

Uncertainties of Parton Distribution Functions

R. Brock, D. Casey, J. Huston, J. Kalk, J. Pumplin, D. Stump, W.K. Tung

Department of Physics and Astronomy

Michigan State University

East Lansing, MI 48824

(May 18, 2000)

We describe preliminary results from an effort to quantify the uncertainties in parton distribution functions and the resulting uncertainties in predicted physical quantities. The production cross section of the W boson is given as a first example. Constraints due to the full data sets of the CTEQ global analysis are used in this study. Two complementary approaches, based on the Hessian and the Lagrange multiplier method respectively, are outlined. We discuss issues on obtaining meaningful uncertainty estimates that include the effect of correlated experimental systematic uncertainties and illustrate them with detailed calculations using one set of precision DIS data.

PACS number(s):

I. INTRODUCTION

Many measurements at the Tevatron rely on parton distribution functions (PDFs) for significant portions of their data analysis as well as the interpretation of their results. For example, in cross section measurements the acceptance calculation often relies on Monte Carlo (MC) estimates of the fraction of unobserved events. As another example, the measurement of the mass of the W boson depends on PDFs via the modeling of the production of the vector boson in MC. In such cases, uncertainties in the PDFs contribute, by necessity, to uncertainties on the measured quantities. Critical comparisons between experimental data and the underlying theory are often even more dependent upon the uncertainties in PDFs. The uncertainties on the production cross sections for W and Z bosons, currently limited by the uncertainty on the measured luminosity, are approximately 4%. At this precision, any comparison with the theoretical prediction inevitably raises the question: How “certain” is the prediction itself?

A recent example of the importance of PDF uncertainty is the interpretation of the measurement of the high- E_T jet cross section at the Tevatron. When the first CDF measurement was published [1], there was a great deal of controversy over whether the observed excess, compared to theory, could be explained by deviations of the PDFs, especially the gluon distribution, from the conventionally assumed behavior; or was it a first signal for new physics [2].

With the unprecedented precision and reach of many of the Run I measurements, understanding the implications of uncertainties in the PDFs has become a burning issue. During Run II (and later at LHC) this issue may strongly affect the uncertainty estimates in precision Standard Model studies, such as the all important W -mass measurement, as well as the signal and background estimates in searches for Higgs and other new physics.

In principle, it is the uncertainties on physical quantities due to parton distributions, rather than on the PDFs

themselves, that is of concern. The PDFs are theoretical constructs that depend on the renormalization and factorization schemes; and there are strong correlations between PDFs of different flavors and from different values of x , which can compensate each other in the convolution integrals that relate them to physical cross sections. On the other hand, since PDFs are universal, if we can obtain meaningful estimates of their uncertainties based on analysis of existing data, then the results can be applied to all processes that are of interest in the future [3,4].

[INTENT OF THIS PARAGRAPH UNCLEAR: IT SEEMS TO DESCRIBE WHAT WE ACTUALLY DO, SINCE GIVEN THE ERROR MATRIX, OR THE EIGEN DIRECTIONS AND EIGENVALUES, THE UNCERTAINTY CAN BE DERIVED FOR AN ARBITRARY QUANTITY, IF THE QUADRATIC APPROXIMATION IS ADEQUATE. PERHAPS THE INTENT WAS TO INTRODUCE THE LAGRANGE MULTIPLIER IDEA HERE??] One can attempt to assess directly the uncertainty on a specific physical prediction due to the full range of PDFs allowed by available experimental constraints. This approach will provide a more reliable estimate for the range of possible predictions for the physical variable under study, and may be the best course of action for ultra-precise measurements such as the mass of the W boson or the W production cross section. However, such results are process-specific and therefore the analysis must be carried out for each case individually.

Until recently, the attempts to quantify either the uncertainties on the PDFs themselves (via uncertainties on their functional parameters, for instance) or the uncertainty on derived quantities due to variations in the PDFs have been rather unsatisfactory. Two commonly used methods are: (1) Comparing the predictions obtained with different PDF sets, *e.g.*, various CTEQ [5], MRS [6] and GRV [7] sets; (2) Within a given global analysis effort, varying individual functional parameters *ad hoc*, within limits considered to be consistent with the existing data, *e.g.* [8]. Neither method provides a systematic

and reliable, quantitative measure of the uncertainties of the PDFs or their predictions.

As a case in point, Fig. 1 shows how the calculated value of the cross section for W boson production at the Tevatron varies with a set of historical CTEQ PDFs as well as the most recent CTEQ [5] and MRST [6] sets. Also shown are the most recent direct measurements from DØ and CDF*.

While it is comforting to see that the predictions have remained within a narrow range, the variation observed cannot be characterized as a meaningful estimate of the uncertainty because (i) the variation with time mostly reflects changes in experimental input to the global analysis, and to gradual refinements of the analysis procedure; and (ii) the perfect agreement between the values of the most recent CTEQ5M1[†] and MRS99 sets must be fortuitous, since each group has also obtained other satisfactory sets that allow much larger variations of the W cross section. The MRST group, in particular has examined the range of this variation by setting a variety of parameters to some extreme values [8]. These studies are useful but can not be considered quantitative or definitive. What is needed are methods that explore thoroughly the possible variations of the parton distribution functions.

It is important to recognize all potential **sources of uncertainty** in the determination of PDFs. Focusing on some of these, while neglecting significant others, may not yield practically useful results. Sources of uncertainty are listed below:

1. **Statistical uncertainties** of the experimental data used to determine the PDFs. These vary over a wide range among the experiments used in a global analysis, but are straightforward to treat.
2. **Systematic uncertainties** within each data set. There are typically many sources of experimental systematic uncertainty, some of which are highly correlated. These uncertainties can be treated by standard methods of probability theory *provided* they are precisely known, which unfortunately is often not the case – either because they may not be randomly distributed and/or because their estimation in practice involves subjective judgements.

*Much of the difference between the DØ and CDF W cross sections is due to the different values of the total $p\bar{p}$ cross sections used

[†]CTEQ5M1 is an updated version of CTEQ5M differing only in a slight improvement in the QCD evolution (cf. note added in proof of [5]). The differences are completely insignificant for our purposes. Henceforth, we shall refer to them generically as CTEQ5M. Both sets can be obtained from the web address <http://cteq.org/>.

3. **Theoretical uncertainties** arising from higher-order PQCD corrections, resummation corrections near the boundaries of phase space, power-law (higher twist) and nuclear target corrections, etc.
4. Uncertainties due to the **parametrization of the non-perturbative PDFs**, $f_a(x, Q_0)$, at some low momentum scale Q_0 . The specific choice of the functional form used at Q_0 introduces implicit correlations between the various x -ranges, which could be as important, if not more so, than the experimental correlations in the determination of $f_a(x, Q)$ for all Q .

Since strict quantitative statistical methods are based on idealized assumptions such as random measurement uncertainties, an important trade-off must be faced in devising a **strategy** for the analysis of PDF uncertainties. If emphasis is put on the “rigor” of the statistical method, then many important experiments cannot be included in the analysis, either because the published errors appear to fail strict statistical tests or because data from different experiments appear to be mutually exclusive in the parton distribution parameter space [4]. If priority is placed on using the maximal experimental constraints from available data, then standard statistical methods may not apply, but must be supplemented by physical considerations, taking into account experimental and theoretical limitations. We choose the latter approach, pursuing the determination of the uncertainties in the context of the current CTEQ global analysis. In particular, we include the full body of the world’s data used in the CTEQ5 analysis in this uncertainty study, and adopt the CTEQ5M1 set as the “best fit” around which the uncertainty studies are performed. In practice, there are unavoidable choices and compromises that must be made in the analysis, which are similar to the subjective judgements often necessary in estimating systematic errors in experimental analyses. The most important consideration is that quantitative results must remain robust with respect to reasonable variations in these choices.

In this Report we describe preliminary results obtained by our group using the two approaches mentioned earlier. [AS OF NOW MENTIONED ONLY IN THE ABSTRACT!] In Section 3 we focus on the error matrix, which characterizes the general uncertainties of the non-perturbative PDF parameters. In Sections 4 and 5 we study specifically the production cross section σ_W for W^\pm bosons at the Tevatron, to estimate the uncertainty of the prediction of σ_W due to PDF uncertainty. We start in Section 2 with a review of some aspects of the CTEQ global analysis on which this study is based.

II. ELEMENTS OF THE BASE GLOBAL ANALYSIS

Since our strategy is based on using the existing framework of the CTEQ global analysis, it is useful to review

some of its features pertinent to the current study [5].

a. Data selection: Table I shows the experimental data sets included in the CTEQ5 global analysis, and in the current study. For neutral current DIS data only the most accurate proton and deuteron target measurements are kept, since they are the “cleanest” and they are already extremely extensive. For charged current (neutrino) DIS data, the significant experiments all use a heavy (Fe) target. Since these data are crucial for the determination of the normalization of the gluon distribution (indirectly via the momentum sum rule), and for quark flavor differentiation (in conjunction with the neutral current data), they play an important role in any comprehensive global analysis. For this purpose, a heavy-target correction is applied to the data, based on measured ratios for heavy-to-light targets from NMC and other experiments. Direct photon production data are not included because of serious theoretical uncertainties, as well as possible inconsistencies between existing experiments [5], [9]. The combination of neutral and charged DIS, lepton-pair production, lepton charge asymmetry, and inclusive large- p_T jet production processes provides a fairly tightly constrained system for the global analysis of PDFs. In total, there are ~ 1300 data points which meet the minimum momentum scale cuts that must be imposed to ensure that PQCD applies. The fractional uncertainties on these points are distributed roughly as dF/F over the range $F = 0.003 - 0.4$.

b. Parametrization: The non-perturbative parton distribution functions $f_a(x, Q)$ at a low momentum scale $Q = Q_0$ are parametrized by a set of functions of x , corresponding to the various flavors a . For this analysis, Q_0 is taken to be 1 GeV. The specific functional forms and the choice of Q_0 are not important, as long as the parametrization is general enough to accommodate the behavior of the true (but unknown) non-perturbative PDFs. The CTEQ analysis adopts the functional form

$$f(x, Q_0) = a_0 x^{a_1} (1 - x)^{a_2} (1 + a_3 x^{a_4}) \quad (1)$$

for most quark flavors as well as for the gluon.[‡] After momentum and quark number sum rules are enforced, there are 18 independent free parameters, hereafter referred to as the “shape parameters” $\{a_i\}$. The PDFs at $Q > Q_0$ are determined from $f_a(x, Q_0)$ by evolution equations from the renormalization group.

