# Average of DIS cross section

S Glazov, DESY.

- Definition of an average.
- $\chi^2$ forms and biases.
- HERA average results.

# Definition of an Average

Consider an experiment which provides set of statistically uncorrelated measurements $M_i^e$ with systematic uncertainties $\alpha_j^e$. The PDF for $M_i^{true}$ and $\alpha_j^{true}$ can be then represented as

$$\chi_e^2(M_i^{true},\alpha_j^{true}) = \sum_i \left( \frac{M_i^{true} - [M_i^e + \sum_j \beta_{ij}^e \alpha_j^{true}]}{\sigma_i^e} \right)^2 + \sum_j (\alpha_j^{true})^2. \quad (1)$$

where $\beta_{ij}^e$ is sensitivity of the measurement at $i$ to the systematic source $j$. For many experiments,

$$\chi_{tot}^2(M_i^{true},\alpha_j^{true}) = \sum_e \chi_e^2(M_i^{true},\alpha_j^{true}). \quad (2)$$

We define as an "average" of the experiments as the following representation of $\chi_{tot}^2$:

$$\chi_{tot}^2(M_i^{true},\gamma_j^{true}) = \chi_{ave}^2 + \sum_i \left( \frac{M_i^{true} - [M_i^{ave} + \sum_j \beta_{ij}^{ave} \gamma_j^{true}]}{\sigma_i^{ave}} \right)^2 + \sum_j (\gamma_j^{true})^2 \quad (3)$$

where $\gamma_j^{true} = \sum_k A_{jk}(\alpha_k^{true} - \alpha_k^{ave})$ and $A$ is an orthogonal matrix.

# Definition of an Average II

In other words and average of experiments has the $\chi^2$ form of a single experiment with the set $M_i^{ave}$ corresponding to all $M_i^e$ but different uncorrelated systematic sources. One can observe:

- An average exists (by checking coefficients in front of $M_i^{true}, \alpha_j^{true}$ in Eq.2 vs Eq.3)
- It is diagonal in $M_i^{true}$ (since Eq.2 contains no off-diagonal terms).
- $M_i^{ave}$ and $\alpha_k^{ave}$ correspond to the minimum of Eq.2, $\chi^2_{ave}$ is the value at the minimum.

If all $\beta_{ij}^e = 0$, this definition coincides with a standard average. Other good properties of an average are preserved: $M_i^{ave}, \sigma_i^{ave}, \sigma_i^{ave\ tot} = \sqrt{(\sigma_i^{ave})^2 + \sum_j (\beta_{ij}^{ave})^2}$ of 3 experiments does not depend on averaging sequence, etc.

Since the form of Eq.1 is simple, finding of the minimum is not hard as well.

## Bias to Smaller $X$-section values

Most of the systematic uncertainties of the X-section measurements, correlated and uncorrelated are given in terms of relative errors. In absolute space, measurement with smaller value has smaller uncertainty $\rightarrow$ bias.

Consider 2 measurements of $X$, one at $X + X\beta$ another at $X - X\beta$, both with the same relative uncertainty $\delta = \sigma/X$. An error weighted average of the two measurements returns

$$\bar{x} = X\frac{1-\beta^2}{1+\beta^2},$$

which for $\beta = 5\%$ corresponds to $0.5\%$ bias.

# $\chi^2$ definitions and biases

A way to avoid the bias is to replace simple quadratic in $M^{true}$ and $\alpha^{true}$ definition of the $\chi^2$

$$\chi^2(M_i^{true}, \alpha_j^{true}) = \sum_i \frac{[M_i^{true} - (M_i + \sum_j \beta_{ij}\alpha_j^{true})]^2}{\sigma_i^2} + \sum_j \left(\alpha_j^{true}\right)^2 \quad (4)$$

