ZEUS NLOQCD fits to extract PDFs, $\alpha_s(M_Z^2)$ and their uncertainties

Amanda Cooper-Sarkar
Oxford University
Durham meeting March 18th 2002

TREATMENT OF CORRELATED SYSTEMATICS

Incorporating systematic errors into the predictions

$$f_i(p,s) = t_i(p) + \sum_k s_k \Delta_{ik}$$

$$\chi^2 = \sum_i \left[\frac{m_i - f_i(p,s)}{\sigma_i}\right]^2 + \sum_k s_k^2$$

OFFSET method: minimize assuming $s_k = 0$, i.e. FITTING DATA AS PUBLISHED

HESSIAN method 1: allow s as well as p to vary. Can be regarded as moving the data points within their systematic errors... letting the theory decide where the 'true' central value of a data point should be..letting the theory decide on the energy scale for the calorimeter for example

OFFSET method calculate

$$M_{ij} = \frac{1}{2}\frac{\partial^2\chi^2}{\partial p_i \partial p_j}, C_{ij} = \frac{1}{2}\frac{\partial^2\chi^2}{\partial p_i \partial s_j}$$

Statistical Covariance Matrix is given by $V = M^{-1}$

Correlated Systematic Errors Covariance Matrix is given by $V = M^{-1}CC^T M^{-1}$
(see Pascaud and Zomer LAL-95-05)

This is the method which gave the ZEUS result

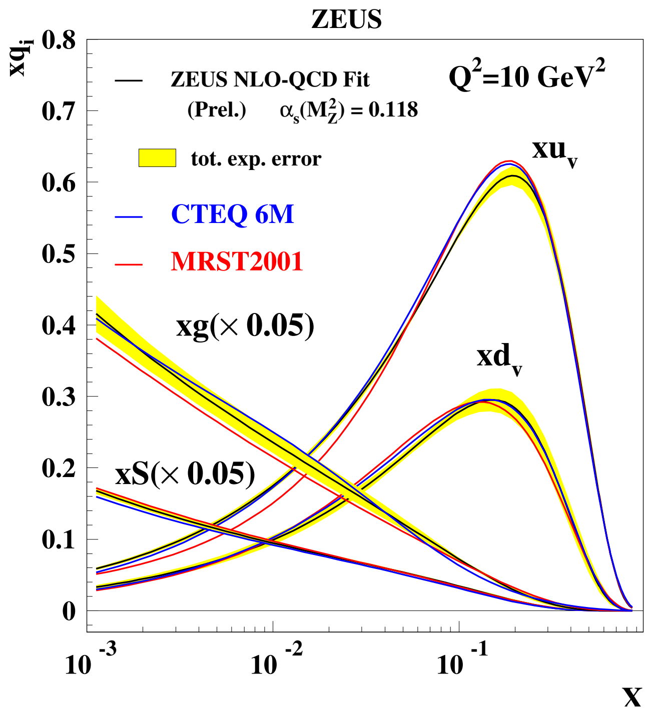
$$\alpha_s(M_Z^2) = 0.1172 \pm 0.0008 \pm 0.0054$$

to be updated to

$\alpha_s(M_Z^2) = 0.1166 \pm 0.0008 \pm 0.0032(sys) \pm 0.0036(norms)$

The method is equivalent to the time honoured procedure 'redo many fits with each s_k set to ± 1 and then add the deviations from the central value in quadrature'. Now called the offset method

It accounts for correlations of the systematic parameters s to the QCD parameters p but not for correlations of systematic parameters to each other. Of course experimentalists strive to make their different sources of systematic uncertainty independent!



