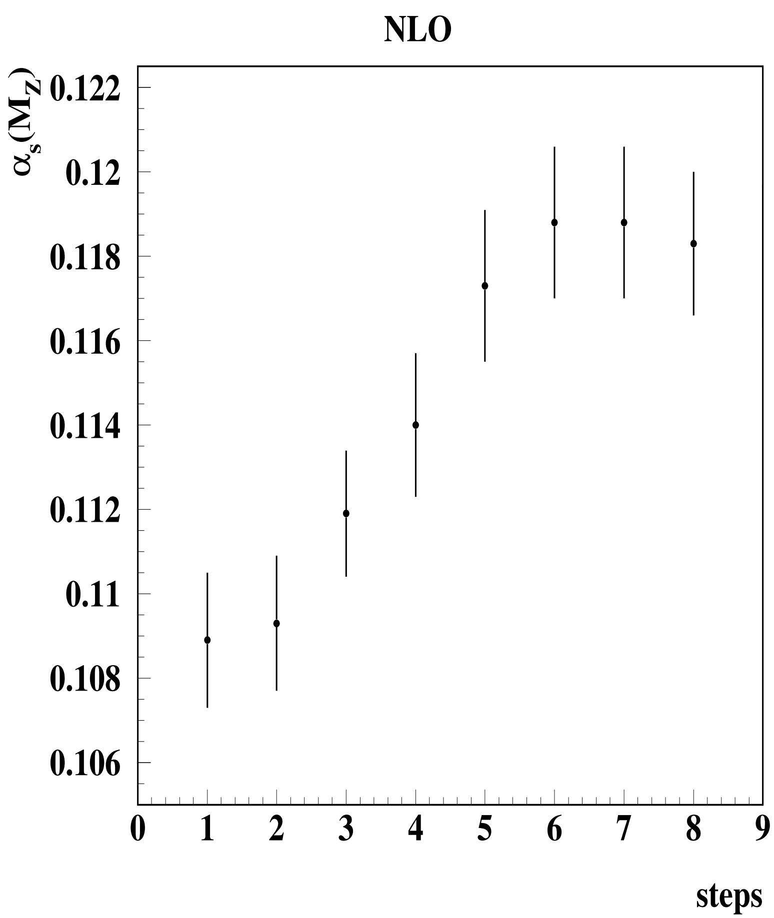
# Sensitivity of the fit to the errors treatment




With the offset method used the central value of $d/u$ is shifted as compared to the estimated based on the standard covariance matrix method and the errors are much bigger due to the offset method is statistically inefficient.

Table 1: The fitted $\alpha_s$ and the PDFs parameters values (I: the fit based on the covariance matrix estimator; II: the fit based on the offset estimator; III: the fit with the statistical and systematic errors combined in quadrature).

| | I | II | III |
|---|---|---|---|
| $a_u$ | $0.693 \pm 0.033(0.027)$ | $0.715 \pm 0.114(0.029)$ | $0.703 \pm 0.035$ |
| $b_u$ | $3.945 \pm 0.050(0.039)$ | $4.119 \pm 0.257(0.038)$ | $4.037 \pm 0.049$ |
| $\gamma_2^u$ | $1.29 \pm 0.44(0.37)$ | $1.39 \pm 1.86(0.40)$ | $1.42 \pm 0.49$ |
| $a_d$ | $0.725 \pm 0.086(0.082)$ | $0.703 \pm 0.172(0.094)$ | $0.717 \pm 0.13$ |
| $b_d$ | $4.93 \pm 0.13(0.12)$ | $4.83 \pm 0.27(0.17)$ | $5.00 \pm 0.17$ |
| $a_G$ | $-0.225 \pm 0.035(0.031)$ | $-0.169 \pm 0.065(0.029)$ | $-0.135 \pm 0.044$ |
| $b_G$ | $6.1 \pm 2.1(1.8)$ | $4.9 \pm 5.6(1.7)$ | $4.07 \pm 1.3$ |
| $\gamma_1^G$ | $-2.63 \pm 0.83(0.71)$ | $-3.41 \pm 0.99(0.45)$ | $-4.06 \pm 0.48$ |
| $\gamma_2^G$ | $4.7 \pm 2.9(2.4)$ | $4.44 \pm 3.4(1.3)$ | $5.41 \pm 1.2$ |
| $A_S$ | $0.166 \pm 0.011(0.0095)$ | $0.167 \pm 0.025(0.011)$ | $0.167 \pm 0.017$ |
| $a_{sd}$ | $-0.1987 \pm 0.0067(0.0050)$ | $-0.1853 \pm 0.0181(0.0050)$ | $-0.1833 \pm 0.0075$ |
| $b_{sd}$ | $5.1 \pm 1.4(1.3)$ | $5.4 \pm 2.8(1.4)$ | $4.9 \pm 2.1$ |
| $\eta_u$ | $1.13 \pm 0.11(0.087)$ | $1.10 \pm 0.23(0.086)$ | $1.16 \pm 0.16$ |
| $b_{su}$ | $10.29 \pm 0.97(0.81)$ | $10.56 \pm 3.2(0.83)$ | $11.2 \pm 1.1$ |
| $\alpha_s(M_Z)$ | $0.1165 \pm 0.0017(0.0014)$ | $0.1138 \pm 0.0044(0.0021)$ | $0.1190 \pm 0.0036$ |




Progressive refinement of the correlated errors treatment in the QCD analysis of the non-singlet SLAC/BCDMS data leads to increase of $\alpha_s$, the final value is in agreement to the LEP data.

The data on the DIS *cross sections* must be used in the fit. With account of the NNLO QCD, the target mass corrections, and the dynamical twist-4(6) terms

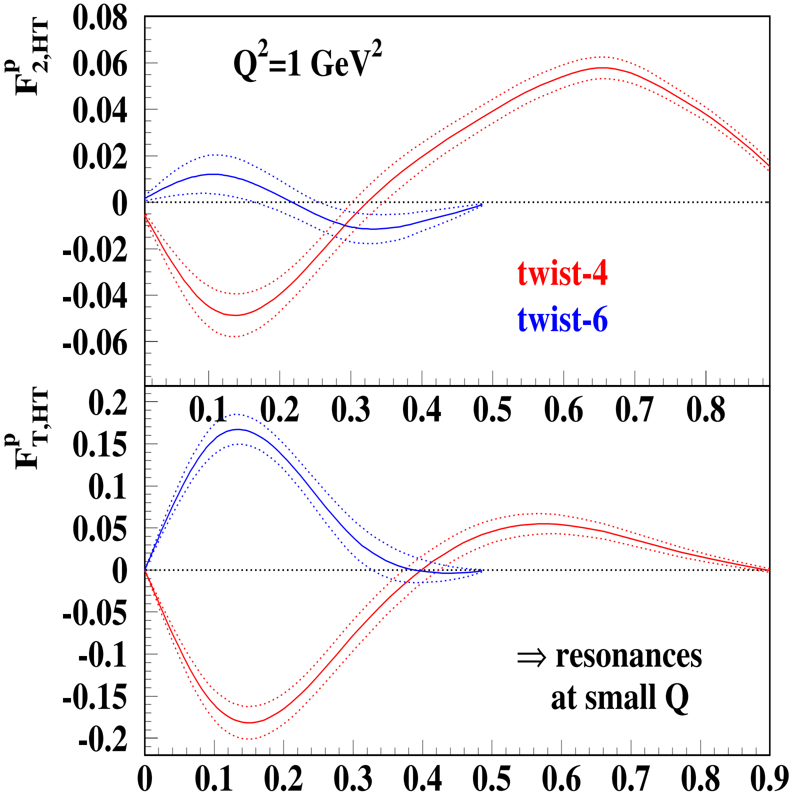
$$\overline{\sigma} = \left[1 - y - \frac{(M_N xy)^2}{Q^2}\right] F_2(x, Q^2) + \frac{y^2}{2} F_{\mathrm{T}}(x, Q^2)$$

$$F_{2,\mathrm{T}}(x, Q) = F_{2,\mathrm{T}}^{\mathrm{LT,TMC}}(x, Q) + \frac{H_{2,\mathrm{T}}^{(2)}(x)}{Q^2} + \frac{H_{2,\mathrm{T}}^{(4)}(x)}{Q^4} \qquad \text{(OPE)}$$