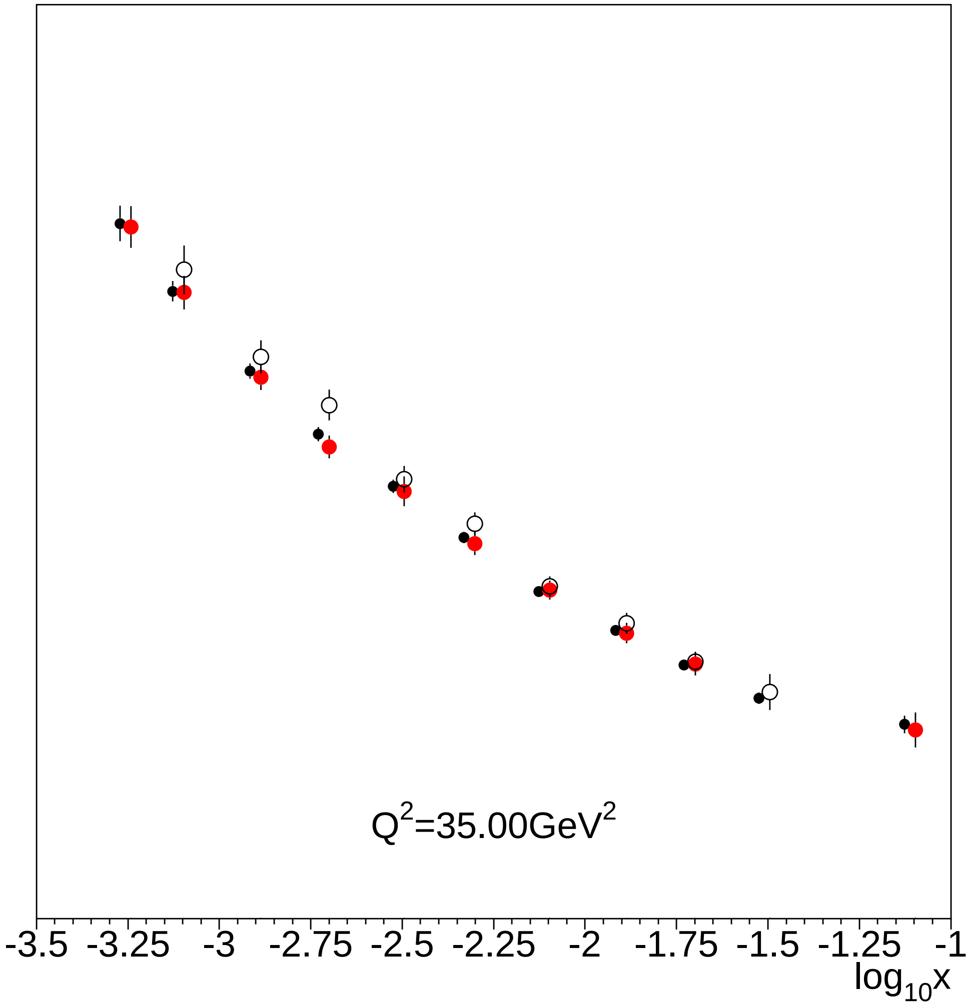
to:

$$\chi^2(M_i^{true}, \alpha_j^{true}) = \sum_i \frac{\left[M_i^{true}\left(1 + \frac{1}{M_i}\sum_j \beta_{ij}\alpha_j^{true}\right) - M_i\right]^2}{\left[M_i^{true}\frac{\sigma_i}{M_i}\right]^2} + \sum_j \left(\alpha_j^{true}\right)^2, \quad (5)$$