The OFFSET method produced the error bands on these curves

HESSIAN method 1

Allow both QCD parameters p and systemtic parameters s to vary in the fit hence the covariance matrix has

$$M_{ij} = \frac{1}{2}\frac{\partial^2\chi^2}{\partial p_i \partial p_j}, C_{ij} = \frac{1}{2}\frac{\partial^2\chi^2}{\partial p_i \partial s_j}, S_{ij} = \frac{1}{2}\frac{\partial^2\chi^2}{\partial s_i \partial s_j}$$

and the errors come from the inverse of the whole matrix. Correlations between different systematic parameters can be introduced by the theory..including correlations between the systematic parameters of different data sets.

This is the method which gave the H1 result

$$\alpha_s(M_Z^2) = 0.1150 \pm 0.0017 \qquad (1)$$

ZEUS and H1 have very different fits apart from the error treatment e.g. ZEUS does a DIS global fit, H1 uses only BCDMS proton data for $y > 0.3$ in addition to their own data. So compare fitting H1 data instead of ZEUS data in the ZEUS global DIS fit, and compare fitting ZEUS data instead of H1 data in a fit done by the H1 procedure

Experiment	no. of data points	no. of systematic	no. of normalization	
	used		uncertainties	uncertainties
ZEUS96/97	242	10	2	
BCDMS p	305	5	5	
NMC p	218	12	4	
NMC D	218			
NMC D/p	129	5	0	
E665 D	47	7	2	
E665 p	47		1	
CCFRxF3	57	18	0	

In a global fit with may data sets we have many sources of systematic error, there are 71 for the ZEUS global fit. Thus fitting the systematic error parameters becomes impractical
CTEQ hep-ph 0101051 to the rescue CTEQ have been working on correct ways to incorporate correlated systematic errors

They suggest an analytic minimisation w.r.t. s parameters, which results in a new formulation of the χ^2

$$\chi^2 = \sum_i \left[\frac{m_i - t_i}{\sigma_i}\right]^2 - BA^{-1}B$$

$$B_k = \sum_i \Delta_{ik}(m_i - t_i)/\sigma_i^2$$

$$A_{kl} = \delta kl + \sum_i \Delta_{ik}\Delta_{il}/\sigma_i^2$$

This is equivalent to the HESSIAN method 1 in principle and avoids the need to minimize 71 extra parameters This will be called HESSIAN method 2

method	data points	χ^2(uncorrel)	-$BA^{-1}B$	χ^2(total)
OFFSET	242	303	-16	303
HESSIAN-1	242	354	-82	272
HESSIAN 2	242	361	-82	279
HESSIAN-1	242	306	-19	287
HESSIAN-2	242	306	-19	287

Table 1: Compare methods for ZEUS only fit

Test on a fit to ZEUS data alone: 242 data points from e^+p NC scattering with well understood systematics plus high Q^2 cross-sections from e^+p CC, e^-p NC and CC, for high x stability

The HESSIAN methods are closely equivalent
In the HESSIAN methods the uncorrelated part of the χ^2 increases w.r.t the OFFSET method and the systematic contribution compensates for this
The figures in red are the results if normalization uncertainties are not included amongst the systematics uncertainties

method	data points	χ^2(uncorrel)	$-BA^{-1}B$	χ^2(total)
OFFSET	242(ZEUS)	328	-27	328
	221(BCDMS)	223	-23	223
HESSIAN-1	242(ZEUS)	2514	-2255	259
	221(BCDMS)	4251	-4073	178
HESSIAN 2	242(ZEUS)	4748	-4478	270
	221(BCDMS)	9808	-9625	171
HESSIAN-1	242(ZEUS)	360	-59	301
	221(BCDMS)	245	-28	217
HESSIAN-2	242(ZEUS)	360	-59	301
	221(BCDMS)	245	-28	217

Table 2: Compare methods for ZEUS +BCDMS fit

Test on a fit of ZEUS plus BCDMS (comparable to H1 fits)

The HESSIAN methods are not completely equivalent when applied to two different data sets
the good χ^2 is made up from a large uncorrelated part of the χ^2 and a large systematic contribution which almost cancels it
However if we do not let normalization uncertainties be part of the systematics uncertainties things look a lot more reasonable.

