Uncertainties of Parton Distribution Functions

Daniel R. Stump
Department of Physics and Astronomy, Michigan State University, East Lansing, MI, 48824

Abstract
Issues related to the analysis of uncertainties of parton distribution functions are discussed, focusing on the methods used in the CTEQ global analysis.

1. INTRODUCTION

High-energy particles interact through their quark and gluon constituents—the partons. By the asymptotic freedom of QCD, the parton cross sections may be approximated by perturbation theory. Then by the factorization theorem of QCD, the parton distribution functions of hadrons are the link between perturbative QCD and experimental measurements. This theory applies to a variety of high-energy processes, including deep-inelastic lepton scattering from nucleons (DIS), Drell-Yan production of $\mu\overline{\mu}$ pairs in nucleon-nucleon collisions (DY), and production of high-p_T jets at $p\bar{p}$ or pp colliders.

The parton distribution functions (PDFs) are important in high-energy physics. Any scattering experiment with nucleons in the initial state will require PDFs for the analysis and interpretation of the experiment. Currently, the HERA accelerator experiments measure ep and $\overline{e}p$ scattering; the Tevatron collider creates $p\overline{p}$ collisions. Tests of the standard model and the search for new physics rely on PDF phenomenology.

QCD global analysis has several goals: to construct an accurate set of PDFs; to know the uncertainties of the PDFs, which come from experimental measurement errors and from theoretical approximations; and to enable predictions, with realistic uncertainties, for future experiments.

The full, systematic study of PDF uncertainties developed slowly [1, 2]. The first practical parton distributions with full error bands were produced by Botje [3] using DIS data. Today many groups and individuals are involved in this active field of research. Both the CTEQ group [4] and the MRST group [5] have provided complete uncertainty analyses of their PDFs; the PDFs from these groups are widely used in high-energy physics. The HERA collaborations, ZEUS [6] and H1 [7], have constructed PDF models based on their own data, including error bands. Other individuals have made important contributions [8, 9]; and the Fermilab group [10] has developed a new methodology for PDFs by Monte Carlo sampling in the parton function space.

This paper will discuss several issues, focusing on the CTEQ methods and results. Section 2 describes the CTEQ input. Section 3 concerns the treatment of systematic errors. Section 4 discusses the compatibility of different data sets within a global fit to data. Section 5 explains the CTEQ uncertainty analysis.

2. GLOBAL ANALYSIS OF QCD

The aim of global analysis of short-distance processes using perturbative QCD is to construct a set of PDFs that yield good agreement with data from many disparate experiments. The

CTEQ6						
	process	data set	CorrMat	N	χ^2	χ^2/N
1	μ DIS	BCDMS F2p	Y	339	378	1.11
2	μ DIS	BCDMS F2d	Y	251	280	1.11
3	$\overline{e}$ DIS	H1 (a)	Y	104	98.6	0.95
4	e DIS	H1 (b)	Y	126	129	1.02
5	$\overline{e}$ DIS	ZEUS	Y	229	263	1.15
6	μ DIS	NMC F2p	Y	201	305	1.52
7	μ DIS	NMC d/p	Y	123	112	0.91
8	$p\overline{p} \to$ jet	D0	Y	90	69	0.77
9	$p\overline{p} \to$ jet	CDF	Y	33	49	1.47
10	$\nu(\overline{\nu})$ DIS	CCFR F2 + F3	Y/N	156	150	0.96
11	Drell-Yan	E605	N	119	95	0.80
12	Drell-Yan	E866 d/p	N	15	6	0.40
13	$p\overline{p} \to W$	CDF (Lasy)	N	11	10	0.91

Table 1: Experimental data sets used in the CTEQ6M global analysis [12]. CorrMat: availability of information on correlations of systematic errors. N is the number of data points. References for the experiments may be found in Ref. [12].

program of global analysis is not a routine statistical calculation, because of systematic errors—both experimental and theoretical. Therefore we must sometimes use physics judgement in producing the PDF model, as an aid to the objective fitting procedures.