c. Fitting: The values of $\{a_i\}$ are determined by fitting the global experimental data to the theoretical expressions which depend on these parameters. The fitting is done by minimizing a global “chi-square” function, χ_{global}^2 . The quotation marks indicate that this function serves as a *figure of merit* of the quality of the global fit,

[‡]An exception is that recent data from E866 seem to require the ratio \bar{d}/\bar{u} to take a more unconventional functional form.

even though it does not have the full significance associated with rigorous statistical analysis, for reasons to be discussed extensively throughout the rest of this report. In practice, this function is defined as:

$$\chi_{\text{global}}^2 = \sum_n \sum_i w_n [(N_n d_{ni} - t_{ni}) / \sigma_{ni}^d]^2 + \sum_n [(1 - N_n) / \sigma_n^N]^2 \quad (2)$$

where d_{ni} , σ_{ni}^d , and t_{ni} denote the data, measurement uncertainty, and theoretical value (dependent on $\{a_i\}$) for the i^{th} data point in the n^{th} experiment. The second term allows the absolute normalization (N_n) for each experiment to vary, constrained by the published normalization uncertainty (σ_n^N). The w_n factors are weights applied to some critical experiments with very few data points, which are known (from physics considerations) to provide useful constraints on certain unique features of PDFs not afforded by other experiments. Experience shows that without some judiciously chosen weights, these experimental data points will have no influence in the global fitting process. The use of these weighing factors, to enable the relevant unique constraints, amounts to imposing certain prior probability (based on physics knowledge) to the statistical analysis.

In the above form, χ_{global}^2 includes for each data point the random statistical uncertainties and the combined systematic uncertainties in uncorrelated form, as presented by most experiments in the published papers. These two uncertainties are *combined in quadrature* to form σ_{ni}^d in (2). Detailed point to point correlated systematic uncertainties are not available in the literature in general; however, in some cases, they can be obtained from the experimental groups. For global fitting, uniformity in procedure with respect to all experiments favors the usual practice of merging them into the uncorrelated uncertainties. For the study of PDF uncertainties, we shall discuss this issue in more detail in Section V and Appendix A.

A. Goodness-of-fit for CTEQ5M

Before trying to calculate the numerical uncertainties, it is important to know some properties of the fit between data and theory. By “theory” we mean CTEQ5M – the best global fit within the parametrization (1) – a theory of the structure of the proton. The goodness-of-fit is measured by the global fitting function (2). We refer to this function as χ^2 , but it is good to remember that its statistical meaning is nontrivial because of the way the systematic errors are treated.

Table II lists the values of χ_n^2/N for the 15 data sets used in the global analysis, where N is the number of data points. This χ_n^2 is

$$\chi_n^2 = \sum_{i=1}^N [(N_n d_{ni} - t_{ni}) / \sigma_{ni}^d]^2; \quad (3)$$

as in (2) the measurement uncertainties combine both statistical and systematic errors, ignoring correlations. Roughly, χ_n^2/N is of order 1 for each experiment, and no experiment has a notably large value of χ_n^2/N . Therefore, there is good qualitative agreement between the theory and data. On the other hand, from the viewpoint of normal statistics, the values of χ_n^2/N deviate from 1 by amounts that would be very improbable if χ_n^2 were a true chi-square variable. The values of N are typically of order 150. For a chi-square distribution with 150 degrees of freedom, the probability for $\chi^2/N < 0.8$ is 0.034; and the probability for $\chi^2/N > 1.2$ is 0.048..

In normal statistics, if χ^2 is significantly greater than N , there are two possible explanations: The theory may be wrong, or the measurement errors (σ_{ni}^d) may be underestimated. If χ^2 is significantly less than N , there is only one explanation: The measurement errors are overestimated. We see examples of both of these cases in Table II.

For several reasons, the large deviations of χ_n^2/N from 1 are not surprising. Perhaps the most obvious is the treatment of the systematic error, which in (2) is combined in quadrature with the statistical error, point-by-point for the data. That is, the *correlations* between systematic errors of different points are ignored. This treatment tends to overestimate the error, and hence makes χ_n^2 smaller than the true deviation. That begins to account for the fact that some experiments have χ_n^2 significantly less than N . And of course the statistical and systematic errors estimated by the experimental analysis may not reflect accurately the true fluctuations of the data.

We conclude from Table II that the fitting function χ_{global}^2 , while it is our best measure of the goodness of fit between the theory and data, should not be treated as a standard chi-square variable of Gaussian statistics.

Another view of the overall fit between CTEQ5M and the data used in the fit is shown in Fig. 2. The graph is a histogram of the variable $\xi \equiv (d - t)/\sigma$ where d is a data value, σ the uncertainty of that measurement (statistical and systematic combined), and t the theoretical value for CTEQ5M. (The jet experiments are omitted from the histogram.) The curve in Fig. 2 has no adjustable parameters; it is the Gaussian with width 1 normalized to the total number of data points (1238, omitting the jet data). Over the entire data set, the theory fits the data within the assigned uncertainties σ_{ni}^d , indicating that those uncertainties are numerically consistent with the actual measurement fluctuations. However, this Figure lumps all the data together, and may hide disagreements between the theory and individual experiments. It just shows that *on average* there is a good fit. But one experiment may be high and another low, resulting in a consistent average. So Fig. 2 may be somewhat mislead-

ing.

Figures 3 and 4 show histograms of ξ for the H1 and ZEUS measurements of $F_2(x, Q)$, respectively. Again the curve is a Gaussian of width 1, with integral equal to the number of data points. These plots show that within the global fit, the fits for individual experiments may have very different characteristics. In the case of H1, the data are much narrower than the Gaussian, which is reflected by $\chi_n^2/N < 1$ in Table II. The uncertainties ascribed to the H1 measurements are evidently overestimated; the error may be exaggerated by our treatment of the systematic error – combining it with the statistical error ignoring correlations between data points. In the case of ZEUS, the data are much wider than the Gaussian, which is reflected by $\chi^2/N > 1$ in Table II. Either CTEQ5M is incorrect, or the uncertainties attributed to the ZEUS measurements are underestimated. In both cases the distribution of ξ is asymmetric about the origin. These Figures demonstrate the limitations of Gaussian statistics for comparing theory and data for parton distribution functions.

III. UNCERTAINTIES ON PDF PARAMETERS: THE ERROR MATRIX

We now describe results from an investigation of the behavior of the χ_{global}^2 function at its minimum, using the standard error matrix approach [21]. This allows us to determine which combinations of parameters are contributing the most to the uncertainty.

At the minimum of χ_{global}^2 , the first derivatives with respect to the $\{a_i\}$ are zero; so near the minimum, χ_{global}^2 can be approximated by

HOW ABOUT CHANGING NAME OF HESSIAN FROM F TO H??

$$\chi_{\text{global}}^2 = \chi_0^2 + \frac{1}{2} \sum_{i,j} F_{ij} y_i y_j \quad (4)$$

where $y_i = a_i - a_{0i}$ is the displacement from the minimum, and F_{ij} is the *Hessian*, the matrix of second derivatives. It is natural to define a new set of coordinates using the complete orthonormal set of eigenvectors of the symmetric matrix F_{ij} as basis vectors. These vectors can be ordered by their eigenvalues e_i . Each eigenvalue is a quantitative measure of the uncertainties in the shape parameters $\{a_i\}$ for displacements in parameter space in the direction of the corresponding eigenvector. The quantity $\ell_i \equiv 1/\sqrt{e_i}$ is the distance in the 18 dimensional parameter space, in the direction of eigenvector i , that makes a unit increase in χ_{global}^2 . ℓ_1 corresponds to displacements in the vector space $\{a_i\}$ in the direction that χ_{global}^2 changes most rapidly, while ℓ_{18} corresponds to the direction where it changes most slowly (flat direction). [MUST CHECK THAT THIS IS IN FACT THE

ORDER USED IN FIG 5] If the only measurement uncertainty were uncorrelated gaussian uncertainties, then ℓ_i would be one standard deviation from the best fit in the direction of the eigenvector. The inverse of the Hessian is the error matrix. [COULD GIVE HERE THE GENERAL FORMULA FOR UNCERTAINTY IN A PHYSICAL QUANTITY USING THIS ERROR MATRIX.]

Because the real uncertainties, for the wide variety of experiments included, are far more complicated than assumed in the ideal situation, the quantitative measure of a given increase in χ_{global}^2 carries little true statistical meaning. However, qualitatively, the Hessian gives an analytic picture of χ_{global}^2 near its minimum in $\{a_i\}$ space, and hence allows us to identify the particular degrees of freedom that need further experimental input in future global analyses.

From calculations of the Hessian we find that the eigenvalues vary over a wide range. Figure 5 shows a graph of the eigenvalues of F_{ij} , on a logarithmic scale. The vertical axis is $\ell_i = 1/\sqrt{e_i}$, the distance of a “standard deviation” along the i^{th} eigenvector. These distances range over 3 orders of magnitude. Large eigenvalues of F_{ij} correspond to “steep directions” of χ_{global}^2 . The corresponding eigenvectors are combinations of shape parameters that are well determined by current data. For example, parameters that govern the valence u and d quarks at moderate x are sharply constrained by DIS data. Small eigenvalues of F_{ij} correspond to “flat directions” of χ_{global}^2 . In the directions of these eigenvectors, χ_{global}^2 changes little over large distances in $\{a_i\}$ space. For example, parameters that govern the large- x behavior of the gluon distribution, or differences between sea quarks, properties of the nucleon that are not accurately determined by current data, contribute to the flat directions. The existence of flat directions is inevitable in global fitting, because as the data improve it only makes sense to increase the flexibility of the parametrizations of $f_a(x, Q_0)$ to fit the improved experimental constraints.

