Uncertainties of predictions from parton distribution functions.
I.   The Lagrange multiplier method

D. Stump, J. Pumplin, R. Brock, D. Casey, J. Huston, and J. Kalk
Department of Physics and Astronomy, Michigan State University, East Lansing, Michigan 48824

H. L. Lai
Department of Physics and Astronomy, Michigan State University, East Lansing, Michigan 48824
and Ming-Hsin Institute of Technology, Hsin-Chu, Taiwan

W. K. Tung
Department of Physics and Astronomy, Michigan State University, East Lansing, Michigan 48824
and Theory Division, CERN, Geneva, Switzerland

We apply the Lagrange multiplier method to study the uncertainties of physical predictions due to the uncertainties of parton distribution functions (PDF's), using the cross section σ_W for W production at a hadron collider as an archetypal example. An effective χ^2 function based on the CTEQ global QCD analysis is used to generate a series of PDF's, each of which represents the best fit to the global data for some specified value of σ_W. By analyzing the likelihood of these ''alterative hypotheses,'' using available information on errors from the individual experiments, we estimate that the fractional uncertainty of σ_W due to current experimental input to the PDF analysis is approximately $\pm 4\%$ at the Fermilab Tevatron, and ± 8–10% at the CERN Large Hadron Collider. We give sets of PDF's corresponding to these up and down variations of σ_W. We also present similar results on Z production at the colliders. Our method can be applied to any combination of physical variables in precision QCD phenomenology, and it can be used to generate benchmarks for testing the accuracy of approximate methods based on the error matrix.

I. INTRODUCTION

All calculations of high-energy processes with initial hadrons, whether within the standard model (SM) or exploring new physics, require parton distribution functions (PDF's) as an essential input. The reliability of these calculations, which underpins both future theoretical and experimental progress, depends on understanding the uncertainties of the PDF's. The assessment of PDF uncertainties has, therefore, become an important challenge to high-energy physics in recent years.

The PDF's are derived from global analyses of experimental data from a wide range of hard processes in the framework of perturbative quantum chromodynamics (PQCD). Quantifying the uncertainties in a global QCD analysis is far from being a straightforward exercise in statistics. There are non-Gaussian sources of uncertainty from perturbation theory (e.g., higher-order and power-law corrections) from choices of parametrization of the nonperturbative input (i.e., initial parton distributions at a low energy scale) from uncertain nuclear corrections to experiments performed on nuclear targets, and from normal experimental statistical and systematic errors. These sources of error need to be studied individually, and eventually combined in a systematic way.

We shall be concerned in this paper with uncertainties of PQCD predictions due to uncertainties of PDF's arising from experimental measurement errors. This problem is considerably more complicated than it appears on the surface. The reason is that in a *global analysis* the large number of data points (~ 1300 in our case) do not come from a uniform set of measurements, but consist of a collection of measurements from many experiments (~ 15) on a variety of physical processes (~ 5–6) with diverse characteristics, precision, and error determination. The difficulty is compounded by a large number of fitting parameters (~ 16) which are not uniquely specified by the theory. Several approaches to this problem have been proposed, with rather different emphases on the rigor of the statistical method, scope of experimental input, and attention to various practical complications [1–7]. Our group has initiated one of these efforts, with the emphasis on utilizing the full constraints of the global data [7]. This work has motivated a closer examination of the standard techniques of error analysis, and necessary improvements and extensions to these techniques, as applied to a complex real world problem such as global QCD analysis of PDF's [8].

In this paper we present a detailed analysis of uncertainties of physical observables due to parton distribution functions, using the Lagrange multiplier method proposed in [7,8]. This method explores the entire multidimensional parton parameter space, using an effective χ^2 function that conveniently combines the global experimental, theoretical, and phenomenological inputs to give a quantitative measure of the goodness-of-fit for a given set of PDF parameters (cf. Sec. II). The method probes directly the variation of the effective χ^2 along a specific direction in the PDF parameter space—that of maximum variation of a specified physical variable. The result is a robust set of *optimized sample PDF's* (or ''alternative hypotheses'') from which the uncertainty of

TABLE I. List of data sets used in the global analysis.