A parton distribution function $f_i(x, Q)$ (where i labels the parton species) depends on two variables: the momentum fraction x carried by the parton and the momentum scale Q at which the nucleon is observed. Heuristically, $f_i(x, Q)$ is the density of parton species i per unit of momentum fraction. The PDFs are parametrized at a low momentum scale Q_0, of order 1 GeV, by a standard functional form with adjustable parameters $(a_0, a_1, a_2, \ldots)$

$$f(x, Q_0) = a_0 x^{a_1} (1 - x)^{a_2} P(x); \tag{1}$$

here $P(x)$ is a smooth function with a few additional free parameters. Separate functions exist for each parton species, so that the total number d of adjustable parameters is of order 20. The Q dependence of $f(x, Q)$ is determined by the QCD evolution equations [11], depending on the renormalization and factorization schemes for the perturbative calculation of parton cross sections. The most widely used CTEQ parton distributions are based on next-leading order (NLO) perturbation theory in the modified minimal subtraction renormalization scheme.

Table 1 lists the experiments used in the CTEQ6 global analysis—the most recent generation of CTEQ parton distributions [12]. Thirteen data sets are employed, and the number N of data points in a set ranges from scores to hundreds. The total number of data points is ~ 1800. These data come from major experiments of the past ten years. The CTEQ6 PDFs agree satisfactorily with all the experiments, from the fact that each χ^2/N is near 1 and from more detailed comparisons.

3. TREATMENT OF EXPERIMENTAL SYSTEMATIC ERRORS

What is a systematic error? Suppose two experimental groups measure the same quantity, but one has a positive systematic error and the other negative. The measurements will not agree within the statistical errors, and no theory could agree with both experiments. If the groups are aware of the possible errors then the measurements will in fact be consistent within the published uncertainties including systematics. But if the errors are not accurately characterized, an incompatibility will appear, recognized as a systematic difference of the results.

The situation described in the previous paragraph is analogous to what often happens in global analysis of PDFs. But in the case of PDFs, the systematic differences are only visible through the process of global analysis. The disparate experiments may measure different energy domains, or altogether different scattering processes; the results cannot be compared directly. Only through the combined fitting of PDFs are the systematic differences revealed.

For a global analysis, a crucial feature of the systematic errors is that they are highly correlated. Therefore we construct the PDFs using a procedure of χ^2 minimization with fitting of systematic errors. First consider an experiment with statistical errors alone; we would define

$$\chi^2 = \sum_{i=1}^{N} \frac{(D_i - T_i)^2}{\sigma_i^2} \qquad \left\{ \begin{array}{l} D_i : \text{data value}(i = 1 \ldots N) \\ T_i : \text{theoretical value} \\ \sigma_i : \text{statistical error (S.D. of } D_i) \end{array} \right. \tag{2}$$

The theory value T_i is a function of d adjustable parameters $\{a_1, \ldots, a_d\} \equiv \mathbf{a}$. Then minimzation of χ^2 with respect to the $\{a_\mu\}$ would give the optimal PDF model $\{a_{0\mu}\}$. This procedure would be ideal if the data value D_i is $T_{0i} + \sigma_i r_i$ where r_i is a random Gaussian fluctuation with $\langle r_i^2 \rangle = 1$.

Normalization error. In any scattering experiment there is an overall normalization uncertainty from the luminosity. To take this possible error into account, we introduce a normalization factor f_{N} and define a new χ^2 function,

$$\chi^2(\mathbf{a}, f_{\mathrm{N}}) = \left(\frac{1 - f_{\mathrm{N}}}{\sigma_{\mathrm{N}}} \right)^2 + \sum_{i=1}^{N} \frac{(f_{\mathrm{N}} D_i - T_i)^2}{\sigma_i^2}, \tag{3}$$

where σ_{N} is the published normalization uncertainty. Minimization of χ^2 with respect to both the model parameters $\{a_\mu\}$ and the normalization factor f_{N} yields an optimized model, within the normalization uncertainty. Rather than fitting T_i to D_i, we fit it to $f_{\mathrm{N}} D_i$. The choice of f_{N} is also optimized by the minimization of χ^2. The penalty term in (3) ensures that f_{N} will not deviate too much from 1.

An overall normalization error of theoretical origin, such as a K-factor from high-order QCD corrections, would be indistinguishable from a luminosity error. Thus the correction for systematic errors may involve both experimental and theoretical errors, and these would be entwined.