Because the eigenvalues of the Hessian have a large range of values, efficient calculation of F_{ij} requires an adaptive algorithm. In principle F_{ij} is the matrix of second derivatives at the minimum of χ_{global}^2 , which could be calculated from very small finite differences. In practice, small computational errors in the evaluation of χ_{global}^2 preclude the use of a very small step size—in part because adaptive numerical integration routines that are used in the pQCD calculations contain noise in the form of small discontinuities caused by changes in the subdivision of integration intervals. Coarse-grained finite differences thus yield a more accurate calculation of the second derivatives. But because the variation of χ_{global}^2 varies markedly in different directions, it is important to use a grid in $\{a_i\}$ space with small steps in steep directions and large steps in flat directions. This grid is generated by an iterative procedure, in which F_{ij} converges to a good estimate of the second derivatives.

From our iteratively-improved calculations of F_{ij} we

find that the minimum of χ_{global}^2 is fairly quadratic over large distances in the parameter space. Figures 6 and 7 show the behavior of χ_{global}^2 near the minimum along each of the 18 eigenvectors. χ_{global}^2 is plotted on the vertical axis, and the variable on the horizontal axis is the distance in $\{a_i\}$ space in the direction of the eigenvector, in units of $\ell_i = 1/\sqrt{e_i}$. There is some nonlinearity, but it is small enough that the Hessian can be used as an analytic model of the functional dependence of χ_{global}^2 on the shape parameters.

Having calculated the eigenvectors of the error matrix, we can examine the eigenvectors and find out which shape parameters of (1) – or rather which linear combinations of the shape parameters – are poorly determined by the data, and which are well determined.

Table III shows the eigenvectors in terms of the original PDF parameters $\{a_i\}$. It should be remembered that these parameters specify the PDF’s at the low Q scale, and applications of PDFs to Tevatron experiments use PDFs at a high Q scale. The evolution equations determine $f(x, Q)$ from $f(x, Q_0)$, so the functional form at Q depends on the $\{a_i\}$ in a complicated way. Nevertheless, there are some interesting observations to make about the eigenvectors of the error matrix. In Table III, each row is an eigenvector of the error matrix. (Recall, the error matrix is the inverse of the Hessian.) The eigenvectors have norm 1; components with squared value less than 0.01 are left blank in the Table. The eigenvectors are ordered by their corresponding eigenvalues, from largest error (the first row) to smallest error (the last row).

By examining the components of the eigenvectors, we can see which shape parameters are well known (with small error) and which are very uncertain (with large error). For example, the first row is dominated by the parameter “A4dou”; this is a fine tuning parameter for the difference between d and u sea quarks, so it is not surprising that its value is not very precise. The largest component in the second row is “A2g”, a parameter that controls the gluon at large x , and it is not surprising that this is poorly known; there is fairly strong admixture of “A2dpu”, which controls u and d sea quarks at large x , and it makes sense that this aspect of the proton is correlated with gluons at large x . As for the well-determined parameters, the last row of the Table is a certain linear combination of “Ng” (the total gluon fraction) and valence u quark parameters “A2u”, “A3u”, “A4u”; these are aspects of the parton structure of the proton that are accurately determined by deep-inelastic scattering. The next-to-the-last row is mainly another, orthogonal, combination of the same parameters. It is interesting to note that the parameter “Ng”, related to the total gluon content of the proton, overlaps only with the three eigenvectors with smallest uncertainty; this gluonic aspect of the proton is well-determined by current data. On the other hand, “A2g”, which governs the large- x behavior of the gluon overlaps only with eigenvectors with large uncertainty, so this aspect of the proton is not well known.

IV. UNCERTAINTY ON σ_W : THE LAGRANGE MULTIPLIER METHOD

In this Section, we determine the variation of χ_{global}^2 as a function of a single measurable quantity. We use the production cross section for W bosons (σ_W) as an archetypal example. The same method can be applied to any other physical observable of interest, for instance the Higgs production cross section, or to certain measured differential distributions. The aim is to quantify the uncertainty on that physical observable due to uncertainties of the PDFs integrated over the entire PDF parameter space.

Again, we use the standard CTEQ5 analysis tools and results [5] as the starting point. The “best fit” is the CTEQ5M1 set. A natural way to find the limits of a physical quantity X , such as σ_W at $\sqrt{s} = 1.8\text{ TeV}$, is to take X as one of the search parameters in the global fit and study the dependence of χ_{global}^2 for the 15 base experimental data sets on X .

Conceptually, we can think of the function χ_{global}^2 that is minimized in the fit as a function of a_1, \dots, a_{17}, X instead of a_1, \dots, a_{18} . This idea could be implemented directly in principle, but a more convenient way to do it in practice is through Lagrange’s method of undetermined multipliers. One minimizes, with respect to the $\{a_i\}$, the quantity

$$F(\lambda) = \chi_{\text{global}}^2 + \lambda X(a_1, \dots, a_{18}) \quad (5)$$

for a fixed value of λ , the Lagrange multiplier. By minimizing $F(\lambda)$ for many values of λ , we map out χ_{global}^2 as a function of X . The minimum of F for a given value of λ is the best fit to the data for the corresponding value of X , *i.e.*, the lowest possible χ_{global}^2 for that value of X .

END OF PUMPLIN’S REVISIONS FOR THIS SECTION.

Figure 8 shows χ_{global}^2 for the 15 base experimental data sets as a function of σ_W at the Tevatron. The horizontal axis is σ_W times the branching ratio for $W \rightarrow$ leptons, in nb. The CTEQ5m prediction is $\sigma_W \cdot BR_{\text{lep}} = 2.374\text{ nb}$. The vertical dashed lines are $\pm 3\%$ and $\pm 5\%$ deviations from the CTEQ5m prediction.

The two parabolas associated with points in Fig. 8 correspond to different treatments of the normalization factor N_n in Eq. 2. The dots (\bullet) are variable norm fits, in which N_n is allowed to float, taking into account the experimental normalization uncertainties, and $F(\lambda)$ is minimized with respect to N_n . The justification for this procedure is that overall normalization is a common systematic uncertainty. The boxes (\square) are fixed norm fits, in which all N_n are held fixed at their values for the global minimum (CTEQ5m). These two procedures represent extremes in the treatment of normalization uncertainty. The parabolas associated with \bullet ’s and \square ’s are just least-square fits to the points.

The other curve in Fig. 8 was calculated using the Hessian method. The Hessian F_{ij} is the matrix of second derivatives of χ_{global}^2 with respect to the shape parameters $\{a_i\}$. The derivatives (first and second) of σ_W may also be calculated by finite differences. Using the resultant quadratic approximations for $\chi_{\text{global}}^2(a)$ and $\sigma_W(a)$, one may minimize χ_{global}^2 with σ_W fixed. Since this calculation keeps the normalization factors constant, it should be compared with the fixed norm fits from the Lagrange multiplier method. The fact that the Hessian and Lagrange multiplier methods yield similar results lends support to both approaches; the small difference between them indicates that the quadratic functional approximations for χ_{global}^2 and σ_W are only approximations.

Figure 9 shows the same thing for pp collisions at $\sqrt{s} = 14\text{ TeV}$, for the LHC.

Finish this paragraph!

For the quantitative analysis of uncertainties, an important question is this: How large an increase in χ_{global}^2 should be taken to define the likely range of uncertainty in the constrained variable X ? There is a well-known theorem of statistics that states that an increase of $\Delta\chi^2 = 1$ in a constrained fit corresponds to 1 standard deviation of the constrained quantity X . However, this theorem does not apply to χ_{global}^2 , because the correlations between systematic errors are ignored in the definition (2) of this function. The theorem would apply if the errors were uncorrelated, Gaussian, and accurately known. At the very least, the first assumption is false for χ_{global}^2 .[§] Theoretically it is possible to define a fitting function χ^2 that includes the correlations between errors, for which $\Delta\chi^2 = 1$ does correspond to 1 standard deviation of the constrained variable. [21] [22] But that function is not χ_{global}^2 , because in χ_{global}^2 the systematic errors are combined with the statistical errors independently for each data point.

It can be shown that if the measurement errors are correlated, and the correlation is not properly taken into account in the definition of χ^2 , then a standard deviation may correspond to an increase $\Delta\chi^2$ much larger than 1. That is evidently the case for χ_{global}^2 . The issue of $\Delta\chi^2$ for a standard deviation, with correlated measurement errors, is explored in Appendix A.

V. COMPARISON OF THE CONSTRAINED FITS TO DATA

Figures 8 and 9 show how the fitting function χ_{global}^2 increases from its minimum value, at the best global fit,

[§]It has been shown by Giele *et al.* [4], that, taken literally, only one or two selected experiments satisfy the standard statistical tests.

as the cross section σ_W for W production is forced away from the prediction of the global fit. Next, we want to use this information, or some other analysis, to estimate the uncertainty in the prediction of σ_W attributable to PDF uncertainty. In ideal circumstances we could say that a certain increase of χ_{global}^2 from the minimum value, call it $\Delta\chi^2$, would correspond to a standard deviation of the global measurement uncertainty. Then a horizontal line on Fig. 8 or 9 at $\chi_{\text{min}}^2 + \Delta\chi^2$ would indicate the probable range of σ_W , by the intersection with the parabola of χ_{global}^2 versus σ_W .

However, such a simple estimate of the uncertainty of σ_W is not possible, because the fitting function χ_{global}^2 does not include the *correlations* between systematic errors. The uncertainty $\sigma_{n_i}^d$ used in the definition (2) of χ_{global}^2 combines in quadrature the statistical and systematic uncertainties for each data point; that is, it treats the systematic uncertainties as uncorrelated. The standard theorems of statistics for random errors do not apply to χ_{global}^2 . This point is discussed in detail in Appendix A.

We cannot use the value of χ_{global}^2 to estimate confidence levels on σ_W . Instead, we believe the best approach is to carry out a thorough statistical analysis, which includes the correlations of systematic errors, comparing individual experiments for which the detailed error information is available, to the constrained fits. The idea is that the constrained fits of Fig. 8 or 9 amount to alternate “hypotheses”; comparison of these different PDF sets with the DIS data should identify which “hypothesis” is best, or which can be ruled out. The virtue of the Lagrange multiplier method is that it gives us *the best fits to existing data* to use as an alternate “hypotheses” for particular values of the constrained quantity.