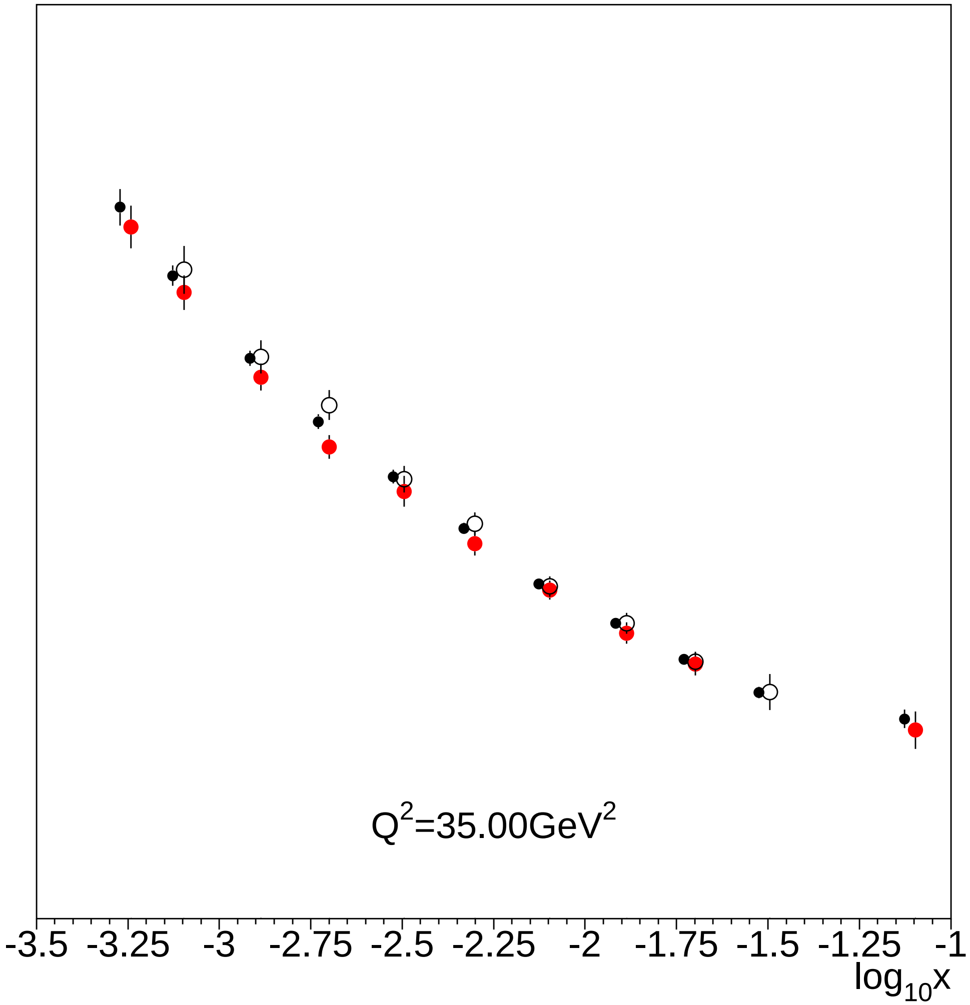
where all relative errors are translated to absolute around the average value.

Bias is studied using toy MC techniques. The main contribution comes from global normalization errors. Average for $\chi^2$ definition of Eq. 5 is found iteratively starting from Eq. 4.