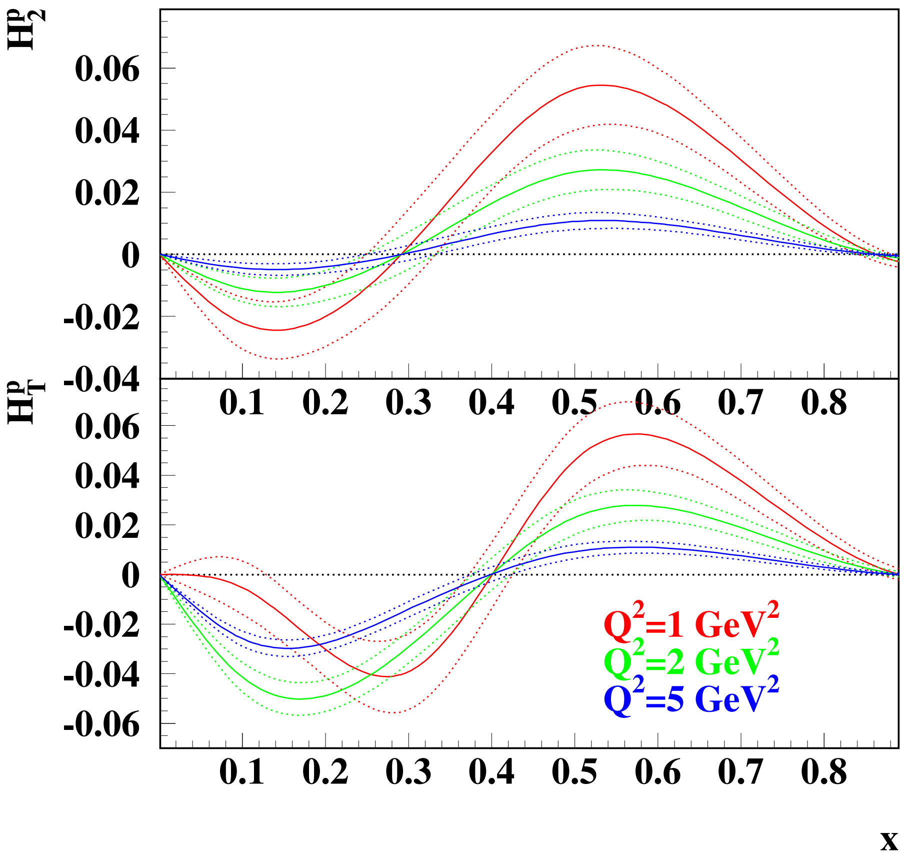
$$F_{2,\mathrm{T}}^{\mathrm{LT,TMC}}(x, Q) = F_{2,\mathrm{T}}^{\mathrm{LT}} \otimes \left[\mathbf{1} + \frac{M_N^2}{Q^2} C^{\mathrm{TMC}}\right] \qquad \text{(TMC)}$$

$$F_{2,\mathrm{T}}^{\mathrm{LT}} = \left[C_{2,\mathrm{T}}^{(0)} + \alpha_{\mathrm{s}} C_{2,\mathrm{T}}^{(1)} + \alpha_{\mathrm{s}}^2 C_{2,\mathrm{T}}^{(2)}\right] \otimes PDFs \qquad \text{(NNLO QCD)}$$

# High-twist terms in the fit with $Q^2 > 1$ GeV$^2$

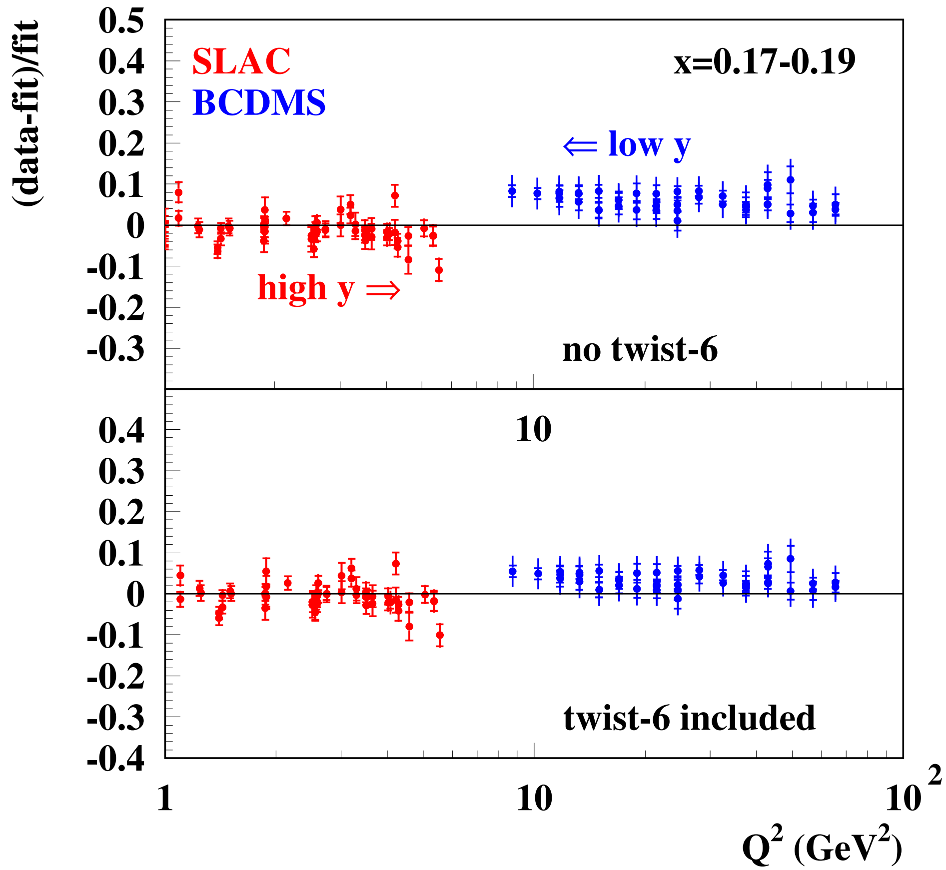



- The HT terms in $F_2$ demonstrate good convergence: $H_2^{(4)}$ is much smaller than $H_2^{(2)}$ and comparable to 0 within the errors.
- For $F_T$ the picture is different: the magnitudes of the twist-4 and twist-6 terms are comparable and somehow compensate each other (*poor convergence of the OPE?*)

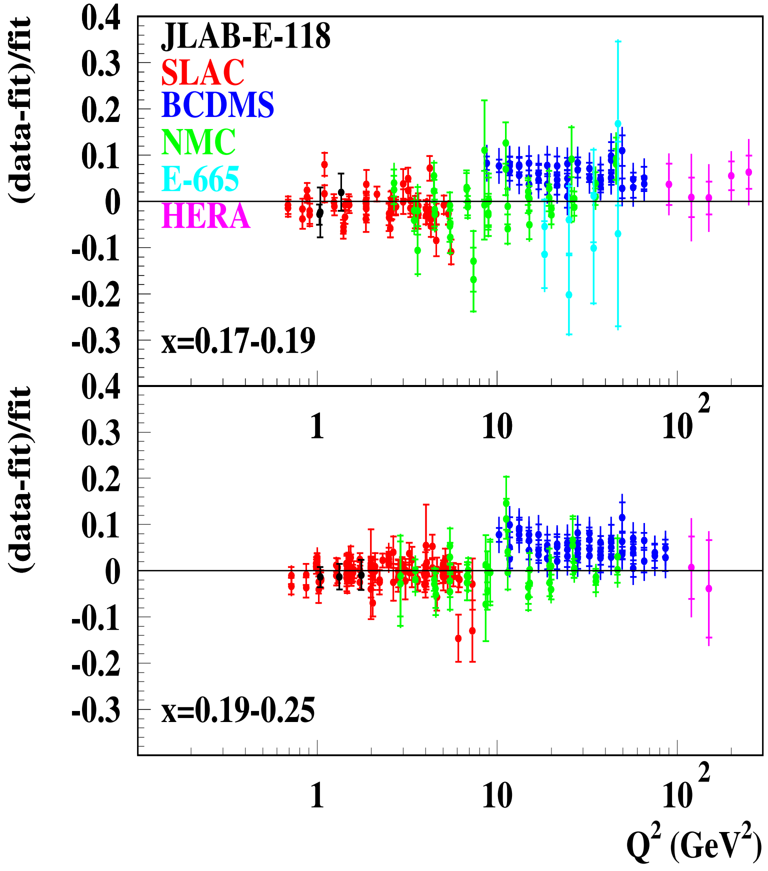



The total HT contribution to $F_\mathrm{T}$ demonstrates weak dependence on $Q$ at $x < 0.3$.

# Impact of the twist-6 terms on pulls of the fit

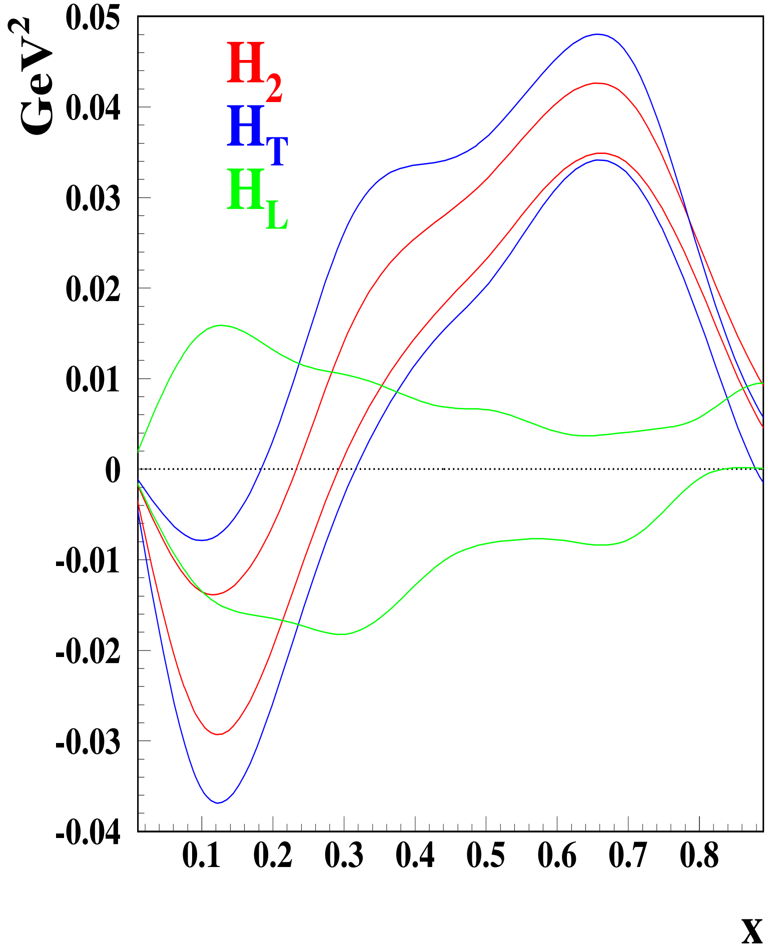



The twist-6 terms in $F_{\mathrm{T}}$ arise due to mismatch of the SLAC and BCDMS data at $Q^2 = 5 \div 10\ \mathrm{GeV}^2$ and *different* $y$.