We will describe here such an analysis for the measurements of $F_2(x, Q)$ by the H1 experiment [12] at HERA, and by the BCDMS [10] experiment on μ scattering.

A. Comparison to H1 data

The H1 experiment has provided a detailed table of measurement uncertainties – statistical and systematic – for their measurements of $F_2(x, Q)$. [12] The CTEQ5 program uses 172 data points from H1 (imposing the cut $Q^2 > 5 \text{ GeV}^2$). For each measurement d_j (where $j = 1 \dots 172$) there is a statistical uncertainty σ_{0j} , an uncorrelated systematic uncertainty σ_{1j} , and a set of 4 correlated systematic uncertainties β_{jk} where $k = 1 \dots 4$. (More precisely, there are 8 correlated uncertainties listed in the H1 table. These correspond to 4 pairs. Each pair consists of one standard deviation in the positive sense, and one standard deviation in the negative sense, of some experimental parameter. For our analysis, we have approximated each pair of uncertainties by a single, symmetric combination, equal in magnitude to the average magnitude of the pair.)

To judge the uncertainty of σ_W at $\sqrt{s} = 1.8 \text{ TeV}$, as

constrained by the H1 data, we will compare the H1 data to the global fits in Fig. 8. The comparison is based on the true, statistical χ^2 , including the correlated uncertainties, which is given by

$$\chi^2 = \sum_j \frac{(d_j - t_j)^2}{\sigma_j^2} - \sum_{kk'} B_k (A^{-1})_{kk'} B_{k'}. \quad (6)$$

The index j labels the data points and runs from 1 to 172. The indices k and k' label the source of systematic uncertainty and run from 1 to 4. The combined uncorrelated uncertainty σ_j is $\sqrt{\sigma_{0j}^2 + \sigma_{1j}^2}$. The second term in (6) comes from the correlated systematic errors. B_k is the vector

$$B_k = \sum_j \frac{(d_j - t_j) \beta_{jk}}{\sigma_j^2}, \quad (7)$$

and $A_{kk'}$ is the matrix

$$A_{kk'} = \delta_{kk'} + \sum_j \frac{\beta_{jk} \beta_{jk'}}{\sigma_j^2}. \quad (8)$$

The derivation of (6), and in particular the proof that this quantity obeys a chi-square distribution, is supplied in Appendix B.

Assuming the published uncertainties σ_{0j} , σ_{1j} and β_{jk} accurately reflect the actual measurement fluctuations, χ^2 would obey a chi-square distribution if the measurements were repeated many times. Therefore the chi-square distribution with 172 degrees of freedom provides a basis for calculating *confidence levels* for the constrained fits in Fig. 8.

Table IV shows χ^2/N for the H1 data compared to seven of the PDF fits in Fig. 8. The center row of the Table is the global best fit – CTEQ5M. The other rows are fits obtained by the Lagrange multiplier method for different values of the Lagrange multiplier. The best fit to the H1 data, *i.e.*, the smallest χ^2 , is not CTEQ5M (the best global fit) but rather the fit with Lagrange multiplier 1000, for which σ_W is 0.8% smaller than the prediction of CTEQ5M. Forcing the value of the W cross section away from the prediction of CTEQ5M causes an increase in χ^2 for the DIS data. At $\sqrt{s} = 1.8 \text{ TeV}$, W production is mainly from $q\bar{q} \rightarrow W^+W^-$ with moderate values of x for q and \bar{q} , *i.e.*, values in the range of DIS experiments. Forcing σ_W higher (or lower) requires a higher (or lower) valence quark density in the proton, in conflict with the DIS data, so χ^2 increases.

The final column in Table IV, labeled “probability”, is computed from the chi-square distribution with 172 degrees of freedom. This quantity is the probability for χ^2 to be greater than the value calculated from the existing data, if the H1 measurements were to be repeated. So, for example, the fit with Lagrange multiplier -3000 , which corresponds to σ_W being 3.2% larger than the CTEQ5M prediction, has probability 0.092. In other words, if the

H1 measurements could be repeated many times, in only 9.2% of trials would χ^2 be greater than or equal to the value that has been obtained with the existing data. This probability represents a confidence level for the value of σ_W that was forced on the PDF by setting the Lagrange multiplier equal to -3000 . At the 9.2% confidence level we can say that $\sigma_W \cdot BR_{\text{lep}}$ is less than 2.450 nb, based on the H1 data. Similarly, at the 21.2% confidence level we can say that $\sigma_W \cdot BR_{\text{lep}}$ is greater than 2.294 nb.

Fig. 10 is a graph of χ^2/N for the H1 data compared to the PDF fits in Table IV. This figure may be compared to Fig. 8. The CTEQ5 prediction of the W production cross section is shown as an arrow, and the vertical dashed lines are $\pm 3\%$ away from the CTEQ5M prediction. The horizontal dashed line is the 68% confidence level on χ^2/N for $N = 172$ degrees of freedom. The comparison with H1 data alone indicates that the uncertainty on σ_W is $\sim 3\%$.

The H1 experiment is a good case, because for H1 we have detailed information about correlations of systematic errors. But it may be somewhat fortuitous that the χ^2 per data point for CTEQ5M is so close to 1 for the H1 data set. In other experiments where χ^2/N is not close to 1, which can easily happen if the estimated systematic uncertainties are not textbook-like, we must supply further arguments about confidence levels.

B. Comparison to BCDMS data

Another experiment that has provided information on the correlated systematic error is BCDMS [10]. This experiment measured $F_2(x, Q)$ from deep-inelastic muon scattering. Here we will compare the data to the constrained fits in Fig. 8, seeking again to evaluate the uncertainty of the prediction of σ_W at $\sqrt{s} = 1.8$ TeV. The method is the same as in the previous section.

The CTEQ5 program uses 168 data points from the BCDMS experiment on μp scattering. The BCDMS data table identifies measurement uncertainties of three kinds: statistical error σ_{0i} , total uncorrelated error σ_i , and a single correlated systematic error β_i . In this case the statistical variable χ^2 , including correlations, is

$$\chi^2 = \sum_{i=1}^N \frac{(m_i - t_i)^2}{\sigma_i^2} - \frac{B^2}{A} \quad (9)$$

where $N = 168$. Here B and A are just numbers,

$$B = \sum_{i=1}^N \beta_i (m_i - t_i) / \sigma_i^2, \quad (10)$$

$$A = \sum_{i=1}^N \beta_i^2 / \sigma_i^2. \quad (11)$$

We compute χ^2 for 7 of the constrained fits in Fig. 8. The results are given in Table V, which has the same meaning as Table IV.

We immediately notice that χ^2 for the BCDMS data is rather large. For CTEQ5M (the best global fit) χ^2/N is 1.109. That doesn't sound very large, until one computes the probability of such a large value for 168 data points; the probability of getting $\chi^2/N \geq 1.099$ with 168 degrees of freedom is only 0.158. This probability is not terribly low, but the low value does raise a question of the validity of CTEQ5M. If χ^2 is large, there are three possible explanations: The theory does not fit the data, or the experimental error is underestimated, or the measurements were very unlucky – fluctuating away from their true values. Before we reject the CTEQ5M PDF's, we should ask whether the actual experimental error might be slightly larger than published. With such a large number of data points (168) even a small inaccuracy in the published uncertainty can make a large gap in the calculated probability.

To appreciate the importance of the accuracy of σ , consider the following example. If σ is underestimated by 5% of its true value (not 5% of the data value but 5% of the uncertainty itself!) then χ^2 should be 10% smaller. The probability for $\chi^2/N > 1.1$ with $N = 168$ is 0.178. Reducing χ^2 by 10% brings it to 1.0, for which the probability is 0.485. A 5% underestimate of σ translates to a factor of 0.37 error on the probability.

Fig. 11 is a graph of χ^2/N for the BCDMS data compared to the PDF fits in Table V. The vertical and horizontal dashed lines have the same meaning as in Fig. 10, and CTEQ5M is indicated by the arrow. The points connected by solid lines are the values of χ^2/N calculated from (9). All values of χ^2 are considerably outside the range of the 68% confidence level.

For experiments with many data points, like 168 for BCDMS, the chi-square distribution is very narrow, so a small inaccuracy in the estimate of σ_j may translate to a large error in probability calculations based on the absolute value of χ^2 . Because the estimation of experimental uncertainties introduces some uncertainty in the value of χ^2 itself, it is not really the *absolute* value of χ^2 that is most relevant, but rather the *relative* value compared to the value at the minimum. Therefore, we believe it makes sense to *renormalize* the values of χ^2 by studying *ratios* of χ^2 's, in order to interpret the variation of χ^2 with the constrained variable σ_W . This procedure is mathematically equivalent to increasing the error estimates so as to make $\chi^2/N = 1$ at the global minimum. The ratios are shown in Fig. 11 as the points connected by dashed lines. Interpreting these points as “corrected” χ^2 's, we find results quite consistent with the earlier H1 calculations: The uncertainty of the prediction of σ_W is about $\pm 3\%$.

One may criticize the renormalization procedure by saying that it is *assuming the conclusion* (that CTEQ5M is a correct description of proton structure). But we assume that the published values of σ_j are subject to some uncertainty, and in the spirit of Bayesian statistics, we apply a correction based on something we believe to be true – that CTEQ5M describes the proton well. Then

the *increase* of χ^2 as σ_W moves away from the predicted value indicates the uncertainty in the prediction.

C. Comparison to all individual experiments, ignoring correlations

For most of the experiments used in the CTEQ5 global analysis, listed in Table I, detailed information on the correlations of systematic errors is not available. Even if it were available, the BCDMS case suggests that the information might not be accurate enough for calculating confidence levels. With many degrees of freedom, a small inaccuracy in σ translates to a large error in the calculated probability.

In this section we will examine how χ_u^2 varies with σ_W for the constrained fits. Here χ_u^2 is data minus theory, weighted by the error,

$$\chi_u^2 = \sum_{i=1}^N \frac{(d_i - t_i)^2}{\sigma_i^2} \quad (12)$$

where σ_i is the statistical and systematic error on the data point d_i , combined in quadrature. That is, χ_u^2 is the uncorrelated chi square.