General systematic errors. High-precision experiments publish *many* systematic errors. In general, the difference between data and theory will be

$$D_i = T_{0i} + \alpha_i r_i + \sum_{j=1}^{K} \beta_{ij} \widehat{r}_j, \tag{4}$$

where α_i is the uncorrelated error of D_i and $\{\beta_{ij}, j = 1 \ldots K\}$ is a set of K systematic (and 100% correlated) errors. The $\{r_i\}$ and $\{\hat{r}_j\}$ denote random fluctuations. To account for the systematic errors, we introduce a systematic shift s_j for each source of error and define a new χ^2 function,

$$\chi'^2(\mathbf{a}, \mathbf{s}) = \sum_{i=1}^{N} \frac{\left(D_i - \sum_j \beta_{ij} s_j - T_i\right)^2}{\alpha_i^2} + \sum_{j=1}^{K} s_j^2. \quad (5)$$

Minimization of χ'^2 with respect to both the functional parameters $\{a_\mu\}$ and the systematic shifts $\{s_j\}$ determines the optimal model. The theory is not fit to the central values of the published data, but to values that are shifted by amounts consistent with the published systematic errors ($\beta_{ij} s_j$). The fitting procedure produces a compatible model within the systematic errors.

Because χ'^2 depends quadratically on the $\{s_j\}$ we can solve for the optimized shifts analytically, $\mathbf{s} \rightarrow \mathbf{s}_0(\mathbf{a})$. Thus the systematic shifts are continually optimized as we vary the functional parameters $\{a_\mu\}$ in seeking the optimal PDFs.

The procedure outlined above accounts for the statistical errors (by weighting with α_i^{-2}), the overall normalization uncertainty (by numerical fitting of $f_{\rm N}$), and the other systematic errors (analytically).

Finally, the *global* χ^2 *function* is the sum of the $\chi_e'^2$ over all the experiments e. We might also apply weighting factors $\{w_e\}$ when combining the experiments, with default values $w_e = 1$. The spirit of global analysis is compromise—the PDF model should fit all data sets satisfactorily. If a model has poor agreement with some data, we may construct another model with better agreement by giving that data an enhanced weight in the global χ^2. The second fit may be judged to be preferable as a standard fit to the global data, even though it is not the best fit to other experiments. In making a subjective decision like this, physics judgement enters into the calculation.

The quality of the final CTEQ6 PDFs can to some extent be gauged by the χ^2 values in Table 1. But just looking at a single number for a data set does not do justice to the theory. More detailed comparisons, e.g., by plotting data and theory superimposed or by the "pull" distributions, reveal the beautiful success of QCD. [12]

4. A STUDY OF COMPATIBILITY

Because we seek to construct PDFs that agree with many disparate experiments, so that the functions contain the best available information on all aspects of the parton structure simultaneously, an important issue is whether the data from different experiments are in agreement with each other. If two data sets do not agree with one another, then no theoretical model can agree with both. The philosophy of the CTEQ program is to use high-precision data from all relevant experiments. We expect to observe minor incompatibilities between experiments because of systematic errors. Then practical PDF models will require compromises among the different experiments. However, we should not encounter true inconsistencies, i.e., which cannot be reconciled by a reasonable model.

Several methods have been used to judge the compatibility of different data sets. One method is to study alternate fits that result from changing the weights of the data sets in the global χ^2 function. One example is shown in Table 2. The data sets and PDFs in this study are not identical to CTEQ6 but are quite similar. The PDF model A is the standard set—for which all data sets have the default weight $w_e = 1$ in the fitting. The PDF model B is an alternate in which the three H1 data sets and two BCDMS data sets are given a large extra weight.[1] The final column $\Delta\chi^2$ gives the difference in χ^2 for each experiment, between models B and A. By heavily weighting H1 and BCDMS, their χ^2's have decreased significantly, by -38.8 units (out of ~ 970). But on the other hand the χ^2's of other experiments have increased significantly, by $+149.7$ (out of ~ 1390). For example, while the agreement with H1 data improved, the quality of the fit to ZEUS data got worse. The result is that the model in best agreement with H1 and BCDMS data will not agree satisfactorily with other experiments. The change in χ^2 can be rather large. There is a minor incompatibility.