- The NMC data cannot resolve this discrepancy in view of relatively big errors.
- The preliminary data by JLAB-E-118 are in agreement with low-Q part of the SLAC data.

## The HT terms in the charged-leptons SFs




In the fit with twist-6 terms set to zero the twist-4 terms in the charged-lepton $F_2$ and $F_{\rm T}$ averaged over proton and neutron are very similar within the errors; therefore the HT term in $F_L$ is comparable to 0.

The high-twist term in

$$R = \frac{\sigma_L}{\sigma_T} = \frac{F_2}{F_T}\left(1 + \frac{4M_N^2 x^2}{Q^2}\right) - 1$$

is also small.

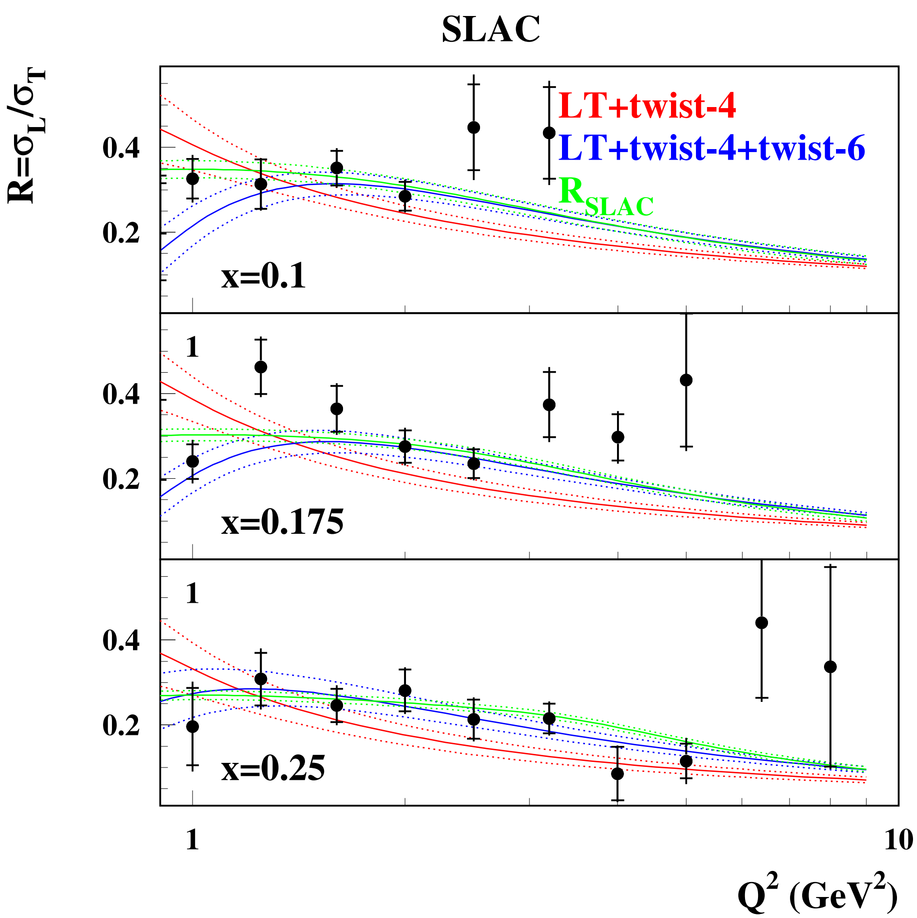
Meanwhile the parameterization of the SLAC data on $R$ includes the log-like and the power-like terms.

$$R^{SLAC} = R_{LOG} + A/Q^2 + B/(Q^4 + C^2)$$

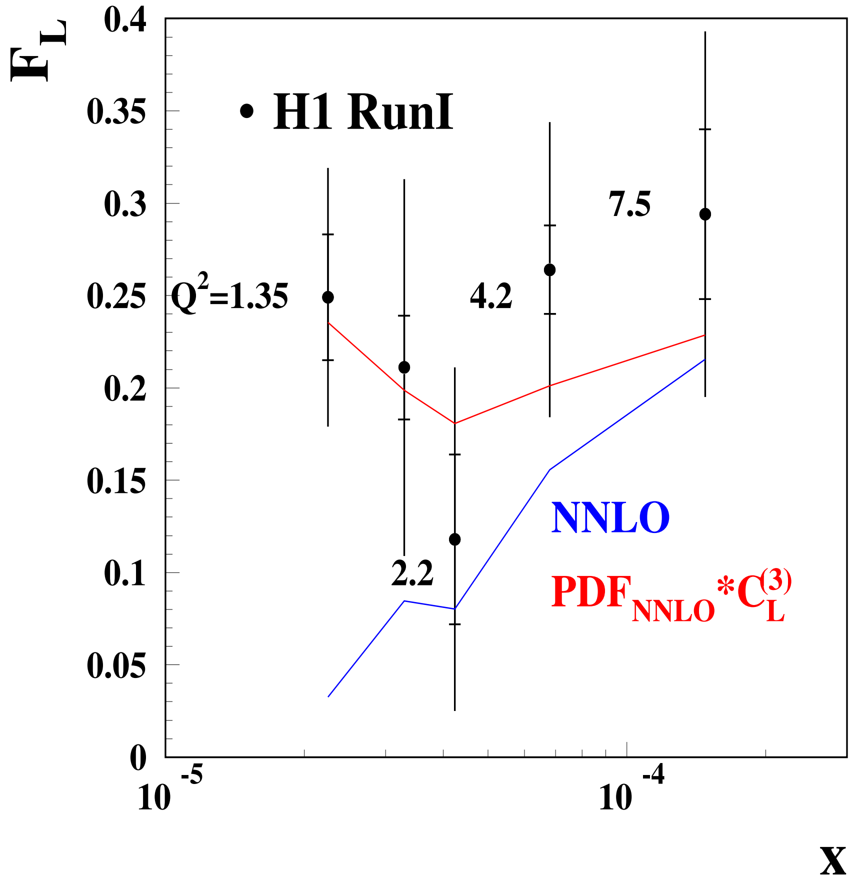
with $A = 0.6$, $B = -0.35$, $C = 0.3$, $Q$ is in GeV.

Further the excess in SLAC data on R at $x \sim 0.2$ with respect to the QCD predictions was considered as evidence of the big HT contribution to R (and $F_{\mathrm{L}}$)

(Miramontes-... 89)




The excess in SLAC data on R at $x \sim 0.2$ with respect to the QCD predictions is evidently connected with the SLAC/BCDMS discrepancy and can be hardly attributed to the HT contribution.



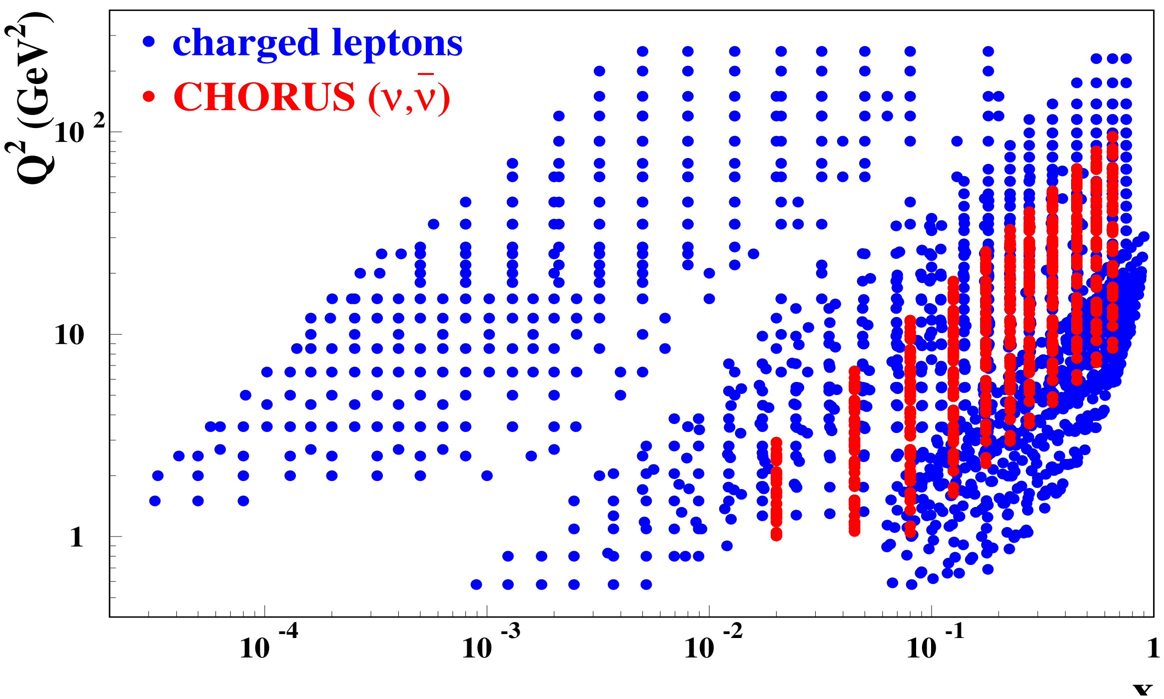

- With the $N^3LO$ corrections to the coefficient functions by Vermasseren-Moch-Vogt the agreement to data is good, alternatively one could expect big power correction.
- In the renormalon model