Experiment	data points	χ^2(uncorrel)	-$BA^{-1}B$	χ^2(total)
ZEUS96/7	242	503	-245	258
BCDMS p	305	392	-72	320
NMC p+D	436	790	-168	622
E665 p+D	94	116	-2	114
CCFRxF3	57	283	-214	69
NMC n/p	129	127	-7	120

Table 3: HESSIAN method 2 not including normalization uncertainty

Now do the global fit by HESSIAN method 2.
Also consider the α_s values for these global fits

$$\alpha_s(M_Z^2) = 0.1166 \pm 0.0033$$

OFFSET method not including norm uncertainty

$$\alpha_s(M_Z^2) = 0.1120 \pm 0.0013$$

HESSIAN method-2 not including norm uncertainty

The α_s values from the HESSIAN method-2 are rather low.

SUMMARY so far

- HESSIAN method 2 gives a much reduced error estimate for α_s (and for PDF parameters)
- However the value of α_s is shifted to low values $\alpha_s \sim 0.112$ for the global fit. Other parameters are not affected as strongly.(The gluon is a little steeper at low x and a little softer at high x)
- Norm uncertainties should not be used in this method.they are not Gaussian
- To compare the fits done by the OFFSET and the HESSIAN method let us use a comparable χ^2 for each fit. Re-evaluate the χ^2 by adding statistical and systematic errors in quadrature for the final parameters as determined by each method.

  Comparing the re-evaluated χ^2 for each method favours the OFFSET method by $\Delta\chi^2 = 283$

Experiment	data points	χ^2/data point Hessian method 2	χ^2/data point offset method
ZEUS96/97	242	1.37	0.83
BCDMS p	305	0.95	0.89
NMC p	218	1.50	1.26
NMC D	218	1.15	0.96
E665 D	47	0.97	0.94
E665 p	47	1.17	1.16
CCFRxF3	57	0.99	0.39
NMC D/p	129	0.97	0.93

Table 4: Table of χ^2 calculated by adding systematic and statistical errors in quadrature for the theoretical parameters determined by the offset method and hessian method 2

Let us consider the errors which the CTEQ/HESSIAN method 2 produces. They are derived from the criteria for 'parameter estimation' $\chi^2 \to \chi^2 + 1$ BUT CTEQ themselves do not believe

that 1 is a reasonable tolerance on a global fit to ~ 1200 data points from diverse sources with unquantifiable theoretical and model uncertainty

The CTEQ papers (hep-ph 0101051 and 0101032 and CTEQ6 hep-ph 0201195 try to quantify a more reasonable tolerance T

$$\chi^2 \to \chi^2 + T^2$$

There are no strict criteria but the idea is that all data sets HAVE to be acceptable and compatible even if strict statistical criteria are not met
See Thorne's talk
The level of tolerance they suggest is

$$T \sim 10$$

To understand this it helps to consider the difference between the criteria for hypothesis testing and parameter estimation

Errors on parameters within a fixed theoretical framework are evaluated from $\chi^2 \rightarrow \chi^2(min) + 1$ (parameter estimation).

But the validity of a theoretical hypothesis is judged such that its χ^2 should be in the range $N \pm \surd(2N)$ where N is the number of degrees of freedom. Hypotheses are then compared such that an acceptable hypothesis is one with a χ^2 within $\chi^2(min) + \surd(2N)$

IN fitting DIS data for PDF parameters and α_s we are not in a clean situation of either parameter estimation or hypothesis testing. We want to stay within a fixed theoretical fraemwork say NLOpQCD

BUT we admit that many model assumptions like the form of the parametrization, value of cuts, value of Q_0^2 etc., could be varied. These variations represent different hyptheses and they are accepted provided they fall within the hypothesis testing criterion. The theoretical parameters obtained for

these different hypotheses can differ from those of the fit with the standard assumptions by more than their errors as evaluated by the parameter estimation criteria.