Giving extra weight to one or a few experiments is equivalent to using that data alone to construct the PDFs. An experimental collaboration that uses only its own data can expect to model that data better than a general set of PDFs from a global analysis. However, their resulting PDFs will not describe the many other data as well.

The inverse of giving extra weight to an experiment is to give it less weight, of which the extreme case is to give it weight 0, i.e., simply to remove it from the global analysis. When such exercises are performed, we find for most data sets that the χ^2's of other experiments may decrease significantly, by amounts of order 10 units in some cases, while the χ^2 for the dropped data increases by a similar amount. Again, the implication is that the best fit to the remaining data may not give a satisfactory fit to the dropped data.

These minor incompatibilities suggest that there are systematic errors, whether of experimental or theoretical origin, that prevent a purely statistical analysis of the uncertainties. Furthermore, the results show that reasonable PDFs may differ in global χ^2 by amounts much larger than 1.

Other clever ways to test the compatibility of different data sets have been devised. One method is to plot χ^2 of the global data versus χ^2 of an individual data set, for alternate PDF models generated by the Lagrange Multiplier method [13]. Another approach is to apply the Bootstrap Method, i.e., to generate alternative PDF models from resampled data points. In the latter method, it has been found that to obtain significant changes in the PDFs, one must apply the resampling procedure to entire data sets rather than just to individual uncorrelated data points.

5. UNCERTAINTY ANALYSIS

As part of the CTEQ global analysis, a group at Michigan State University has developed several computational methods for analyzing fully the uncertainties of parton distributions [4]. In these calculations, we continue to use $\chi^2_{\rm global}$ as a figure of merit for the quality of model PDFs. The methods explore $\chi^2_{\rm global}$ in the neighborhood of its minimum, as illustrated in Fig. 1(a). The point represents the standard fit, denoted S_0, which corresponds to the position in parameter space where χ^2 is minimum; and the ellipse indicates nearby points that are also deemed to be acceptable fits to the global data set. The computational problem is to explore the full d-dimensional neighborhood of S_0. When this problem has been solved, we can assess the implications for PDF uncertainties.

[1] In generating the alternate PDFs, the overall normalization factors $f_{{\rm N},e}$ were kept fixed at their values for the standard set. All other systematic shifts were allowed to readjust to the new PDFs.

	data set	N	$\chi^2[A]$	$\chi^2[B]$	$\Delta\chi^2$
1	BCDMS p	339	366.1	362.7	-3.4
2	BCDMS d	251	273.6	258.5	-15.1
3	H1 (a)	104	97.8	85.4	-12.4
4	H1 (b)	126	127.3	123.0	-4.3
5	H1 (c)	129	108.9	105.3	-3.6
6	ZEUS	229	261.1	288.5	$+27.5$
7	CDHSW F2	85	65.6	84.8	$+19.2$
8	NMC p	201	295.5	303.5	$+8.0$
9	NMC d/p	123	115.4	111.6	-3.8
10	CCFR F2	69	84.9	139.4	$+54.5$
11	E605	119	94.7	96.7	$+2.0$
12	E866 pp	184	239.2	242.9	$+3.7$
13	E866 d/p	15	5.00	5.6	$+0.6$
14	D0 jet	90	62.6	84.6	$+22.0$
15	CDF jet	33	56.1	55.1	-1.0
16	CDHSW F3	96	76.4	87.5	$+11.0$
17	CCFR F3	87	27.0	32.7	$+5.9$
18	CDF W	11	8.7	8.9	$+0.2$

Table 2: Experimental data sets used in compatibility studies. The data sets and PDFs are not identical to CTEQ6, but quite similar. PDF set A: the standard set. PDF set B: PDFs obtained by giving large extra weight to the H1 and BCDMS data sets in the fitting process.