$$H(x) = \Lambda^2 \int_x^1 C_{IRR}(z) q(x/z) \frac{dz}{z}$$

and large HT terms at large $x$ automatically leads to the same at small one.

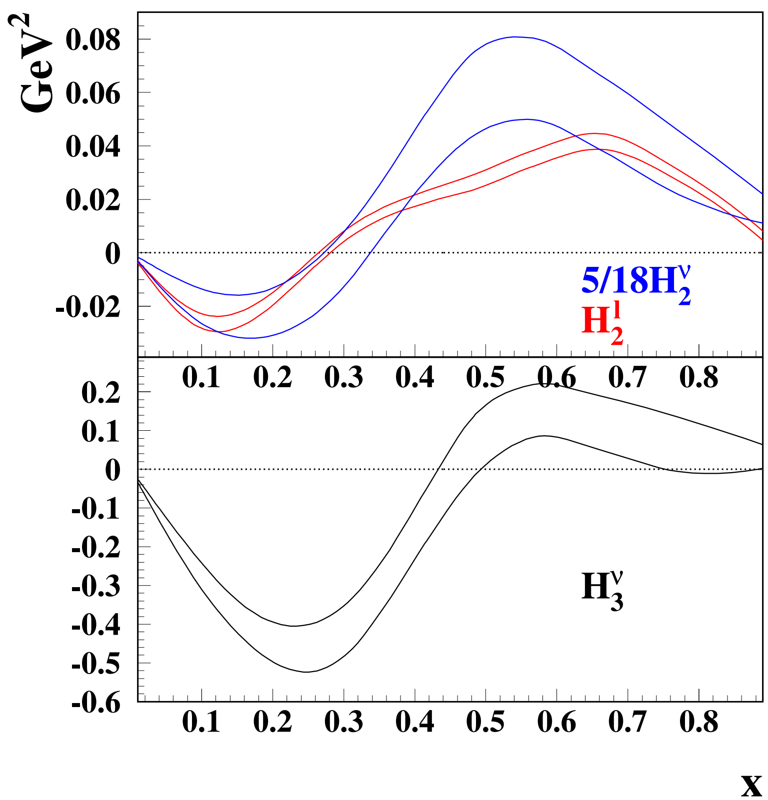
(Stein-... 98).

# The $\nu N$ DIS data



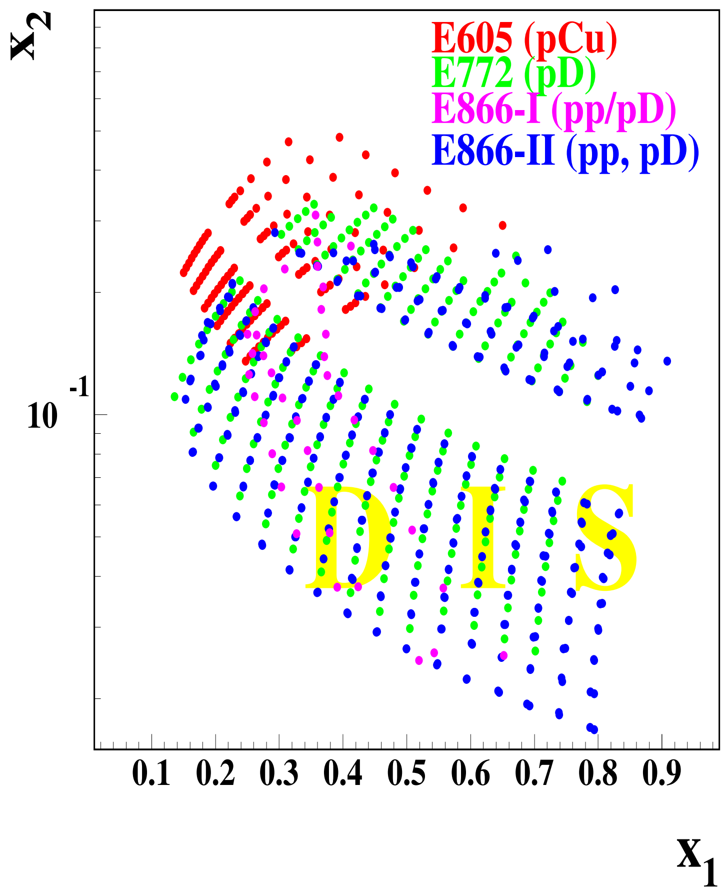

$$\overline{\sigma} = \left[1 - y - \frac{(M_N xy)^2}{Q^2}\right] F_2(x, Q^2) + \frac{y^2}{2} F_{\mathrm{T}}(x, Q^2) \pm y \left[1 - \frac{y}{2}\right] xF_3(x, Q^2)$$

# The twist-4 terms in the $\nu N$ DIS




- $H_2^{\nu N} = H_{\mathrm{T}}^{\nu N}$, motivated by the charged-leptons fit.
- $H_2^{\nu N}$ is in remarkable agreement to $H_2^{lN}$ rescaled with the quarks charge.
- $\int H_3^{\nu N}(x)dx = -0.10 \pm 0.03$ GeV$^2$, in agreement to the early calculations by Braun-Kolesnichenko.

# The fixed-target Drell-Yan data in the global fit

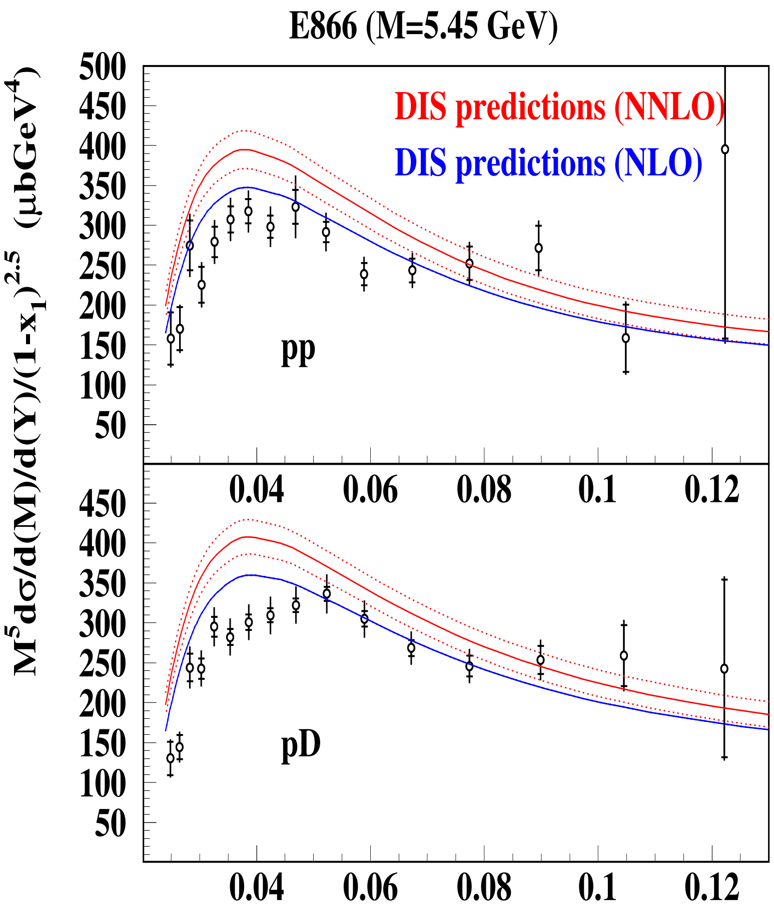



- The c.s. of Drell-Yan process provide complimentary constraint to the DIS data and allows separation of the quark and anti-quark distributions