# Average of H1 and ZEUS for $Q^2 = 35$ GeV$^2$ (Eq. 4)



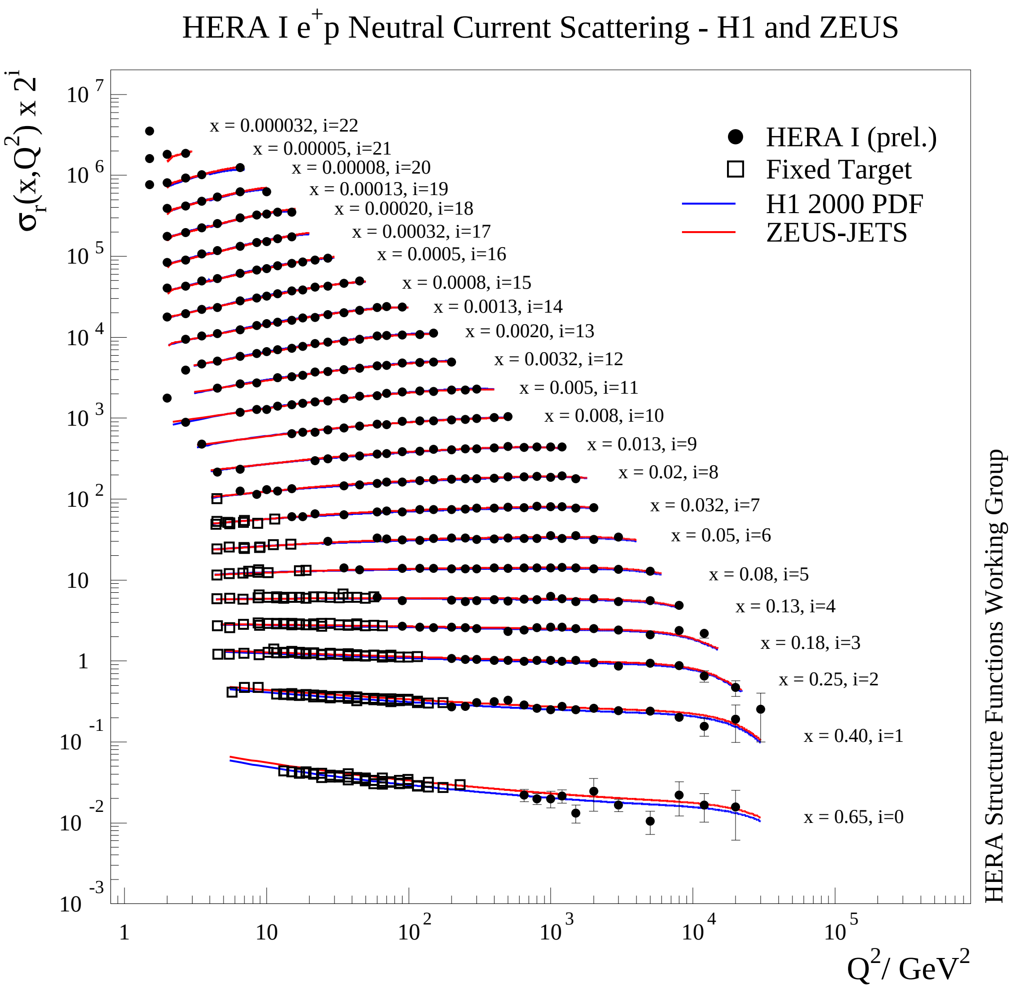
# Average of H1 and ZEUS for $Q^2 = 35\ \mathrm{GeV}^2$ (Eq. 5)



## Preliminary combination of H1 and ZEUS

- NC and CC, $e^{\pm}p$ data combined simultaneously. This allows to follow updates of correlated systematic uncertainties.
- $E_p = 820$ GeV data corrected to $E_p = 920$ GeV data. Correction may be restricted to $y < 0.35$ only.
- Data is interpolated to a common grid. Good agreement of interpolation factors for H1 and ZEUS parameterizations. In future, we will use the fit to the average for the interpolation.
- Global normalizations are treated as multiplicative. For other systematic uncertainties, both assumptions are considered, added as systematic uncertainty
- Correlation between various H1 and ZEUS systematic sources is considered. Largest effect from assumption on $\gamma p$ background and hadronic energy scale (common MC models). Added as extra uncertainties.