But we believe that we are *overestimating* the measurement uncertainty in (12) by ignoring the correlations between systematic errors. Therefore the *absolute* value of χ_u^2 cannot be used to set confidence levels on the fits. What is more important is the *relative* value of χ_u^2 for the constrained fits, compared to CTEQ5M. If CTEQ5M is a correct description of the proton, then the value of σ_W calculated from CTEQ5M should be the correct value. Forcing σ_W away from that prediction produces a poorer fit to the existing data on PDFs, as shown in Fig. 8. Now, how does the fit change for each experiment used in the CTEQ5 program? For this we compare the *ratio* of χ_u^2 for the constrained fit to χ_u^2 for CTEQ5M.

Figure 12 shows, for each of the 15 experiments in Table I, the ratio $\chi_u^2/\chi_{\text{CTEQ5M}}^2$ versus σ_W . By plotting the ratios, most of the points fit on the same scale, from 0.8 to 1.2. For most experiments, χ_u^2 increases as σ_W is forced away from the prediction of CTEQ5M, but there are some notable exceptions. For most DIS experiments, changing σ_W by $\pm 3\%$ causes χ_u^2 to increase significantly. For some experiments – for example jet production in the CDF and D0 experiments – there is little change in χ_u^2 as σ_W is forced to differ from the prediction. If we may treat the ratio as a “corrected” χ^2 then the results of 15 experiments are reasonably consistent with the statement that the uncertainty on σ_W , due to uncertainty in PDFs, is $\pm 3\%$.

We conclude from Fig. 12 that it is important to use many experiments to determine the parton distribution functions. By doing a *global* fit, differences between experiments tend to average out. The figure demonstrates that the individual experiments do differ significantly

from the average, and hence from each other. Unless there is a rationale for excluding one experiment in favor of another, the best we can do is to construct the best fit to the collective data.

D. Comments on Variable norm fits

Figures 8 and 9 show that the value of χ_{global}^2 for a constrained fit can be significantly reduced by allowing the overall normalization of the experiment to vary. In the previous analysis we have considered “fixed norm fits” – constrained fits with the overall normalization of each experiment the same as in CTEQ5M.

The overall normalization the most obvious source of systematic error in an experiment, so the CTEQ5 program introduces a normalization factor N_n for each experiment, and includes these factors as fitting parameters along with the shape parameters. Each experiment has published the estimated uncertainty in normalization, so the fitting requires that the value of N_n lie within the published range of uncertainty. Typical values of N_n for CTEQ5M are a few percent above or below 1.

The added flexibility of the fit that results from allowing N_n to vary, in the constrained fits with a Lagrange multiplier, produces much lower values of χ_{global}^2 compared to the fixed normalization fit, as we see from Figs. 8 and 9. But we need to examine the normalization factors for the constrained fits, to see whether they are reasonable.

Complete this discussion

VI. CONCLUSIONS

It has been widely recognized by the HEP community, and it has been emphasized at this workshop, that PDF phenomenology must progress from the past practice of periodic updating of *representative* PDF sets to a systematic effort to map out the uncertainties, both on the PDFs themselves and on physical observables derived from them. For the analysis of PDF uncertainties, we have only addressed the issues related to the treatment of experimental uncertainties. Equally important for the ultimate goal, one must come to grips with uncertainties associated with theoretical approximations and phenomenological parametrizations. Both of these sources of uncertainties induce highly correlated uncertainties, and they can be numerically more important than experimental uncertainties in some cases. Only a balanced approach is likely to produce truly useful results. Thus, great deal of work lies ahead.

This report described first results from two methods for quantifying the uncertainty of parton distribution functions associated with experimental uncertainties. The specific work is carried out as extensions of the CTEQ5

global analysis. The same methods can be applied using other parton distributions as the starting point, or using a different parametrization of the non-perturbative PDFs. We have indeed tried a variety of such alternatives. The results are all similar to those presented above. The robustness of these results lends confidence to the general conclusions.

The Hessian, or error matrix method reveals the uncertainties of the shape parameters used in the functional parametrization. The behavior of χ_{global}^2 in the neighborhood of the minimum is well described by the Hessian if the minimum is quadratic.

The Lagrange multiplier method produces constrained fits, *i.e.*, the best fits to the global data set for specified values of some observable. The increase of χ_{global}^2 , as the observable is forced away from the predicted value, indicates how well the current data on PDFs determines the observable.

The constrained fits generated by the Lagrange multiplier method may be compared to data from individual experiments, taking into account the uncertainties in the data, to estimate confidence levels for the constrained variable. For example, we estimate that the uncertainty of σ_W attributable to PDFs is $\pm 3\%$.

Further work is needed to apply these methods to other measurements, such as the W mass or the forward-backward asymmetry of W production in $p\bar{p}$ collisions. Such work will be important in the era of high precision experiments.

APPENDIX A

The global fitting function χ_{global}^2 defined in (2) resembles the standard statistical variable χ^2 , and it is tempting to try to apply theorems of Gaussian statistics to analyze the significance of the fit between theory and experiment. However, the familiar theorems do not apply, because of the correlations between measurement errors. The purpose of this Appendix is to explore this issue. The material here is textbook probability and statistics, but nontrivial nonetheless. In particular, the effect of correlated errors can be a source of confusion if it is not clearly understood.

For simplicity we shall describe the simplest case, namely, measurement of a single observable. The arguments are easily extended to cases where multiple quantities are measured, such as the determination of parton distribution functions.

Consider an observable m that is measured N times. We shall refer to N measurements of m as one “experiment”. Let the true value of m be m_0 . The measurements are $m_1, m_2, m_3, \dots, m_N$. The deviations from the true value are $\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_N$, where $\alpha_i = m_i - m_0$. In general the measurement errors are correlated, so in the Gaussian approximation the probability distribution of the fluctuations is

$$dP = \mathcal{N} \exp \left\{ -\frac{1}{2} \sum_{i,j=1}^N \alpha_i C_{ij} \alpha_j \right\} d^N \alpha. \quad (13)$$

Here C_{ij} is a real symmetric matrix, referred to below as the correlation matrix. The constant $\mathcal{N} = \sqrt{\text{Det } C} / (2\pi)^{N/2}$ is determined by the normalization condition $\int dP = 1$.

We will require the variance matrix $\langle \alpha_i \alpha_j \rangle$, where the notation $\langle Q \rangle$ means the average of Q in the probability distribution (13). For this Gaussian distribution,

$$\langle \alpha_i \alpha_j \rangle = (C^{-1})_{ij}. \quad (14)$$

(This result and the normalization integral are most easily derived by using the eigenvectors of C as integration variables.)

The mean square fluctuation E_i of the i^{th} measurement m_i is

$$E_i \equiv \langle \alpha_i^2 \rangle = (C^{-1})_{ii}. \quad (15)$$

To find the best estimate of the value of m from these N measurements, *ignoring the correlations in the measurement errors*, we define a chi-square function $\chi_u^2(m)$ by

$$\chi_u^2(m) = \sum_{i=1}^N \frac{(m_i - m)^2}{E_i}. \quad (16)$$

The value of m that minimizes $\chi_u^2(m)$, call it \bar{m} , is then the best estimate of m_0 based on this information. The function $\chi_u^2(m)$ is directly analogous to the fitting function χ_{global}^2 in the CTEQ program, in the sense that it does not include any information about the correlations between errors. The minimum of $\chi_u^2(m)$ occurs at a weighted average of the measurements,

$$\bar{m} = \frac{\sum_{i=1}^N m_i / E_i}{\sum_{i=1}^N 1 / E_i}. \quad (17)$$

If all the E_i 's are equal then \bar{m} is just the average of the measurements.

Now, what are the fluctuations of the mean \bar{m} ? That is, if the “experiment” consisting of N measurements could be replicated many times, what would be the distribution of \bar{m} 's obtained in those many trials? It turns out that \bar{m} has a Gaussian distribution

$$\frac{dP}{d\bar{m}} = \frac{1}{\sqrt{2\pi}\Sigma} \exp [-(\bar{m} - m_0)^2 / (2\Sigma^2)]. \quad (18)$$

The standard deviation Σ of \bar{m} is

$$\Sigma^2 = \frac{1}{D^2} \sum_{ij} \frac{(C^{-1})_{ij}}{E_i E_j} \quad (19)$$

where

$$D = \sum_i \frac{1}{E_i}. \quad (20)$$

The question we wish to answer is this: *How much does $\chi_u^2(m)$ increase, when m moves away from the minimum at \bar{m} by the amount $\pm \Sigma$ that corresponds to one standard deviation of the mean?* The answer to this question is

$$\Delta\chi_u^2 = \Sigma^2 D. \quad (21)$$

This result follows easily from the definition (16), since

$$\chi_u^2(\bar{m} + \Sigma) - \chi_u^2(\bar{m}) = -2\Sigma \sum_i \frac{m_i - \bar{m}}{E_i} + \Sigma^2 \sum_i \frac{1}{E_i}, \quad (22)$$

and the linear term in Σ is 0 by the definition of \bar{m} . So far the discussion has been quite general. We will now discuss some illustrative special cases.

Example 1: Suppose the measurement errors are uncorrelated; that is,

$$C_{ij} = \delta_{ij}/E_i. \quad (23)$$

Then the standard deviation of the mean \bar{m} is $\Sigma = 1/\sqrt{D}$. Therefore for the uncorrelated case, the increase of χ_u^2 corresponding to one standard deviation of the mean is $\Delta\chi_u^2 = 1$. This is the “normal” statistical result: the 1 S.D. range corresponds to an increase by 1 in $\Delta\chi^2$.

An even more special case—and the most familiar one—is when the errors are uncorrelated and constant. This case has $E_i = \sigma^2$ independent of i , where σ is the standard deviation of single measurements. The correlation matrix is $C_{ij} = \delta_{ij}/\sigma^2$. In this case D is N/σ^2 , and the standard deviation of the mean is $\Sigma = \sigma/\sqrt{N}$.