For example consider the change in the central value of α_s for a few different hypotheses

change of hypthesis	$\Delta\chi^2$	$\Delta\alpha_s$
values of cuts Q_0^2 etc	~ 10	0.002
heavy quark scheme	~ 50	0.001
higher twist terms	~ 50	-0.003

This is not a problem for the OFFSET method in which the uncorrelated errors evaluated by the parameter estimation criteria are augmented by the contribution of the systematic uncertainties. The shifts in parameter values are within the total error $\Delta\alpha_s = 0.0033$

However it is a problem for HESSIAN method 2 -such shifts now exceed the optimal error estimates of this method $\Delta\alpha_s = 0.0013$.

CTEQ propose using $\chi^2 \rightarrow \chi^2(min) + T^2$, for $T = 10$ to give a more reasonable estimate of the parameter error for HESSIAN method 2.

ZEUS have chosen to use

$$\chi^2 \rightarrow \chi^2 + \surd(2N)$$

i.e to use the hypothesis testing criterion.
In the present case since $\surd(\surd(2N)) \sim 7$ these

choices are not very different.

So what do we get if we changed the tolerance for parameter estimation in minuit from $T = 1$ to $T = 7$, using the HESSIAN method 2?

We get

$$\alpha_s(M_Z^2) = 0.1120 \pm 0.0033$$

The error is now the same size as that determined by the OFFSET method

And yes this does apply to all the PDF parameters as well!

And it is till true if we do allow normalization uncertainties into both methods- both methods then give errors ~ 0.005

We are trying to make conservative estimates of the error on the PDF shapes and the value of α_s, within a theoretical framework. The OFFSET method does this. It produces errors which correspond to raising the tolerance to the ball park which CTEQ and many others consider reasonable.

However is there any reason to prefer the OFFSET method to HESSIAN method 2, with tolerance $T \sim 7 - 10$?
Five things to consider: 2 theoretical, 2 experimental

- Look at Alekhin's plot. The OFFSET method is not efficient because it is inflenced more by the points with large systematic error than the point with small systematic error.
  EXPERIMENTALISTS DO YOU FEEL SLIGHTLY UNEASY?

- The Hessian method is stated to work when the systematic errors on the data are not extemely large compared to statistical. How large is large? Is the factor 3 we are achieving at ZEUS large? Consider Zarnecki's toy model comparing the methods as a function of the ratio $\sigma^2_{sys}/\sigma^2_{stat}$

- One cannot compare the χ^2 for these two methods, but one can recalcuate the χ^2 for the parameter values determined by each of these methods in the same way, by adding syatematic and statistical errors in quadrature. The OFFSET method is favoured by a $\Delta\chi^2$ of 283

- When the Hessian method is applied to two data sets which are measuring the same 'truth' (e.g. the H1 and ZEUS data sets measuring the reduced cross-section for ep scattering over the same region of the x, Q^2 plane) the systematic shifts which result from fitting the systematic error parameters pull the data sets apart!

- If the method is used with different input data sets e.g. ZEUS alone, ZEUS +BCDMS, ZEUS in a global fit (CTEQ6)

  the systematic shifts which it determines differ. Thus the 'true' position of the data depends not only on the theory but on which data sets are being fitted

ZEUS sys param	ZEUS+BCDMS	CTEQ6	ZEUS ONLY
1	1.31	1.67	-0.36
2	0.71	-0.67	1.17
3	1.32	-1.25	1.20
4	0.05	-0.44	0.40
5	-0.04	0.00	0.32
6	0.93	-1.07	0.39
7	1.45	1.28	-1.40
8	0.22	0.62	0.20
9	0.11	-0.40	0.04
10	-0.13	0.21	-0.06
11	-1.18	N.A.	0.5
12	-0.42	N.A.	0.7