$$\sigma_{\mathrm{DIS}} \sim q(x) + \bar{q}(x)$$

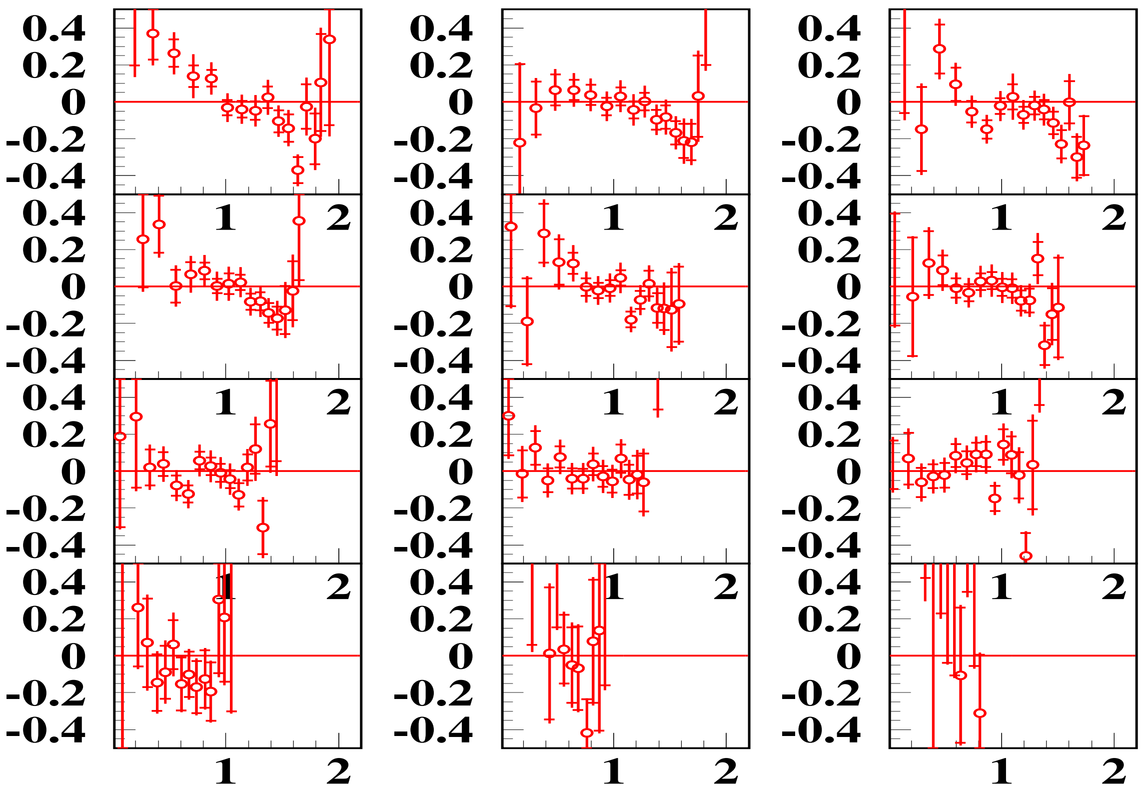
$$\sigma_{\mathrm{DY}} \sim q(x_1)\bar{q}(x_2) + q(x_2)\bar{q}(x_1)$$

- At large $x_1$ and small $x_2$ $\sigma_{\mathrm{DY}} \sim q(x_1)\bar{q}(x_2)$ and $\Delta q(x_1), \Delta\bar{q}(x_2) \sim O(1\%)$ from DIS

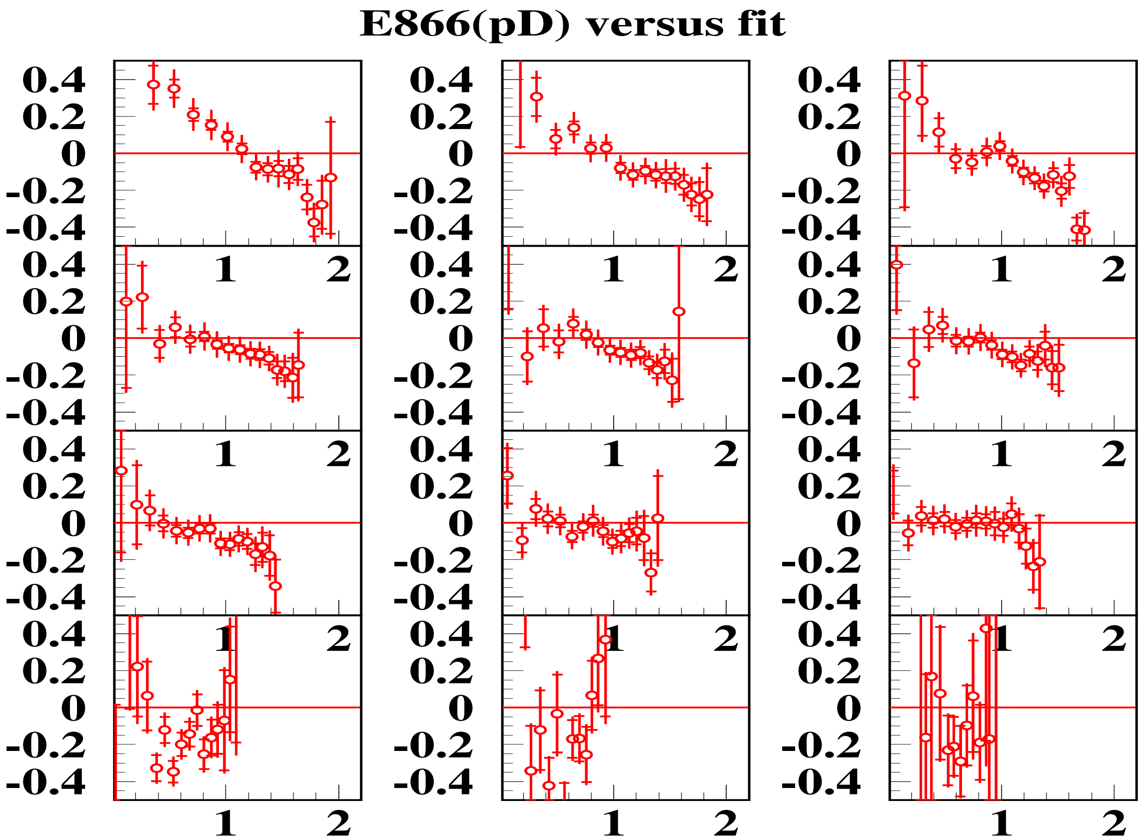



- The NNLO corrections by Anastasiou-Dixon-Melnikov-Petriello are crucial for this comparison: *In the NLO agreement is better!*
- The absolute cross section DY data by E-866 at low $\mu\mu$ masses are incompatible with the global DIS data and cannot be included into the fit.

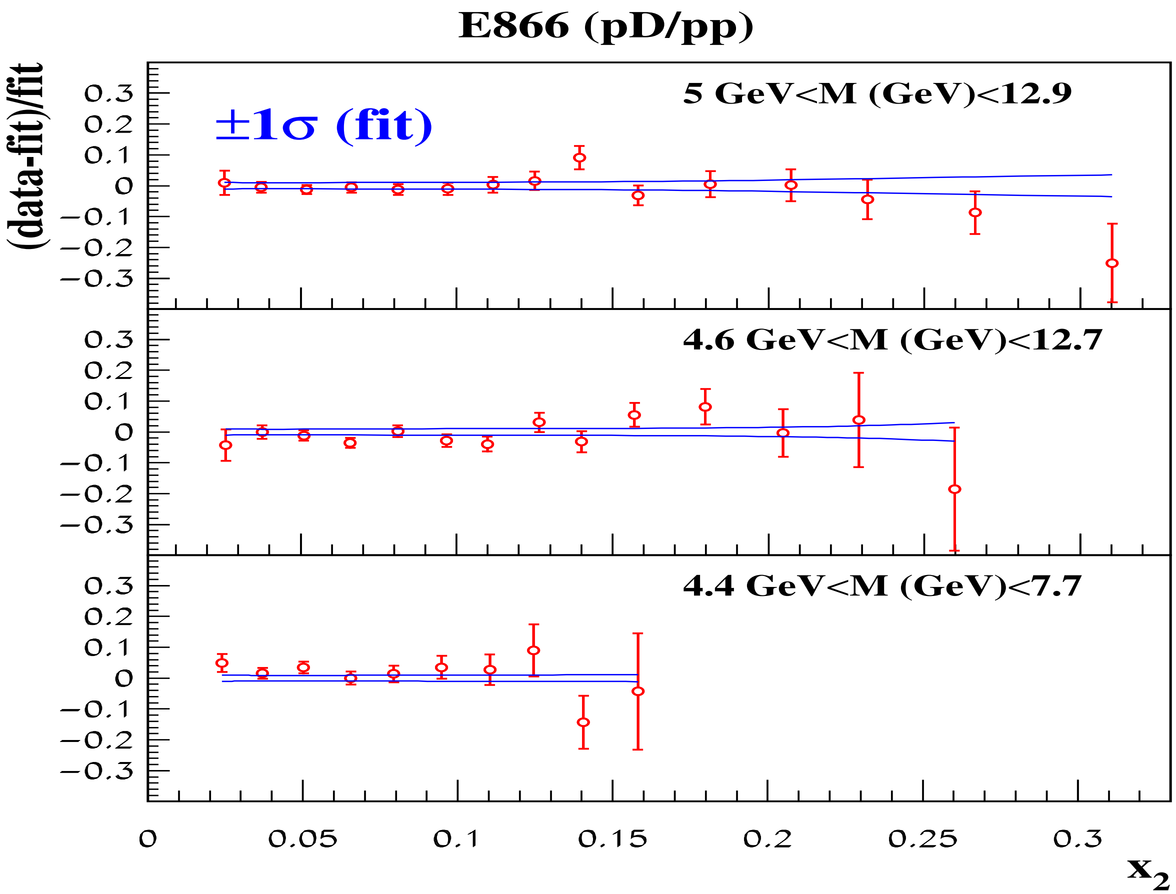
E866(pp) versus fit



$\chi^2 = 322(209)/184$ with(out) account of the errors correlation.