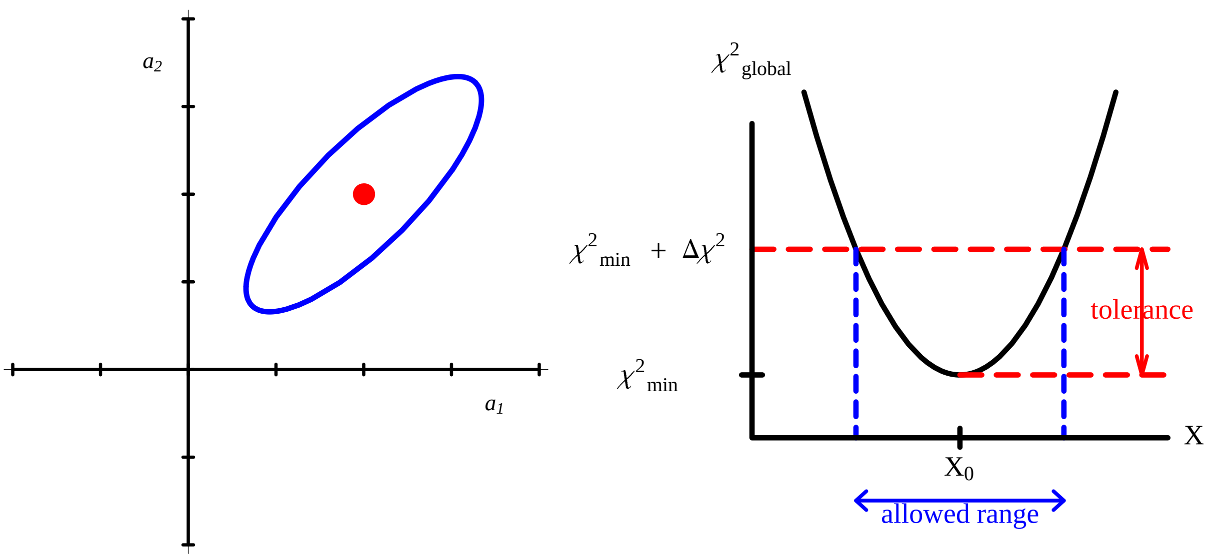

Fig. 1: (a) The neighborhood of the minimum of χ^2 in parameter space. (b) Tolerance and allowed range.

The Hessian Method. The Hessian is the matrix of second derivatives of χ^2,

$$H_{\mu\nu} = \frac{1}{2}\frac{\partial^2\chi^2}{\partial a_\mu \partial a_\nu}\bigg|_0 \tag{6}$$

where μ and ν range from 1 to d. The classical error formula for the deviation of an observable $X(\mathbf{a})$ from its predicted value is

$$(\Delta X)^2 = \left(\Delta\chi^2\right)\sum_{\mu,\nu}\frac{\partial X}{\partial a_\mu}\left(H^{-1}\right)_{\mu\nu}\frac{\partial X}{\partial a_\nu} \tag{7}$$

where $\Delta\chi^2$ is any specified increase in the value of χ^2 from the minimum. We obtain better numerical convergence by using the eigenvectors of the Hessian as the basis for the parameter space. To compute the variations of $X(\mathbf{a})$ along the eigenvector directions, we generate, for each eigenvector κ, a pair of points—along the $+$ and $-$ directions of the eigenvector—denoted $S_\kappa^{(\pm)}$. These are displaced from the standard point S_0 by a distance $T = \sqrt{\Delta\chi^2}$. Then the finite-difference analog of (7) is

$$(\Delta X)^2 = \frac{1}{4}\sum_{\kappa=1}^{d}\left[X\left(S_\kappa^{(+)}\right) - X\left(S_\kappa^{(-)}\right)\right]^2 \tag{8}$$

which we call the "master formula" for calculating PDF uncertainties. The eigenvector basis sets $\{S_\kappa^{(\pm)}\}$ have been published for general use by high-energy physics.

The Lagrange Multiplier method. The second method is used to analyze the uncertainties of PDF-dependent predictions, in a way that does not require the linearized error analysis. The idea is to perform a constrained fit, by minimization of the function $\chi^2(\mathbf{a})+\lambda X(\mathbf{a})$ with respect to variations of the functional parameters $\{a_\mu\}$, for a fixed value of the Lagrange multiplier λ. The result is the best fit to data for which the observable X has the value given by that fit. As we vary the Lagrange multiplier λ, we thus trace out the curve of $\chi^2_{\rm constrained}$ versus X. The corresponding points in parameter space are treated as alternative models for the parton structure.