The criterion $\Delta\chi^2 = 1$ for one standard deviation of a measured quantity is a standard result, often used in the analysis of precision data. But if χ^2 is defined ignoring the correlations between measurement errors, then the criterion $\Delta\chi^2 = 1$ is only valid for uncorrelated errors. We will next consider two examples with correlated errors, to show that $\Delta\chi_u^2$ is not 1 for such cases.

Example 2: Suppose measurements 1 and 2 are correlated, 3 and 4 are correlated, 5 and 6 are correlated, etc. Then the correlation matrix is

$$\begin{aligned} C_{ij} &= 1/\sigma^2 \text{ for } i = j \\ C_{ij} &= C_{ji} = c/\sigma^2 \text{ for } i \text{ odd and } j = i + 1 \\ C_{ij} &= 0 \text{ otherwise} \end{aligned} \quad (24)$$

where $-1 < c < 1$ since the determinant of C must be positive. The inverse matrix C^{-1} can be constructed using the fact that C is block diagonal, consisting of $N/2$ 2×2 blocks. Then it can be shown that

$$\Sigma = \frac{\sigma}{\sqrt{N}\sqrt{1+c}} \text{ and } \Delta\chi_u^2 = 1 - c. \quad (25)$$

The increase of χ_u^2 for one standard deviation of the mean ranges from 0 to 2, depending on c . The criterion $\Delta\chi^2 = 1$ does not apply to this example with correlated errors. A standard increase of χ_u^2 may be smaller or larger than 1.

Example 3: For an even more striking example, suppose the N measurements that constitute a single “experiment” are, for $i = 1, 2, 3, \dots, N$

$$m_i = m_0 + y + \beta_i \quad (26)$$

where the β_i are randomly distributed with standard deviation σ , and the measurements are systematically off by the amount y . Suppose that y has a Gaussian distribution with standard deviation s for replications of the “experiment”. In this example,

$$C_{ij} = \frac{1}{\sigma^2} \left(\delta_{ij} - \frac{s^2}{Ns^2 + \sigma^2} \right) \quad (27)$$

$$(C^{-1})_{ij} = \sigma^2 \delta_{ij} + s^2 \quad (28)$$

The standard deviation of the individual measurements m_i 's is

$$\langle m^2 \rangle - \langle m \rangle^2 = \langle \alpha^2 \rangle - \langle \alpha \rangle^2 = \sigma^2 + s^2. \quad (29)$$

Therefore $\chi_u^2(m)$, ignoring the correlations, is

$$\chi_u^2(m) = \sum_{i=1}^N \frac{(m - m_i)^2}{\sigma^2 + s^2}. \quad (30)$$

The minimum of $\chi_u^2(m)$ occurs at \bar{m} , which is just the average of the individual measurements. The variance of \bar{m} , averaged over many replications of the “experiment” is

$$\Sigma^2 = \langle \bar{m}^2 \rangle - \langle \bar{m} \rangle^2 = s^2 + \frac{\sigma^2}{N}. \quad (31)$$

The increase of χ_u^2 as m moves from \bar{m} to $\bar{m} \pm \Sigma$, *i.e.*, by one standard deviation of the mean, is

$$\begin{aligned} \Delta\chi_u^2 &= \chi_u^2(\bar{m} + \Sigma) - \chi_u^2(\bar{m}) \\ &= \frac{\sigma^2 + Ns^2}{\sigma^2 + s^2}. \end{aligned} \quad (32)$$

In the limit $s/\sigma \ll 1$, the error correlations in this model become negligible and $\Delta\chi^2$ reduces to the conventional value of 1. But in the limit $s/\sigma \gg 1$ where the error correlations are dominant, $\Delta\chi^2$ approaches N .

END OF PUMPLIN'S REVISIONS FOR THIS APPENDIX

Thus for this example, which has 100% correlation between errors, the increase of χ_u^2 is much larger than 1. If s and σ are comparable, then $\Delta\chi_u^2$ is of order N .

Theoretically, correlated measurement errors can be incorporated into a fitting function, by defining

$$\chi_c^2(m) = \sum_{i=1}^N \sum_{j=1}^N (m - m_i) C_{ij} (m - m_j). \quad (33)$$

The minimum of χ_c^2 occurs at \tilde{m} , which is the best estimate of m_0 . The standard deviation of this mean is $\tilde{\Sigma}$, and it can be shown that the increase in $\chi_c^2(m)$ as m moves from \tilde{m} by $\pm\tilde{\Sigma}$ is $\Delta\chi_c^2 = 1$. But this treatment of correlated errors is applicable only if the correlation matrix C_{ij} is known accurately. For the full list of experiments in the global analysis of parton distribution functions, the published correlations of systematic errors are not sufficiently complete or accurate to use χ_c^2 as the fitting function. [4], [22]

In this Appendix we have only considered the measurement of a single quantity. The determination of parton distribution functions seeks to measure *many* quantities, *i.e.*, the ≈ 18 shape parameters, and ≈ 10 normalization constants, in the parametrization (1). The above arguments can be extended to measurements of multiple quantities. If the measurement errors are uncorrelated, then the increase of χ^2 by 1 from the minimum defines a hyperellipse in parameter space – the *error ellipse* – corresponding to one standard deviation of linear combinations of the parameters. However, if the errors are correlated then $\Delta\chi_u^2 = 1$ is not the correct criterion for a standard deviation.

The Lagrange multiplier method finds the best fit to the data, subject to a constrained value of some quantity. Again, if the errors are uncorrelated then one standard deviation of the constrained quantity corresponds to an increase of χ^2 by 1 from the global (unconstrained) minimum. But if the errors are correlated then $\Delta\chi_u^2 = 1$ is not the correct criterion for one standard deviation of the constrained variable.

One reason we have described this familiar, even elementary, statistics, is to avoid certain misconceptions. CTEQ5M is a parametrized fit to 1295 data points with approximately 30 fitting parameters. The minimum value of χ_{global}^2 is approximately 1200. Naively it seems that an increase of χ_{global}^2 by merely 1, say from 1200 to 1201, could not possibly represent a standard deviation of the fit. Naively one might suppose that a standard deviation would have $\Delta\chi^2 \sim \sqrt{1295}$, or perhaps $\Delta\chi^2 \sim \sqrt{30}$, rather than 1. However, this is a misconception. If the errors are determined accurately and uncorrelated, or if the correlations are known and incorporated into χ^2 , then indeed $\Delta\chi^2 = 1$ does represent a standard deviation. But this theorem is irrelevant to our problem, because the errors are correlated. It must be understood that this familiar theorem of statistics cannot be applied, because we have insufficient knowledge of the correlations of systematic errors.

APPENDIX B

The purpose of this appendix is to derive the appropriate definition of χ^2 for data with correlated systematic errors. The defining condition is that χ^2 should obey a chi-square distribution.

Let $\{m_i\}$ be a set of measurements, where $i = 1, 2, 3, \dots, N$. Let t_i be the true, *i.e.*, theoretical value of the i^{th} measured quantity. Several kinds of measurement errors will contribute to the difference between m_i and t_i . The uncorrelated error is denoted by σ_i . Also, there are some number K of correlated errors, denoted $\beta_{1i}, \beta_{2i}, \dots, \beta_{Ki}$. Thus the i^{th} measurement can be written as

$$\begin{aligned} m_i &= t_i + \text{errors} \\ &= t_i + \sigma_i r_i + \sum_{j=1}^K \beta_{ji} r'_j \end{aligned} \quad (34)$$

where r_i and r'_j are independently fluctuating variables. We assume that each of these fluctuations has a Gaussian distribution with width 1,

$$p(r) = \frac{e^{-r^2/2}}{\sqrt{2\pi}}. \quad (35)$$

Note that r'_j is independent of i ; that is, the errors $\beta_{j1}, \beta_{j2}, \dots, \beta_{jN}$ are 100% correlated for all data points.

The probability distribution of the measurements is

$$\begin{aligned} dP &= \int \prod_{i=1}^N f(r_i) dr_i \prod_{j=1}^K f(r'_j) dr'_j \\ &\times \prod_{i=1}^N \delta \left(m_i - t_i - \sigma_i r_i - \sum_{j=1}^K \beta_{ji} r'_j \right) d^N m. \end{aligned} \quad (36)$$

Is an exercise in Gaussian integration to evaluate the integrals. The result is

$$dP = \mathcal{N} \exp \left[-\frac{1}{2} \chi^2 \right] d^N m \quad (37)$$

where \mathcal{N} is a normalization constant. The quantity χ^2 is

$$\chi^2 = \sum_{i=1}^N \frac{(m_i - t_i)^2}{\sigma_i^2} - \sum_{j=1}^K \sum_{j'=1}^K B_j (A^{-1})_{jj'} B_{j'} \quad (38)$$

where B_j is a vector with K components

$$B_j = \sum_{i=1}^N \beta_{ji} (m_i - t_i) / \sigma_i^2, \quad (39)$$

and $A_{jj'}$ is a $K \times K$ matrix

$$A_{jj'} = \delta_{jj'} + \sum_{i=1}^N \beta_{ji} \beta_{j'i} / \sigma_i^2. \quad (40)$$

To check that (38) makes sense we can consider a special case. Suppose the number K of systematic errors is N , and each systematic error contributes to just one measurement. Then the matrix of systematic errors has the form

$$\beta_{ji} = \delta_{ji} b_i. \quad (41)$$

This situation is equivalent to an additional set of *uncorrelated* errors $\{b_i\}$. The vector B_j is

$$B_j = \frac{b_j (m - m_j)}{\sigma_j^2} \quad (42)$$

and the matrix $A_{jj'}$ is

$$A_{jj'} = \delta_{jj'} \left[1 + \frac{b_j^2}{\sigma_j^2} \right]. \quad (43)$$

Substituting these results into (38) we find

$$\chi^2 = \sum_i \frac{(m - m_i)^2}{\sigma_i^2 + b_i^2}, \quad (44)$$

which makes sense: The uncorrelated errors just combine in quadrature.

The statistical quantity χ^2 has a chi-square distribution with N degrees of freedom. Thus this variable may be used to set confidence levels of the theory for the given data. But to use this variable, the measurement errors σ_i and β_{ji} , for $i = 1, 2, \dots, N$ and $j = 1, 2, \dots, K$, must be known from the experiment.