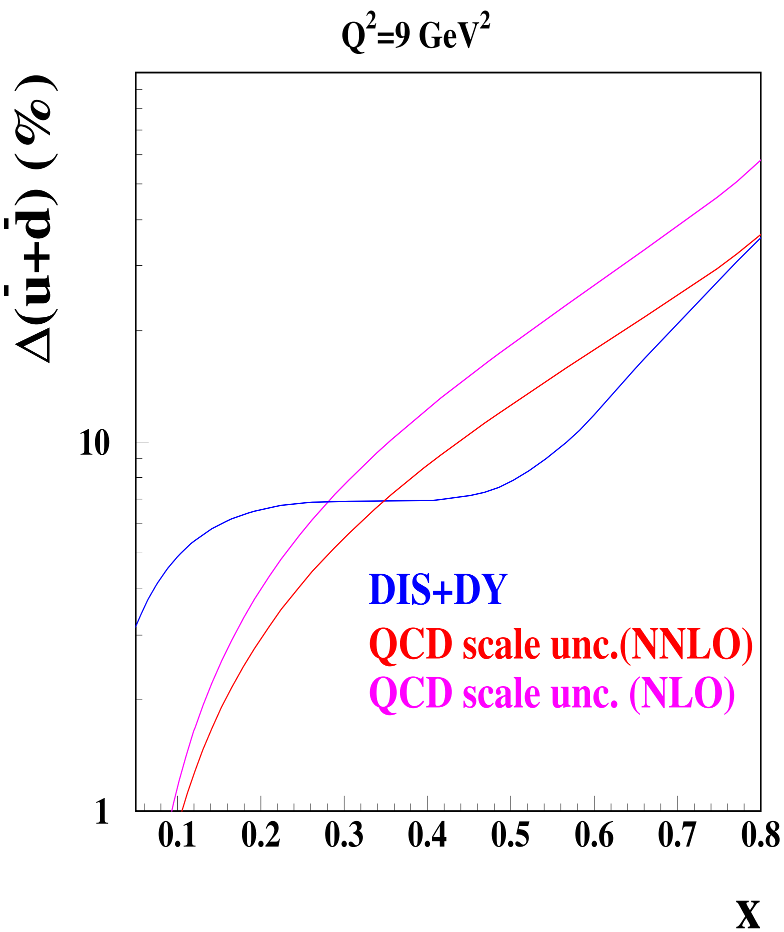



The deuteron data have the same trend as the proton ones.




In the ratio the systematic shifts in the E866 pp and pD data cancel.

# Impact of the DY data on the sea distribution




- The combination of DIS and DY data by E605 (pCu) and E866 (pD/pp) constraint the sea with the error $< 20$ % at $x \lesssim 0.7$.
- The errors in PDFs due to variation of the QCD scales are comparable to the experimental ones (*NNLO corrections are crucial*).

## Strange sea from the dimuon neutrino data

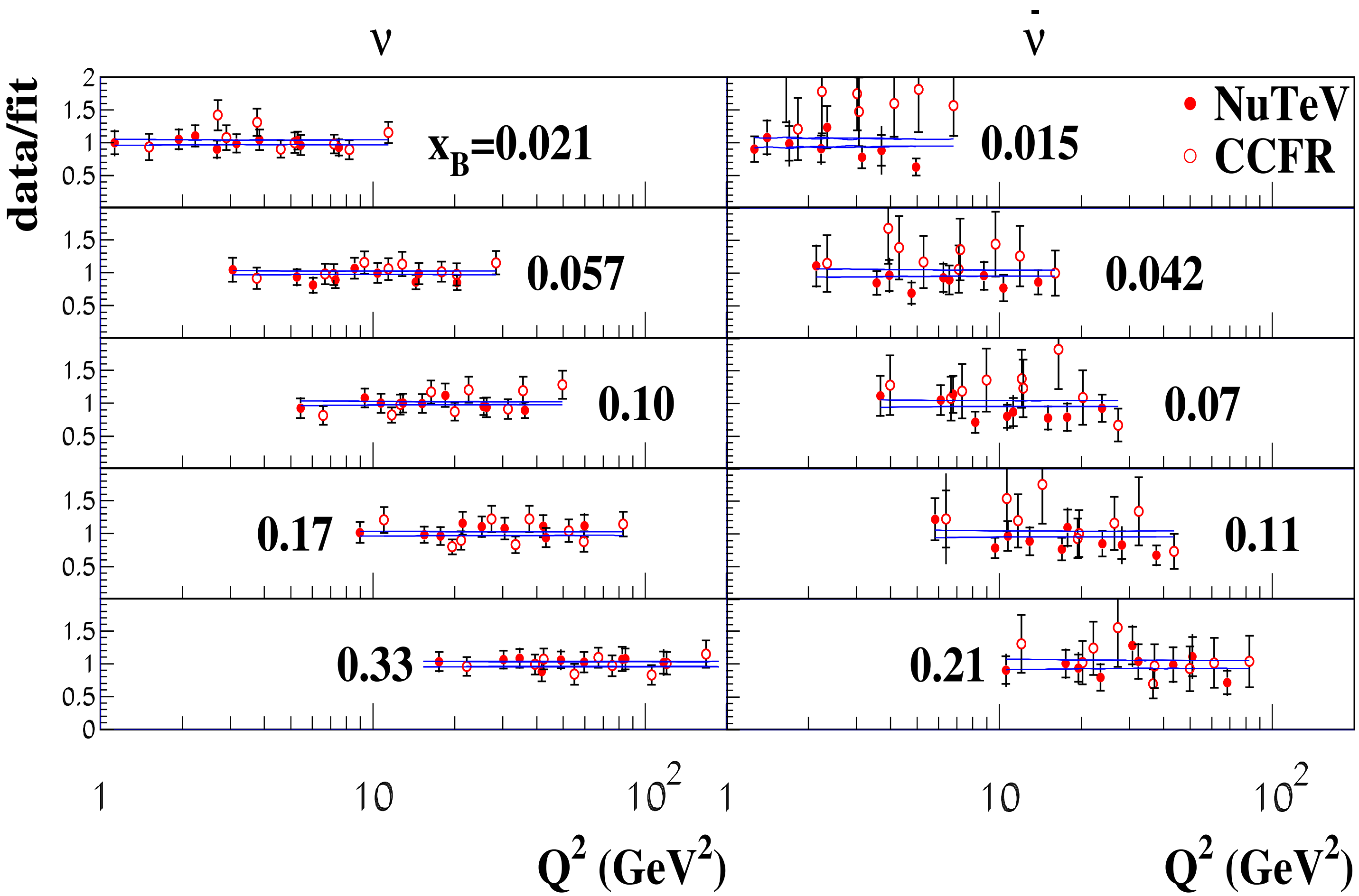
$$\frac{d\sigma^{\overset{(-)}{\nu}N}_{\mu\mu}}{dxdy} = B_\mu \frac{d\sigma^{\overset{(-)}{\nu}N}_{\mathrm{charm}}}{dxdy}$$