5.1 The question of tolerance

How should the uncertainty of PDFs be assessed? One simple, and too naive, idea is to use the increase of $\chi^2_{\rm global}$ to define the uncertainty. Let $X(a)$ be some quantity of interest that depends on the PDFs. We have, by either the Hessian or Lagrange multiplier method, the curve of $\chi^2_{\rm global}$ versus X, as illustrated in Fig. 1(b), for constrained fits over a range of values of X. The predicted value of X is X_0, the value for the standard set S_0, for which χ^2 is minimum. Then for any tolerated increase in χ^2, denoted "tolerance" $\Delta\chi^2$ in Fig. 1(b), there is a corresponding allowed range for the quantity X. So, if we choose a canonical increase in χ^2 for the tolerance, the error on X is directly computed as $\pm\Delta X$.

For normal, i.e., Gaussian, errors, the tolerance would be $\Delta\chi^2 = 1$. However, it is well-known from experience that the changes in PDFs for which the global χ^2 increases by 1 unit are quite trivial. For example, we have seen in Section 4 that omitting or including a particular data set in the global analysis can easily induce changes in the global χ^2 much larger than 1. Nevertheless it is important to understand that the $\Delta\chi^2 = 1$ criterion is appropriate for normal statistics, even with systematic errors. Consider an experiment to measure an observable θ, with N measurements $\{\theta_i\}$ that are assumed to have statistical errors $\{\sigma_i\}$ and a common systematic error β; then we write $\theta_i = T + \sigma_i r_i + \beta\widetilde{r}$ where T is the true value and r_i and $\widetilde{r}$ are random fluctuations. The combined measurement is determined by minimization of $\chi'^2(\theta, s)$ in the manner described earlier,

$$\chi'^2(\theta, s) = \sum_{i=1}^{N} \frac{(\theta_i - \beta s - \theta)^2}{\sigma_i^2} + s^2 \qquad \begin{cases} \theta : \text{observable} \\ s : \text{systematic shift} \end{cases} \tag{9}$$

The combined measurement θ_c and its standard deviation $\Delta\theta_c$ are then

$$\theta_c = \frac{\sum_i \theta_i/\sigma_i^2}{\sum_i 1/\sigma_i^2} \qquad \text{and} \qquad (\Delta\theta_c)^2 = \frac{1}{\sum_i 1/\sigma_i^2} + \beta^2. \tag{10}$$

It can be shown that the increase in χ'^2 for one SD of θ_c is

$$\chi'^2\left[\theta_c \pm \Delta\theta_c, s_0(\theta_c \pm \Delta\theta_c\right] - \chi'^2\left[\theta_c, s_0(\theta_c)\right] = 1; \tag{11}$$

here $s_0(\theta)$ is the optimal systematic shift as a function of θ.

The criterion $\Delta\chi^2 = 1$ depends on the assumption that the statistical and systematic errors are known. Suppose, however, that the true statistical and systematic errors are rather $\widehat{\sigma}$ and $\widehat{\beta}$ (independent of i). Then the SD of θ_c becomes $(\Delta\theta_c)^2 = \widehat{\sigma}^2/N + \widehat{\beta}^2$; and the increase of χ'^2 is $(\widehat{\sigma}^2 + N\widehat{\beta}^2)/(\sigma^2 + N\beta^2)$. For example, if the systematic error were *omitted* from the fitting procedure of χ'^2 (i.e., $\beta = 0$) then the increase in χ'^2 would be large for large N.

Rather than base the uncertainty estimate on one single number—the global χ^2 value, which we already suspect has non-ideal behavior from the minor systematic differences or incompatibilities between different data sets—we go back to inspect the χ^2's of the individual data sets. By the Hessian or Lagrange Multiplier method we generate a series of alternate fits, which differ by the value of the observable X of interest. Then each data set defines a "prediction" (from the fit with lowest χ^2) and a "range" (from a specified increase of its χ^2). The estimated uncertainty of X is the intersection of the allowed ranges from the separate experiments. This procedure, while it does not produce a precise confidence level, does give a reasonable estimate of the uncertainty of the prediction attributable to PDFs, in the spirit of a 90% confidence level. So long as no data set can rule out a value of X, a reasonable set of PDFs will exist that describes the global data satisfactorily with that value of X.