Experiment	Process	Label	No. data pts.	Reference
BCDMS	DIS μp	BCDMSp	168	[14]
BCDMS	DIS μd	BCDMSd	156	[14]
H1	DIS ep	H1	172	[15]
ZEUS	DIS ep	ZEUS	186	[16]
NMC	DIS μp	NMCp	104	[17]
NMC	DIS $\mu p/\mu n$	NMCr	123	[17]
NMC	DIS $\mu p/\mu n$	NMCrx	13	[17]
CCFR	DIS νp	CCFR2	87	[18]
CCFR	DIS νp	CCFR3	87	[18]
E605	DY pp	E605	119	[19]
NA51	DY pd/pp	NA51	1	[20]
E866	DY pd/pp	E866	11	[21]
CDF	$W_{lep-asym.}$	CDFw	11	[22]
D0	$\bar{p}p \rightarrow jetX$	D0jet	24	[23]
CDF	$\bar{p}p \rightarrow jetX$	CDFjet	33	[24]

the physical variable can be assessed quantitatively without the approximations inherent in the traditional error matrix approach. For concreteness, we consider the cross section σ_W of W boson production at the Tevatron as the archetypal example (cf. Sec. III).

The definition of the effective χ^2 function, and the inputs that go into it, do not permit a direct statistical interpretation of its numerical value. To obtain meaningful confidence levels for the optimized sample PDF sets, it is necessary to conduct a series of likelihood analyses of these sample PDF's, using all available information on errors for the individual experiments. The results from these analyses serve as the basis to assign an overall uncertainty range on the physical variable, and a corresponding tolerance measure for the effective χ^2 function used in the analysis, that are consistent with the experiments used in the current global QCD analysis (cf. Sec. IV).

This method can be applied to any physical variable, or combination of physical variables, in precision QCD phenomenology. In Sec. V we present results on W production at the Large Hadron Collider (LHC), and Z production at the Fermilab Tevatron and the CERN LHC. We compare the uncertainties obtained in all cases, and comment on previous estimates in the context of these results. In Sec. VI we present parton distribution sets that are optimized to give high and low values of the W and Z cross sections, while remaining consistent with current experiments according to our analysis.

The Lagrange multiplier method provides a useful tool to test the reliability of the more traditional method of error propagation via the error matrix [1,4,9], which relies on the quadratic expansion of the χ^2 function around its minimum. In a companion paper [10] we perform an in-depth analysis of the uncertainties of the PDF's in the error matrix approach, using the much improved numerical method for calculating the Hessian that was developed in [8]. There we demonstrate how the more specialized Lagrange multiplier method can set useful benchmarks for the general purpose error matrix approach.

II. THE GLOBAL QCD ANALYSIS

We adopt the same experimental and theoretical input as the CTEQ5 analysis [11]: 15 data sets from 11 experiments on neutral-current and charged-current deep inelastic-scattering (DIS), lepton-pair production [Drell-Yan (DY)], lepton asymmetry in W production, and high p_T inclusive jet production processes are used (cf. Table I). The total number of data points is $N = 1295$. We denote the experimental data values by $\{D\} = \{D_I; I = 1, \ldots, N\}$. The theory input is next-leading-order (NLO) PQCD, and the theory value for the data point I will be denoted by T_I. The theory depends on a set of parameters $\{a\} \equiv \{a_i; i = 1, \ldots, d\}$. These parameters characterize the nonperturbative QCD input to the analysis; they determine the initial PDF's $\{f(x, Q_0; \{a\})\}$ defined at a low-energy scale Q_0, below the energy scale of the data, which we choose to be $Q_0 = 1$ GeV. When we need to emphasize that the theoretical values depend on the PDF parameters, we write $T_I(a)$ to indicate the dependence on $\{a\}$.