A chi-square distribution with many degrees of freedom is a very narrow distribution, sharply peaked at $\chi^2 = N$. Therefore small inaccuracies in the values of the σ_i 's and β_{ji} 's may translate into a large error on the confidence levels computed from the chi-square distribution.

It is equation (38) that we use in Section V to compare the constrained fits produced by the Lagrange multiplier method to data from DIS experiments.

-
- [1] CDF Collaboration (Abe et al.), *Phys. Rev. Lett.* **77**, (1996) 439.
- [2] J. Huston, E. Kovacs, S. Kuhlman, H. L. Lai, J. F. Owens, D. Soper, W. K. Tung, *Phys. Rev. Lett.* **77**, 444(1996); E. W. N. Glover, A. D. Martin, R. G. Roberts, and W. J. Stirling, "Can partons describe the CDF jet data?", hep-ph/9603327.
- [3] S. Alekhin, *Eur. Phys. J.* **C10**, 395 (1999) [hep-ph/9611213]; and contribution to Proceedings of *Standard Model Physics (and more) at the LHC*, 1999.
- [4] W. T. Giele and S. Keller, *Phys. Rev.* **D58**, 094023; contribution to this Workshop by W. T. Giele, S. Keller and D. Kosower; and private communication.
- [5] H. L. Lai, J. Huston, S. Kuhlmann, J. Morfin, F. Olness, J. F. Owens, J. Pumplin and W. K. Tung, hep-ph/9903282, (to appear in *Eur. J. Phys.*); and earlier references cited therein
- [6] A. D. Martin and R. G. Roberts and W. J. Stirling and R. S. Thorne, *Eur. Phys. J.* **C4**, (1998) 463, hep-ph/9803445; and earlier references cited therein.
- [7] M. Gluck and E. Reya and A. Vogt, *Eur. Phys. J.* **C5**, (1998) 461, hep-ph/9806404.
- [8] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, "Parton Distributions and the LHC: W and Z Production", hep-ph/9907231.
- [9] J. Huston, et al, *Phys. Rev.* **D51**, (1995) 6139, hep-ph/9501230. L. Apanasevich, C. Balazs, C. Bromberg, J. Huston, A. Maul, W. K. Tung, S. Kuhlmann, J. Owens, M. Begel, T. Ferbel, G. Ginther, P. Slattery, M. Zielinski, *Phys.Rev.* **D59**, 074007 (1999); P. Aurenche, M. Fontanaz, J.Ph. Guillet, B. Kniehl, E. Pilon, M. Werlen, *Eur. Phys. J.* **C9**, 107 (1999).
- [10] BCDMS Collaboration (A.C. Benvenuti, et al.), *Phys.Lett.* **B223**, 485 (1989); and *Phys. Lett.* **B237**, 592 (1990).
- [11] NMC Collaboration: (M. Arneodo et al.) *Phys. Lett.* **B364**, 107 (1995).
- [12] H1 Collaboration (S. Aid et al.): "1993 data" *Nucl. Phys.* **B439**, 471 (1995); "1994 data", DESY-96-039, e-Print Archive: hep-ex/9603004; and H1 Webpage.
- [13] ZEUS Collaboration (M. Derrick et al.): "1993 data" *Z. Phys.* **C65**, 379 (1995) ; "1994 data", DESY-96-076 (1996).
- [14] CCFR Collaboration (W.C. Leung, et al.), *Phys. Lett.* **B317**, 655 (1993); and (P.Z. Quintas, et al.), *Phys. Rev. Lett.* **71**, 1307 (1993).
- [15] E605: (G. Moreno, et al.), *Phys. Rev.* **D43**, 2815 (1991).
- [16] E866 Collaboration (E.A. Hawker, et al.), *Phys. Rev. Lett.* **80**, 3175 (1998).
- [17] NA51 Collaboration (A. Baldit, et al.), *Phys. Lett.* **B332**, 244 (1994).
- [18] CDF Collaboration (F. Abe, et al.), *Phys. Rev. Lett.* **74**, 850 (1995).
- [19] See [1] and F. Bedeschi, talk at 1999 Hadron Collider Physics Conference, Bombay, January, 1999.
- [20] D0 Collaboration: B. Abbott et al., FERMLAB-PUB-98-207-E, e-Print Archive: hep-ex/9807018
- [21] D.E. Soper and J.C. Collins, "Issues in the Determination of Parton Distribution Functions", CTEQ Note 94/01; hep-ph/9411214.
- [22] M. Botje, "A QCD analysis of HERA and fixed target structure function data", hep-ph/9912439.

TABLES

Process	Experiment	Measurable	N_{data}
DIS	BCDMS [10]	F_{2H}^μ, F_{2D}^μ	324
	NMC [11]	F_{2H}^μ, F_{2D}^μ	240
	H1 [12]	F_{2H}^e	172
	ZEUS [13]	F_{2H}^e	186
	CCFR [14]	$F_{2Fe}^\nu, x F_{3Fe}^\nu$	174
Drell-Yan	E605 [15]	$sd\sigma/d\sqrt{\tau}dy$	119
	E866 [16]	$\sigma(pd)/2\sigma(pp)$	11
	NA-51 [17]	A_{DY}	1
W-prod.	CDF [18]	Lepton asym.	11
Incl. Jet	CDF [19]	$d\sigma/dE_t$	33
	D0 [20]	$d\sigma/dE_t$	24

TABLE I. List of processes and experiments used in the CTEQ5 global analysis. The total number of data points is 1295.

experiment	data set	N	χ_n^2/N
BCDMS	bcd1h2	168	0.863
BCDMS	bcd1d2	156	1.44
H1	h1f2	172	0.626
ZEUS	zeusf2	186	1.35
NMC	NmcF2p	104	1.07
NMC	NmcF2r	123	0.904
NMC	NmcF2rX	13	0.894
CCFR	ccfrf2	87	0.836
CCFR	ccfrf3	87	0.359
E605	e605	119	0.798
NA51	NA51	1	0.437
CDF	cdfLasy	11	0.790
E866	e866b.mod	11	0.485
D0	D0Ibst	24	0.872
CDF	cdf1jet	33	1.68

TABLE II. χ_n^2/N for each experiment, for the global fit CTEQ5M. N is the number of data points used from the experiment.

A1g	A2g	A3g	Ng	A1dpu	A2dpu	A3dpu	Ndmu	A1dou	A2dou	A3dou	A4dou	A2u	A3u	A4u	A2d	A3d	A4d
.	0.16	0.20	0.19	.	-.94
-0.11	-0.84	.	.	.	0.46	0.18	.	0.13
.	0.17	-0.66	.	.	0.37	0.56	.	-0.18	-0.12	0.10
.	.	0.14	.	.	-0.12	-0.15	0.13	-0.68	-0.47	0.41	-0.27
.	-0.28	-0.71	.	.	-0.45	-0.46
.	-0.25	0.12	-0.14	-0.71	-0.46	-0.42
0.14	-0.31	0.10	.	-0.23	-0.59	0.61	0.18	.	-0.11	.	.	-0.11	.	.	0.13	.	.
.	-0.10	.	.	.	-0.16	0.13	-0.49	-0.34	0.66	0.33	-0.15
-0.12	.	.	.	0.10	.	.	0.11	-0.16	0.21	.	.	-0.52	-0.35	-0.44	.	-0.35	0.41
.	-0.14	0.41	0.12	0.40	.	-0.51	0.59
.	0.71	.	0.38	0.47	0.17	.	0.22	-0.11	.	.	-0.13
0.67	.	.	.	-0.46	0.13	-0.18	.	-0.13	0.11	-0.13	.	.	0.10	.	-0.33	0.21	0.26
-0.47	.	.	.	0.16	-0.10	0.13	0.14	-0.23	.	-0.29	.	0.21	0.29	-0.12	-0.40	0.43	0.25
.	0.36	-0.10	0.47	.	-0.11	-0.45	0.26	-0.36	0.37	0.25
.	0.36	-0.32	0.20	-0.36	.	.	-0.53	0.49	.	.	-0.20
0.47	.	.	0.34	0.71	0.23	.	-0.27	.	.	.
0.22	.	.	-0.78	0.41	-0.27	0.21	0.20	.	.	.
.	.	.	-0.50	-0.10	0.60	-0.42	-0.43	.	.	.

TABLE III. Table of eigenvector mapping to PDF parameters.

Lagrange multiplier	$\sigma_W \cdot B$ in nb	$\chi^2/172$	probability
3000	2.294	1.0847	0.212
2000	2.321	1.0048	0.468
1000	2.356	0.9676	0.605
0	2.374	0.9805	0.558
-1000	2.407	1.0416	0.339
-2000	2.431	1.0949	0.187
-3000	2.450	1.1463	0.092

TABLE IV. Comparison of H1 data to the PDF fits with constrained values of σ_W . (B is the branching ratio for $W \rightarrow$ leptons.)

Lagrange multiplier	$\sigma_W \cdot B$ in nb	$\chi^2/168$	probability
3000	2.294	1.2057	0.036
2000	2.321	1.1669	0.068
1000	2.356	1.1115	0.153
0	2.374	1.1094	0.158
-1000	2.407	1.1210	0.135
-2000	2.431	1.1590	0.077
-3000	2.450	1.2131	0.031

TABLE V. Comparison of BCDMS data to the PDF fits with constrained values of σ_W . (B is the branching ratio for $W \rightarrow$ leptons.)

FIGURES

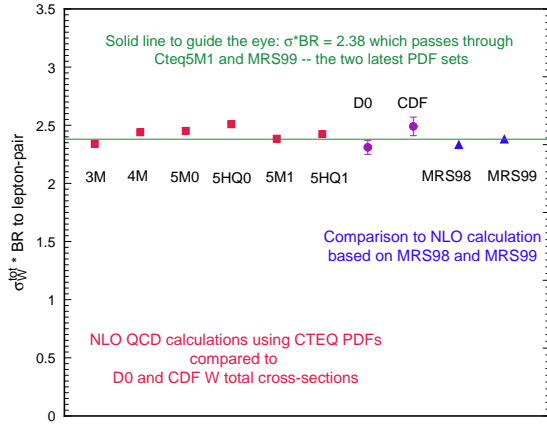


FIG. 1. Predicted cross section for W boson production for various PDFs.