Figure 2 illustrates our uncertainty estimates for two cases. Figure 2(a) shows the 90% confidence ranges, for the separate data sets, of the cross section σ_W for production of a W boson at the Tevatron [4]; the dashed lines are our estimated uncertainty range on σ_W. Figure 2(b) shows the $\Delta\chi^2 = 1$ ranges for the values of the strong coupling $\alpha_S(M_Z)$. The shaded band is the PDG world-averaged 1-sigma range. The vertical lines are our estimated uncertainty range from global analysis. The global fitting is nicely consistent with the PDG value—another success of the standard model—but we find that the uncertainties in global analysis are larger than those of the experimental techniques used in the PDG determination of $\alpha_S(M_Z)$.

Returning to the question of the increase of the global χ^2, we find that extreme alternate PDF sets, i.e., that would be judged unacceptable by our studies, tend to differ in χ^2 from the minimum value by an amount of order 100 (for $N_{\text{tot}} \sim 1800$ for CTEQ6). Therefore we have adopted $\Delta\chi^2 = 100$ as a typical tolerance for constructing the eigenvector basis sets $\{S_\kappa^{(\pm)}\}$

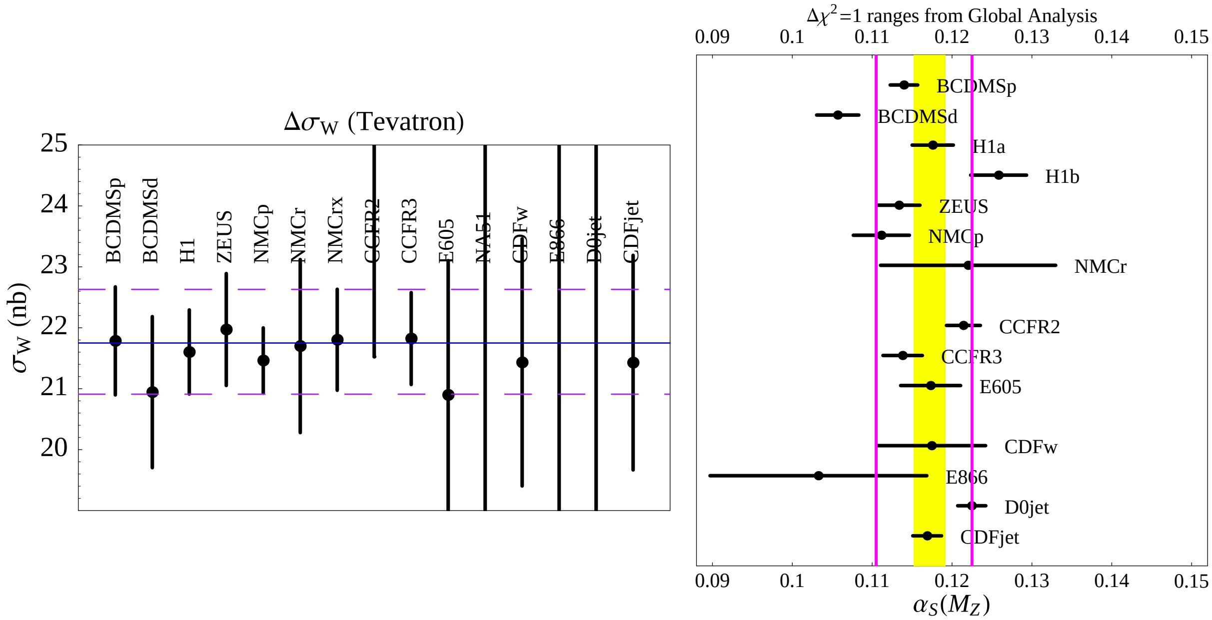

Fig. 2: (a) Allowed ranges for σ_W from separate data sets in a global analysis; these are 90% confidence ranges. (b) Allowed ranges for $\alpha_S(m_Z)$ from separate data sets; these are $\Delta\chi^2 = 1$ ranges.

published for general use in PDF applications. Obviously this increase in χ^2 is very large compared to the normal 1.[2]

All PDF research groups must confront the issue of "tolerance," and different groups have made different judgements. The CTEQ and MRST groups, which employ data from many processes, examine the agreement with separate data sets to estimate the uncertainty. Other studies, based on fewer processes, apply the Gaussian criterion $\Delta\chi^2 = 1$ [9, 7]. Further understanding of this issue will emerge as the groups compare their physical predictions.