The parametrization of $\{f(x, Q_0)\}$ is somewhat arbitrary, motivated by physics, numerical considerations, and economy. Another parametrization might be employed, and differences among the possible parametrizations are, in principle, a source of theoretical uncertainty in their own right. For most of this study we focus on a single parametrization, but we comment on the effect of changing the parametrization at the end of Sec. IV. The number d of the parameters $\{a\}$ is chosen to be commensurate with current experimental constraints. For this study we use $d = 16$. The detailed forms adopted for the initial functions $\{f(x, Q_0; \{a\})\}$ are not of particular concern in this study, since we shall be emphasizing results obtained by ranging over the full parameter

space.[1] The explicit formulas are given in Appendix C (where relevant PDF's from the results of our study are presented). The $T_I(\{a\})$ are calculated as convolution integrals of the relevant NLO QCD matrix elements and the universal parton distributions $\{f(x,Q;\{a\})\}$ for all Q. The latter are obtained from the initial functions $\{f(x,Q_0;\{a\})\}$ by NLO QCD evolution.

The *global analysis* consists of a systematic way to determine the best values for the $\{a\}$, and the associated uncertainties, by fitting $\{T(a)\}$ to $\{D\}$. Because of the wide range of experimental and theoretical sources of uncertainty mentioned in the Introduction, there are a variety of strategies to deal with the complex issues involved [1–4,7]. In the next two sections, the primary tool we employ is conventional χ^2 analysis. The important task is to define an effective χ^2 function, called $\chi^2_{\text{global}}(a)$, that conveniently combines the theoretical and global experimental inputs, as well as relevant physics considerations based on prior knowledge, to give an overall measure of the goodness-of-fit for a given set of PDF parameters.

Experience in global analysis of PDF's during the past two decades has demonstrated that the PDF's obtained by the minimization of such a suitably chosen χ^2_{global} provide very useful up-to-date hadron structure functions which, although not unique, are representative of good fits between theory and experiments. Now we must quantify the uncertainties of the PDF's and their predictions; i.e., we must expand the scope of the work from merely identifying typical solutions to systematically mapping the PDF parameter space in the neighborhood around the minimum of χ^2.

The simplest possible choice for the χ^2 function would be

$$\chi^2(a) = \sum_{I=1}^{N} \frac{[D_I - T_I(a)]^2}{\sigma_I^2}, \tag{1}$$

where σ_I is the error associated with data point I. Through $T_I(a)$, $\chi^2(a)$ is a function of the theory parameters $\{a\}$. Minimization of $\chi^2(a)$ would identify parameter values for which the theory fits the data. However, the simple form (1) is appropriate only for the ideal case of a uniform data set with uncorrelated errors. For data used in the global analysis, most experiments combine various systematic errors into one effective error for each data point, along with the statistical error. Then, in addition, the fully correlated normalization error of the experiment is usually specified separately. For this reason, it is natural to adopt the following definition for the effective χ^2 (as done in previous CTEQ analyses):

$$\chi^2_{\text{global}}(a) = \sum_{n} w_n \chi^2_n(a)$$

$$(n \text{ labels the different experiments}), \tag{2}$$

$$\chi^2_n(a) = \left(\frac{1 - \mathcal{N}_n}{\sigma_n^N}\right)^2 + \sum_I \left(\frac{\mathcal{N}_n D_{nI} - T_{nI}(a)}{\sigma_{nI}^D}\right)^2. \tag{3}$$

For the nth experiment, D_{nI}, σ_{nI}^D, and $T_{nI}(a)$ denote the data value, measurement uncertainty (statistical and systematic combined) and theoretical value (dependent on $\{a\}$) for the Ith data point, σ_n^N is the experimental normalization uncertainty, and $\mathcal{N}_n$ is an overall normalization factor (with default value 1) for the data of experiment n. The factor w_n is a possible weighting factor (with default value 1) which may be necessary to take into account prior knowledge based on physics considerations or other information. The *a priori* choices represented by the w_n values are present, explicitly or implicitly, in any data analysis. For instance, data inclusion or omission (choices which vary for different global analysis efforts) represent extreme cases, assigning either 100% or 0% weight to each available experimental data set. Similarly, choices of various elements of the analysis procedure itself represent subjective input. Subjectivity of this kind also enters into the analysis of systematic errors in experiments.