# Combined HERA data

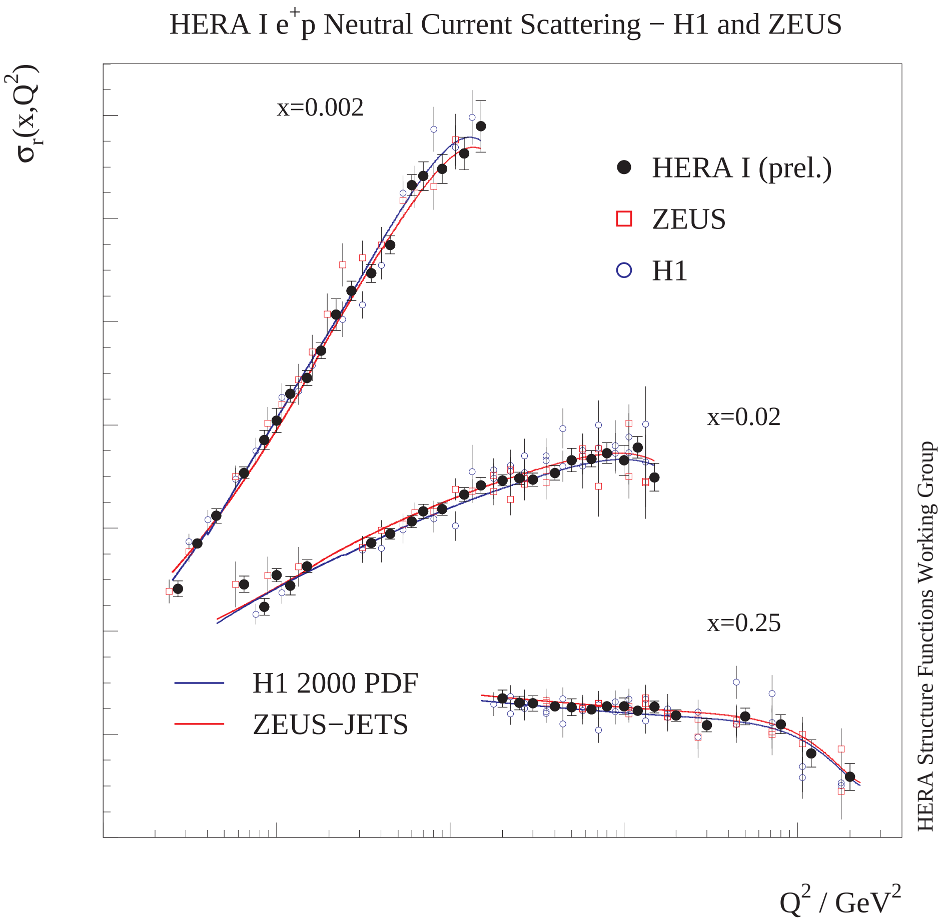



$\chi^2/dof = 510/599$

Experiments agree too well (over-consistency). That is related to conservative estimate of uncorrelated systematic uncertainties.

$\chi^2$ average data vs theory removes over-consistency of the data, provides more stringent test.

# Combination of HERA data




Shifts of some systematic sources:

| | | Shift | Error |
|---|---|---|---|
| 15 | hl lumi1 | 1.4373 | 0.5694 |
| 16 | hl lumi2 | 0.6875 | 0.6063 |
| 17 | h1 lumi3 | -1.0738 | 0.6260 |
| 18 | h1 lumi4 | -0.0797 | 0.7782 |
| 19 | h1 lumi5 | -0.5273 | 0.6088 |
| 1 | z lumi1 | 0.0461 | 0.5966 |
| 39 | z lumi2 | 0.0797 | 0.7782 |
| 43 | z lumi3 | -0.4064 | 0.3871 |
| ... | | | |
| 2 | h1_Ee_Spacal | 0.7473 | 0.3361 |
| .... | | | |
| 10 | h1_BG_Spacal | -0.4111 | 0.8168 |
| 28 | zd9_bg | -0.3002 | 0.4263 |

Experiments cross-calibrate each other. Uncertainties on systematic sources are strongly reduced for unique kinematic dependences, if the other experiments has better precision in a similar phase space.

# Summary

- Averaging provides a model independent tool to study consistency of the data.
- Over-consistency or inconsistency of the data is separated from the data to theory comparison.
- Hessian or Offset methods can be applied for error estimation after the data is combined.
- A systematic study of effects of correlations between different datasets.

Preliminary results based on published H1/ZEUS data are available since last summer. Next step is to publish the combined data tables...

→ Combined HERA data should allow for better estimation of PDF uncertainties, more strict constraint to the theory.