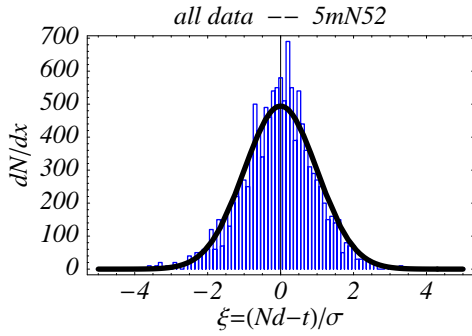


FIG. 2. Histogram of $\xi \equiv (Nd - t)/\sigma$ for all data points from the CTEQ5M fit.

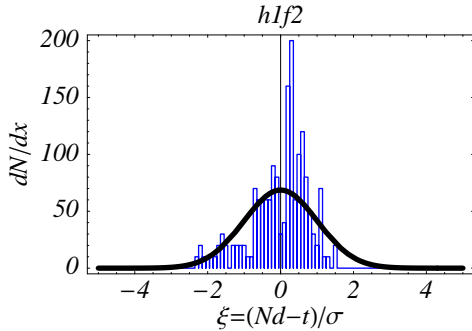


FIG. 3. Histogram of the $\xi \equiv (Nd - t)/\sigma$ for data points from the H1 experiment.

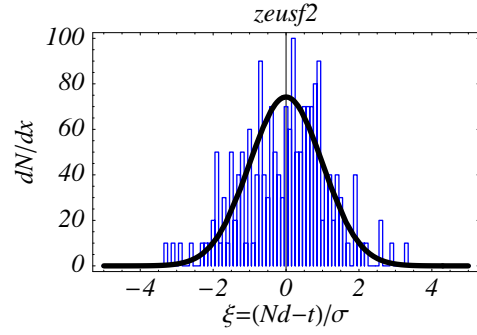


FIG. 4. Histogram of $\xi \equiv (Nd - t)/\sigma$ for data points from the ZEUS experiment.

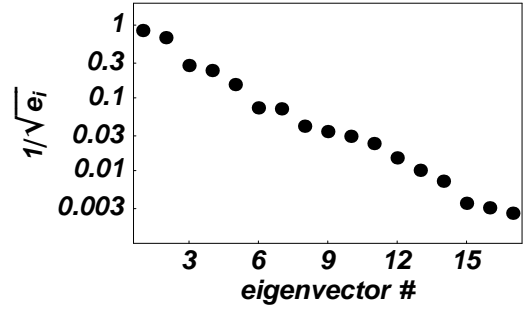


FIG. 5. Plot of the eigenvalues of the Hessian. The vertical axis is $\ell_i = 1/\sqrt{e_i}$.

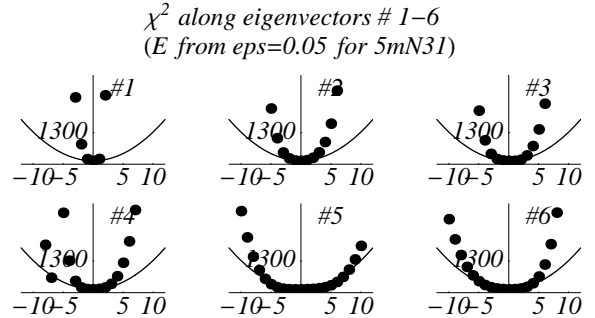


FIG. 6. Value of χ^2 along the six eigenvectors with the largest eigenvalues.

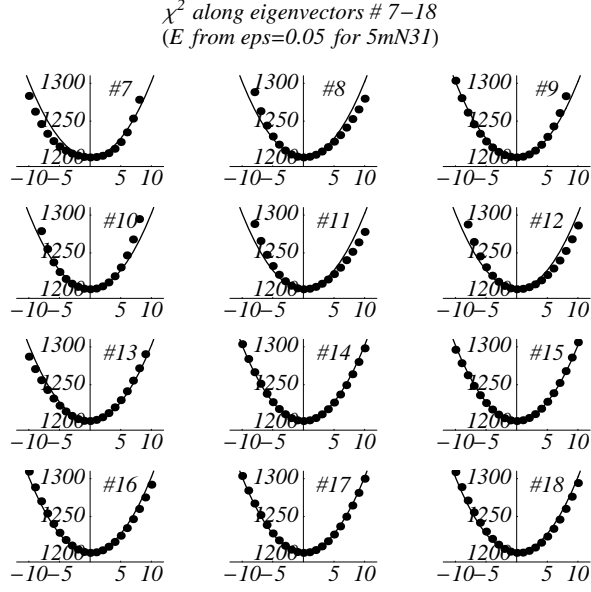


FIG. 7. Value of χ^2 along the 12 eigenvectors with the smallest eigenvalues.

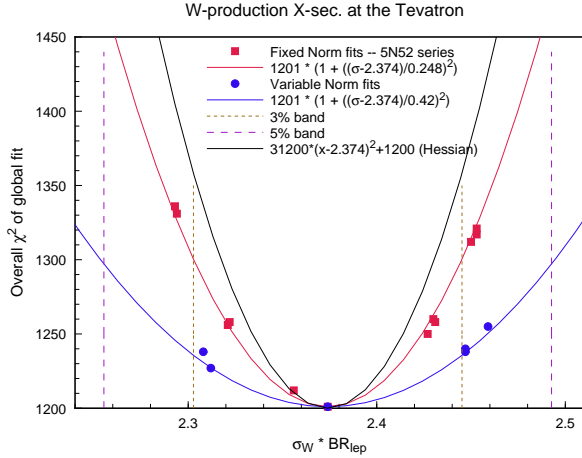


FIG. 8. χ^2 of the base experimental data sets versus $\sigma_W \cdot BR_{lep}$, the W production cross-section at the Tevatron times lepton branching ratio, in nb.

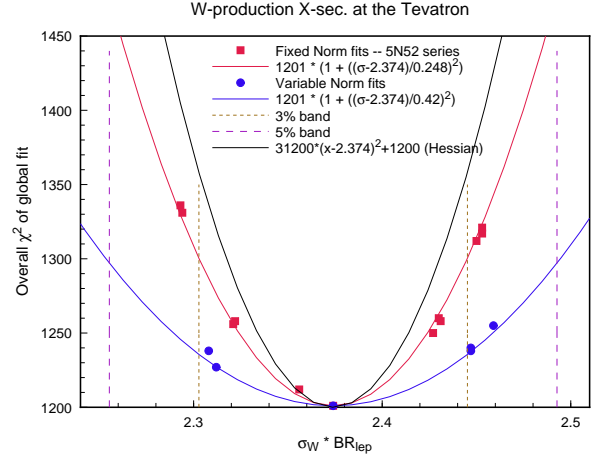


FIG. 9. χ^2 of the base experimental data sets versus $\sigma_W \cdot BR_{lep}$, the W production cross-section at the LHC times lepton branching ratio, in nb.

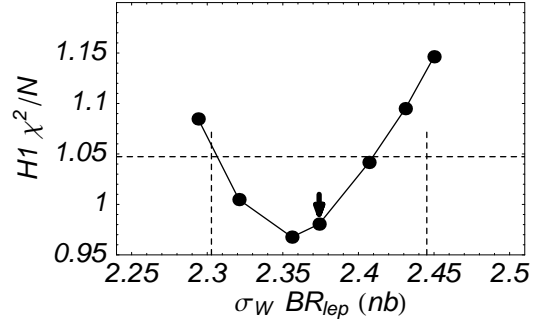


FIG. 10. χ^2/N of the H1 data, including error correlations, compared to PDFs obtained by the Lagrange multiplier method for constrained values of σ_W .

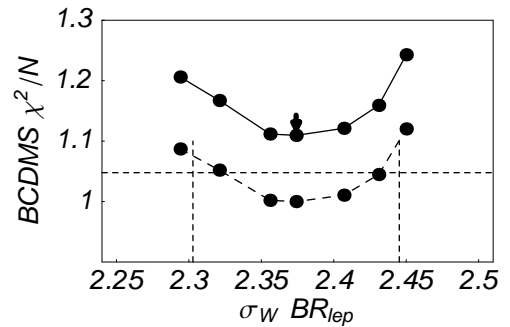


FIG. 11. χ^2/N of the BCDMS data, including error correlations, compared to PDFs obtained by the Lagrange multiplier method for constrained values of σ_W . The points connected by solid lines are the values of χ^2/N . The points connected by dashed lines are the ratios $(\chi^2/N)/(\chi^2/N)_{CTEQ5M}$.

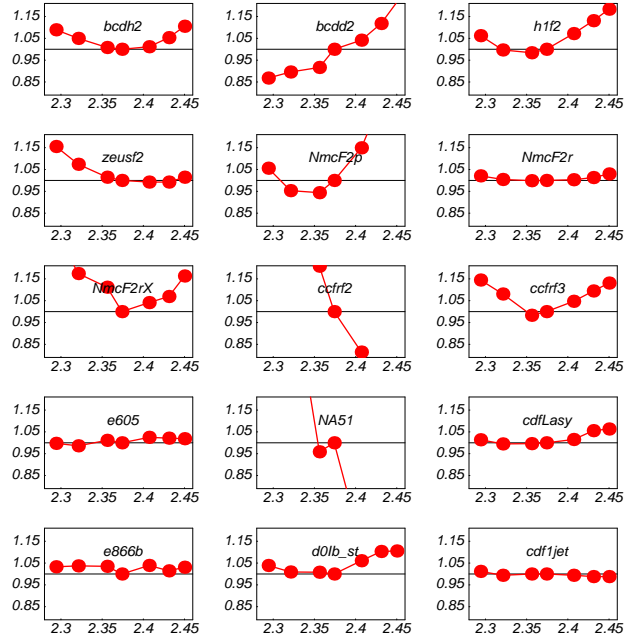


FIG. 12. The abscissa is $\sigma_W BR_{lep}$ in nb. The ordinate is $\chi_u^2 / \chi_{CTEQ5M}^2$.