The function $\chi^2_{\text{global}}(a)$ allows the inclusion of all experimental constraints in a uniform manner while allowing flexibility for incorporating other relevant physics input. We will make use of this function to explore the neighborhood of the best fit, and to generate sample PDF's pertinent to the uncertainty of the prediction of a specific physical variable of interest. However, the numerical value of this effective χ^2 function should not be given an *a priori* statistical interpretation, because correlations between measurement errors, and correlated theoretical errors, are not included in its definition. In particular, the likelihood of a candidate PDF set $\{a\}$ cannot be determined by the value of the increase $\Delta\chi^2_{\text{global}}(a)$ above the minimum.[2] Instead, the evaluation of likelihoods and estimation of global uncertainty will be carried out in a separate step in Sec. IV, after sets of optimal sample PDF's for the physical variable of interest have been obtained.

III. THE LAGRANGE MULTIPLIER METHOD

The Lagrange multiplier method is an extension of the χ^2 minimization procedure, that relates the range of variation of a physical observable X dependent upon the PDF's, to the variation of the function $\chi^2_{\text{global}}(a)$ that is used to judge the goodness of fit of the PDF's to the experimental data and PQCD.

---

[1]In other words, for this paper, the PDF parameters $\{a\}$ play mostly the role of ''internal variables.'' In contrast, they occupy the center stage in the companion paper [10].

[2]The often quoted theorem of Gaussian error analysis, that an increase of χ^2 by one unit in a constrained fit to data corresponds to one standard deviation of the constrained variable, is true only in the absence of correlations. When existing correlations are left out, the relevant size of $\Delta\chi^2$ can be much larger than 1. Appendix A discusses this point in some detail.
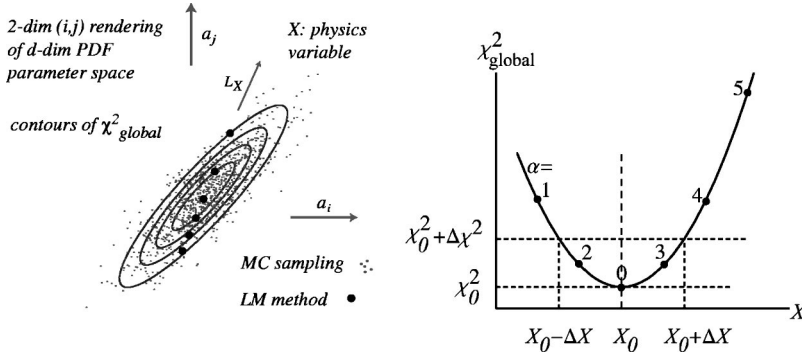
FIG. 1. Left: The LM method provides sample points along a single curve L_X in the multidimensional PDF parameter space, relevant to the observable X. Right: For a given tolerance $\Delta\chi^2_{\text{global}}$, the uncertainty in the calculated value of X is $\pm\Delta X$. The solid points correspond to the sample points on the curve L_X in the left plot.

A. The method

The method has been introduced in [7,8]. The starting point is to perform a global analysis as described in Sec. II, by minimizing the function $\chi^2_{\text{global}}(a)$ defined by Eq. (2), thus generating a set of PDF's that represents the best estimate consistent with current experiment and theory. We call this set the "standard set,"[3] denoted S_0. The parameter values that characterize this set will be denoted by $\{a^{(0)}\}\equiv\{a_i^{(0)};i=1,\ldots,d\}$, and the absolute minimum of χ^2_{global} will be denoted by χ_0^2. Now, let X be a particular physical quantity of interest. It depends on the PDF's, $X=X(a)$, and the best estimate (or prediction) of X is $X_0=X(a^{(0)})$. We will assess the *uncertainty* of this predicted value by a two-step analysis. First, we use the Lagrange multiplier method to determine how the minimum of $\chi^2_{\text{global}}(a)$ increases, i.e., how the quality of the fit to the global data set decreases, as X deviates from the best estimate X_0. Second, in Sec. IV, we analyze the appropriate tolerance of χ^2_{global}.