6. CONCLUSION

Global analysis of QCD requires multivariate fitting to varied data with both statistical and systematic errors. Naturally any such problem is intricate. A routine statistical analysis will not solve the problem.

This review of CTEQ methods has emphasized the experimental errors—how measurement errors propagate to predictions that depend on PDFs. Other sources of PDF uncertainty exist, from theoretical assumptions. The MRST group has recently published an extensive study of theoretical uncertainties (the second paper in [5]). The paper identifies and analyzes four categories of theoretical uncertainty, and concludes that the theoretical uncertainty is as large as, or perhaps larger than, the experimental uncertainty. In any case theoretical and experimental errors are entwined by the fitting procedure; a theory error may be compensated by correlated data shifts within the published range of systematic errors.

Because PDFs are important for the interpretation of current and future collider experiments [14], improved theoretical calculations and methods of uncertainty analysis will no doubt

[2]When varying the PDFs to generate alternate fits, we keep the normalization factors $f_{N,e}$ fixed. The change of χ^2 would be smaller if the $\{f_{N,e}\}$ were allowed to readjust. For normal errors the optimized systematic shifts must be allowed to float when the criterion is $\Delta\chi^2 = 1$.

be developed, especially as data with ever higher precision become available.

ACKNOWLEDGEMENTS

I would like to thank Louis Lyons and the other conference organizers; also my MSU colleagues Jon Pumplin, Wu-Ki Tung and Joey Huston; and Sergei Alekhin for useful conversations about PDF uncertainties during a visit to MSU.

References

[1] J. C. Collins and D. E. Soper, arXiv:hep-ph/9411214 (1994).

[2] C. Pascaud and F. Zomer, Technical Note LAL-95-05 (1995).

[3] M. Botje, Eur. Phys. J. **C14**, 285 (2000) [hep-ph/9912439].

[4] J. Pumplin, D. R. Stump and W. K. Tung, Phys. Rev. D **65**, 014011 (2002) [hep-ph/0008191]; D. R. Stump *et al*, Phys. Rev. D **65**, 014012 (2002) [hep-ph/0101051]; J. Pumplin *et al*, Phys. Rev. D **65**, 014013 (2002) [hep-ph/0101032].

[5] A. D. Martin, R. G. Roberts, W. J. Stirling and R. S. Thorne, Eur.Phys.J. C **28**, 455-473 (2003) [hep-ph/0211080]; and arXiv:hep-ph/0308087 (2003).

[6] S. Chekanov *et al* [ZEUS Collaboration], Phys. Rev. D **67**, 012007 (2003).

[7] C. Adloff *et al* [H1 Collaboration], arXiv:hep-ex/0304003 (2003).

[8] V. Barone, C. Pascaud, and F. Zomer, Eur. Phys. J. **C12**, 243 (2000) [hep-ph/9907512].

[9] S. I. Alekhin, Eur. Phys. J. **C10**, 395 (1999) [hep-ph/9611213]; Phys. Rev. D **63**, 094022 (2001) [hep-ph/0011002].

[10] W. T. Giele and S. Keller, Phys. Rev. **D58**, 094023 (1998) [hep-ph/9803393]; W. T. Giele, S. Keller and D. Kosower, arXiv:hep-ph/0104052 (2001).

[11] V.N. Gribov and L.N. Lipatov, Sov. J. Nucl. Phys. **15**, 438 (1972); L.N. Lipatov, Sov. J. Nucl. Phys. **20**, 94 (1975); G. Altarelli and G. Parisi, Nucl. Phys. **B126**, 298 (1977); Yu.L. Dokshitzer, Sov. Phys. JETP **46**, 641 (1977).

[12] J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, P. Nadolsky and W. K. Tung, JHEP **0207**, 012 (2002) [hep-ph/0201195].

[13] J. C. Collins and J. Pumplin, arXiv:hep-ph/0105207 (2001).

[14] D. R. Stump, J. Huston, J. Pumplin, W. K. Tung, H. L. Lai, S. Kuhlmann and J. F. Owens, arXiv:hep-ph/0303013 (submitted to JHEP).