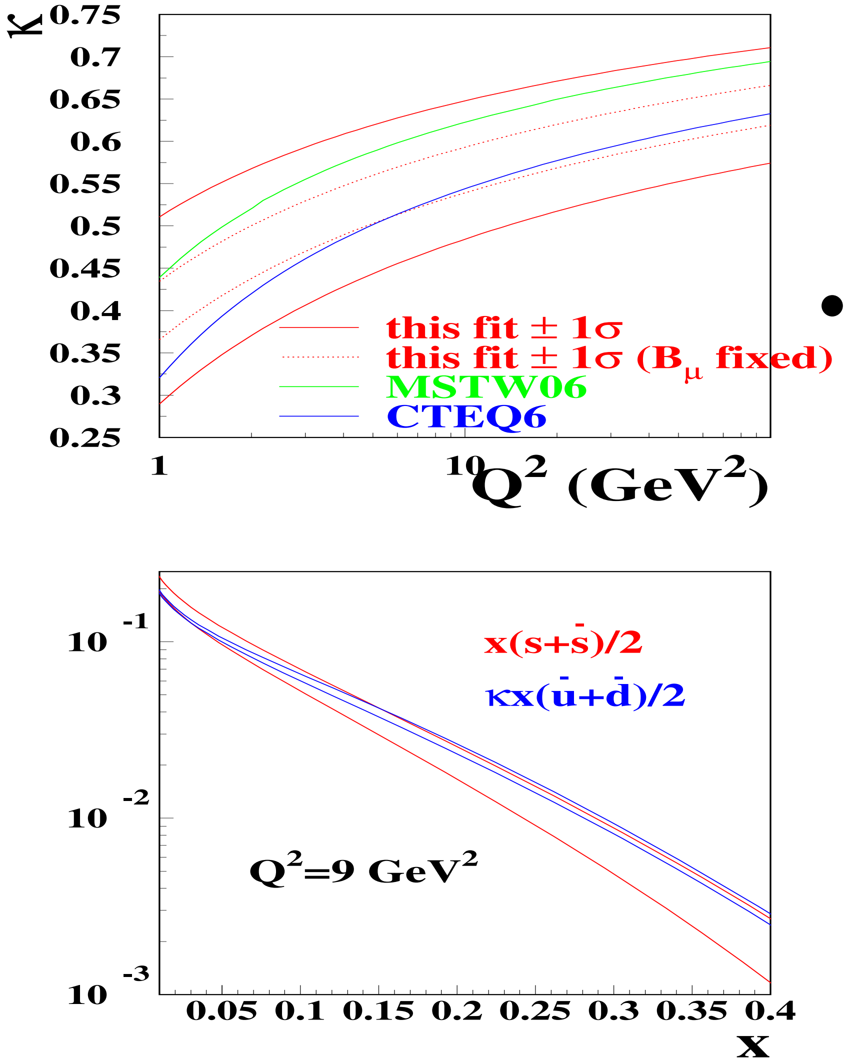
$$B_\mu = \sum f_h Br(h \to \mu X)$$

$$\frac{d\sigma^{\overset{(-)}{\nu}N}_{charm}}{dxdy} = \frac{G_F^2 M E_{\overset{(-)}{\nu}}}{\pi(1+Q^2/M_W^2)^2}\left[\left(1 - y - \frac{Mxy}{2E}\right) F_{2,c}^{\overset{(-)}{\nu}N}(x,Q) + \right.$$
$$\left. + \frac{y^2}{2} F_{T,c}^{\overset{(-)}{\nu}N}(x,Q) \overset{+}{\underset{(-)}{}} y\left(1 - \frac{y}{2}\right) x F_{3,c}^{\overset{(-)}{\nu}N}(x,Q)\right]$$

$$F_{2,c}^{\overset{(-)}{\nu}N}(x,Q) = 2\xi\left[|V_{cs}|^2 \overset{(-)}{s}(\xi,\mu) + |V_{cd}|^2 \frac{\overset{(-)}{u}(\xi,\mu) + \overset{(-)}{d}(\xi,\mu)}{2}\right]$$

$$F_{T,c}^{\overset{(-)}{\nu}N} = \overset{+}{\underset{(-)}{}} x F_{3,c}^{\overset{(-)}{\nu}N} = \frac{x}{\xi} F_{2,c}^{\overset{(-)}{\nu}N}$$


ν
ν̄
data/fit
NuTeV
CCFR
x_B=0.021
0.057
0.10
0.17
0.33
0.015
0.042
0.07
0.11
0.21
Q^2 (GeV^2)
Q^2 (GeV^2)




- We do not see energy and beam dependence of $B_\mu$, the average value of $B_\mu = 0.091 \pm 0.010$ is obtained.
- The value of the strange sea suppression factor

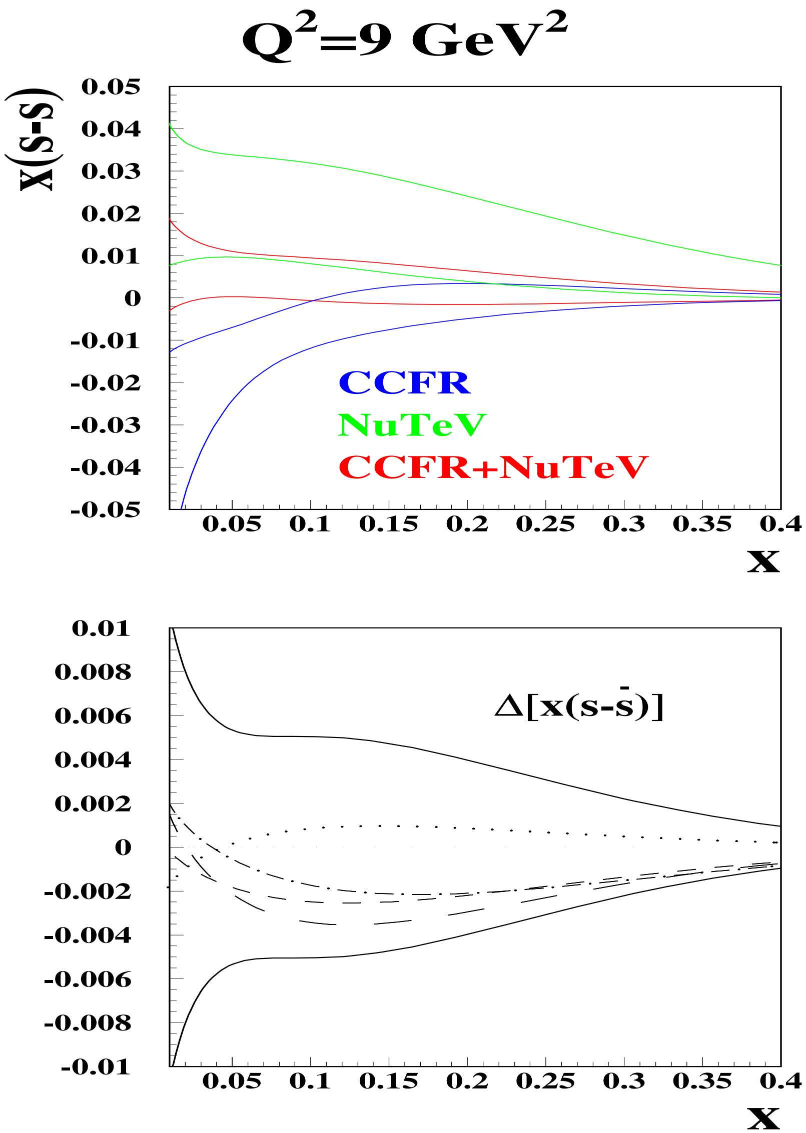
$$\kappa = \frac{\int_0^1 x\,[s(x) + \overline{s}(x)]\,dx}{\int_0^1 x\,\left[\overline{u}(x) + \overline{d}(x)\right]\,dx}$$

depends on $Q$ and $\kappa(Q^2 = 20\ \mathrm{GeV}^2) = 0.59 \pm 0.08$. This is bigger than $\kappa(20\ \mathrm{GeV}^2) = 0.48^{+0.06}_{-0.05}$ obtained in the CCFR NLO QCD fit.

## The constraints on $B_\mu$ from the emulsion experiments

| Measurement | $E_\nu > 5$ GeV | $E_\nu > 30$ GeV |
|---|---|---|
| CHORUS (KayisTopaksu 05) | $7.30 \pm 0.82\%$ | $8.50 \pm 1.08\%$ |
| CHORUS (DiCapua 08) | $9.11 \pm 0.93\%$ | |
| E531 (Bolton 97) | $7.86 \pm 0.49\%$ | $8.86 \pm 0.57\%$ |
| Weighted average | $7.94 \pm 0.38\%$ | $8.78 \pm 0.50\%$ |

With the additional input for $B_\mu$ we get substantial improvement in the strange sea magnitude, $\kappa(20 \text{ GeV}^2) = 0.62 \pm 0.04$.



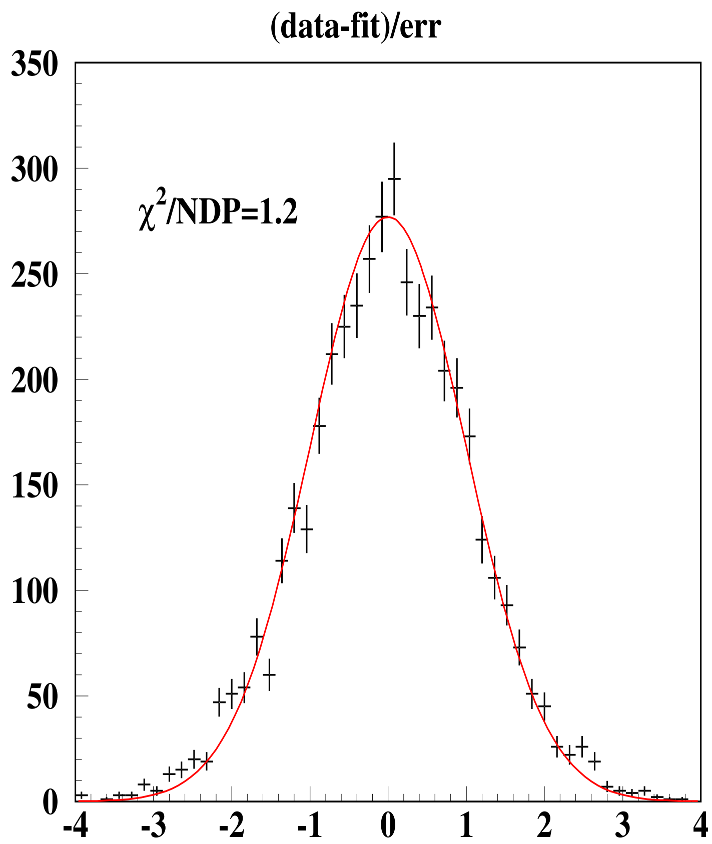

- The $s/\overline{s}$ asymmetry is comparable to zero within the errors, at $Q^2 = 20$ GeV$^2$ $\int_0^1 x[s(x) - \overline{s}(x)]dx = 0.0010(13)$ $(0.0013(9)$ with the additional input for $B_\mu$).
- The asymmetry has different sign for the CCFR and NuTeV cases and is sensitive to the choice of the QCD scale, value of $B_\mu$, to the nuclear corrections, etc.