As explained in [7,8], the first step is taken by introducing a Lagrange multiplier variable λ, and minimizing the function

$$\Psi(\lambda,a)=\chi^2_{\text{global}}(a)+\lambda X(a) \tag{4}$$

with respect to the original d parameters $\{a\}$ for fixed values of λ. In practice we minimize $\Psi(\lambda,a)$ for many values of the Lagrange multiplier $\lambda_1,\lambda_2,\ldots,\lambda_M$. For each specific value λ_α, the minimum of $\Psi(\lambda_\alpha,a)$ yields a set of parameters $\{a_{\min}(\lambda_\alpha)\}$, for which we evaluate the observable X and the related χ^2_{global}. We use the shorthand $(X_\alpha,\chi^2_{\text{global},\alpha})$ for this pair. $\chi^2_{\text{global},\alpha}$ represents the lowest achievable χ^2_{global}, for the global data, for which X has the value X_α, taking into account all possible PDF's in the *full d-dimensional parameter space* of points $\{a\}$. In other words, the result $\{a_{\min}(\lambda_\alpha)\}$ is a *constrained fit*—with X constrained to be X_α. We can equivalently say that X_α is an extremum of X if χ^2_{global} is constrained to be $\chi^2_{\text{global},\alpha}$. We denote the resulting set of PDF's by S_α.

We repeat the calculation for many values of λ, following the chain

$$\lambda_\alpha\rightarrow\min[\Psi(\lambda_\alpha,a)]\rightarrow a_{\min}(\lambda_\alpha)\rightarrow X_\alpha \quad\text{and}\quad \chi^2_{\text{global},\alpha}$$

for $\alpha=1,2,3,\ldots,M$. The result is a parametric relationship between X and χ^2_{global}, through λ. We call this function $\chi^2_{\text{global}}(X)$; so $\chi^2_{\text{global}}(X_\alpha)=\chi^2_{\text{global},\alpha}$ is the minimum of $\chi^2_{\text{global}}(a)$ when X is constrained to be X_α. The absolute minimum of χ^2_{global}, which we denote χ_0^2, is the minimum of $\Psi(\lambda=0,a)$, occurring at $\{a\}=\{a^{(0)}\}$. Thus the procedure generates a set of optimized sample PDF's along the curve of maximum variation of the physical variable X in the d-dimensional PDF parameter space (with $d=16$ in our case). These PDF sets $\{S_\alpha\}$ are exactly what is needed to assess the range of variation of X allowed by the data. In other words, the Lagrange multiplier method provides optimal PDF's tailored to the physics problem at hand, in contrast to an alternative method [3] that generates a large sample of PDF's by the Monte Carlo method. The underlying ideas of these two complementary approaches are illustrated in the plot on the left side of Fig. 1.

$\chi^2_{\text{global}}(X)$ is the lowest achievable value of $\chi^2_{\text{global}}(a)$ for the value X of the observable, where $\chi^2_{\text{global}}(a)$ represents our measure of the goodness-of-fit to the global data. Therefore, the allowed range of X, say from $X_0-\Delta X$ to $X_0+\Delta X$, corresponding to a chosen tolerance of the goodness of fit $\Delta\chi^2_{\text{global}}=\chi^2_{\text{global}}-\chi_0^2$, can be determined by examining a graph of χ^2_{global} versus X, as illustrated in the plot on the right-hand side of Fig. 1. This method for calculating ΔX may be more robust and reliable than the traditional error propagation because it does not approximate $X(a)$ and $\chi^2_{\text{global}}(a)$ by linear and quadratic dependence on $\{a\}$, respectively, around the minimum.