## The NNLO fit of PDFs

- The global inclusive DIS data on the *cross sections* with the transferred momentum $Q^2 > 1$ GeV$^2$.
  The fixed target Drell-Yan data by FNAL-E-605 (p Cu) and FNAL-E-866 (pD/pp).
  Inclusive $\nu N$ *cross section* data by CHORUS.
  Data on dimuon production in the $\nu N$ interactions by the CCFR and NuTeV collaborations.

- Splitting Functions (up to $O(\alpha_s^3)$) (Moch-Vermasseren-Vogt 04)
  Massless quarks coefficient functions (up to $O(\alpha_s^3)$) (Vermasseren-Moch-Vogt 05)
  Heavy quark NC coefficient functions (up to $O(\alpha_s^2)$) (Laenen-Riemersma-Smith-van Neerven 92-93)
  Heavy quark CC coefficient functions (up to $O(\alpha_s)$) (Gottschalk 81)

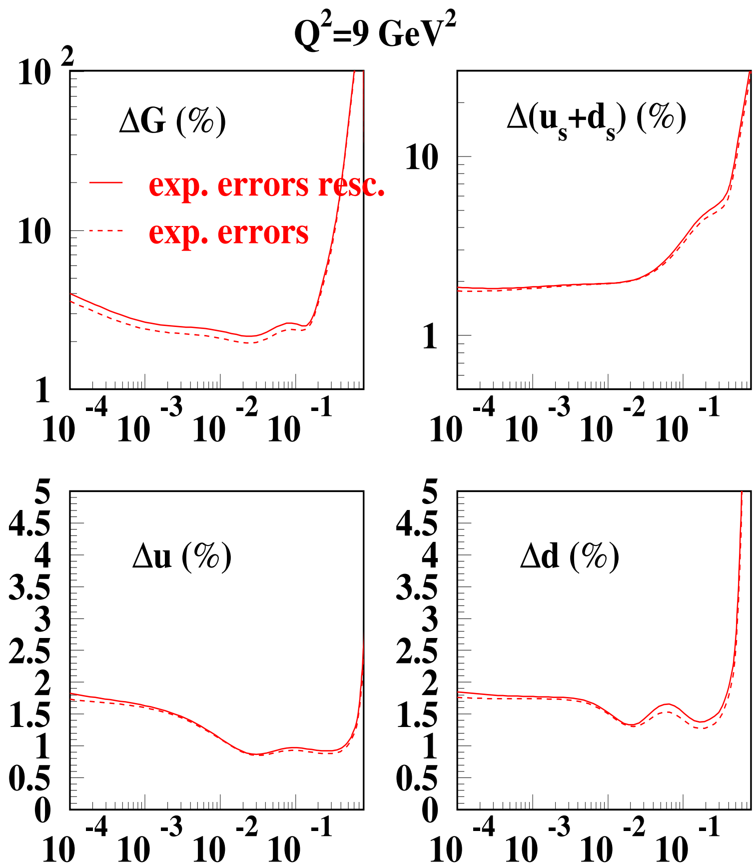
# The pulls distributions




For the fit of PDFs to the global DIS data distribution of the errors, including systematic ones, does obey the Gaussian one. One can evidently apply criterion $\Delta\chi^2 = 1$ for the errors estimation.

| Experiment | NDP | $\chi^2/NDP$ | scale factor |
|---|---|---|---|
| SLAC-E-49A | 214 | 0.51 | – |
| SLAC-E-49B | 389 | 1.30 | 1.15 |
| SLAC-E-61 | 44 | 0.31 | – |
| SLAC-E-87 | 218 | 0.99 | – |
| SLAC-E-89A | 148 | 1.45 | 1.21 |
| SLAC-E-89B | 216 | 1.10 | 1.05 |
| SLAC-E-139 | 22 | 0.83 | – |
| SLAC-E-140 | 46 | 1.22 | 1.11 |
| BCDMS | 605 | 1.16 | 1.08 |
| NMC | 578 | 1.39 | 1.19 |
| E665 | 130 | 1.59 | 1.30 |
| H1(96-97) | 147 | 1.32 | 1.16 |
| ZEUS(96-97) | 161 | 1.45 | 1.22 |
| E605 | 119 | 1.46 | 1.22 |
| E866 | 39 | 1.34 | 1.16 |
| NuTeV($\mu\mu$) | 89 | 0.47 | – |
| CCFR($\mu\mu$) | 89 | 0.69 | – |
| CHORUS | 1084 | 1.23 | 1.11 |

# The PDFs uncertainties




- The PDFs errors are calculated using the covariance matrix method with account of the systematic errors correlations, $\Delta\chi^2 = 1$.
- Rescaling of the errors has marginal impact on the PDFs errors.

## Summary

- The NNLO fit of the PDFs to the combined charged leptons DIS, fixed-target Drell-Yan, dimuon neutrino data by NuTeV and CCFR, and inclusive neutrino data by CHORUS demonstrates reasonable consistency of the data down to $Q = 1$ GeV: $\chi^2/NDP = 1.2$ for $NDP = 4339$; the rescaling factors on the errors are within 1.3 (cf. PDG tables).

- The HT terms extracted from the fit demonstrate remarkable universality: $H_2^{lN} \approx H_{\mathrm{T}}^{lN} \approx 5/18 H_{2,\mathrm{T}}^{\nu N}$ and are stable with respect to the NNLO corrections; this gives a confidence in their dynamical nature.

- The errors on PDFs obtained with the standard criterion $\Delta\chi^2 = 1$ are in agreement with the qualitative estimates and provide a good tool for the precision studies at the LHC: $\Delta p \sim 2\%$ at $x \lesssim 0.1$.

# The shape of PDFs

The general form of PDFs is

$$xp(x, 3GeV) = Ax^{\alpha}(1-x)^{\beta}x^{P(x)}$$

$$P(x) = ax + bx^2 + ...$$

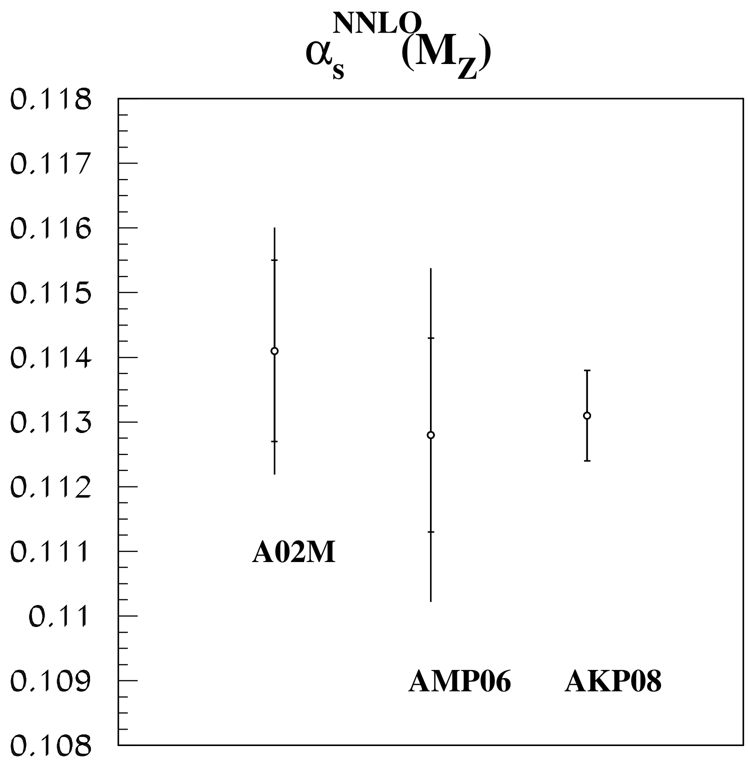
The polynomial factor provides sufficient flexibility of PDFs with respect to data used in the fit keeping the low-$x$ and high-$x$ asymptotics.

The sea quarks isospin asymmetry is motivated by the Regge asymptotic

$$x(\overline{u} - \overline{d})(x, 3GeV) = Ax^{0.5\pm0.2}(1-x)^{\beta}x^{P(x)}$$

In some variants of the fit the strange/anti-strange asymmetry is

$$(s - \overline{s})(x, 3GeV) = Ax^{-0.5}(1-x)^{\beta}(1-\gamma x).$$




The value of $\alpha_s$ is stable with respect to the previous determinations. For AKP08 $\alpha_s(3\ \mathrm{GeV}) = 0.2306(32)$ in the NNLO that correspond to $\alpha_s(M_Z) = 0.1131 \pm 0.0007(\mathrm{exp})$, in agreement to 0.1134(18) obtained by Blümlein-Boettcher-Guffanti in the non-singlent fit.