Although the parameters $\{a\}$ do not appear explicitly in this analysis, the results do depend, in principle, on the choice of parameter space (including the dimension, d) in which the minimization takes place. In practice, if the degrees of freedom represented by the parametrization are chosen to match the constraining power of the global data sets used, which must be true for a sensible global analysis, the results are quite stable with respect to changes in the parametrization choices. The sensitivity to these choices is tested, as part of the continuing effort to improve the global analysis.

The discussion so far has left open this question: What is the appropriate tolerance $\Delta\chi^2_{\text{global}}$ to define the "error" of the prediction X_0? This question will be addressed in Sec. IV.

_____

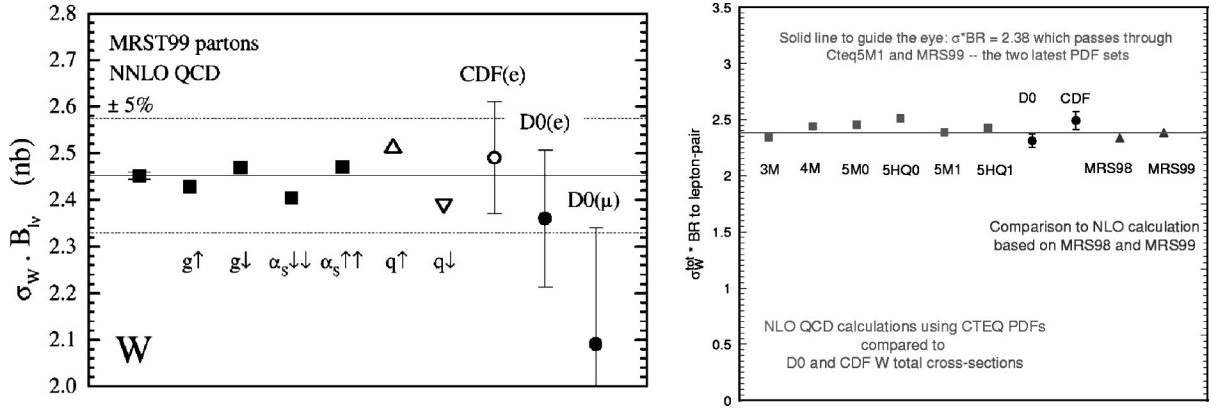[3]This standard set is very similar to the published CTEQ5M1 set [11].

FIG. 2. Calculated cross section for $W^{\pm}$ boson production (multiplied by the branching ratio for $W^{-} \to e \bar{\nu}$) at the Tevatron, for various current and historical PDF's. The two plots are from Refs. [13] and [7], respectively.

Our method can obviously be generalized to study the uncertainties of a collection of physical observables $(X_1, X_2, \ldots, X_s)$ by introducing a separate Lagrange multiplier for each observable. Although the principle stays the same, the amount of computational work increases dramatically with each additional observable.

B. A case study: The W cross section

In this section we examine the cross section σ_W for inclusive $W^{\pm}$ production at the Tevatron ($p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV) to illustrate the method and to lay the ground work for the quantitative study of uncertainties to be given in Sec. IV. Other examples will be described in Sec. V. Preliminary results of this section have been reported previously [7,8].

Until recently the only method for assessing the uncertainty of σ_W due to PDF's has been to compare the calculated values obtained from a number of different PDF's, as illustrated in Fig. 2, in which the plots are taken from existing literature.[4] The PDF's used in these comparisons are either the "best fits" from different global analysis groups [11,12] (hence are not pertinent to uncertainty studies) or are chosen by some simple intuitive criteria [13]. The meaning and reliability of the resulting range of σ_W are not at all clear. Furthermore, these results do not provide any leads on how the uncertainties can be improved in the future. The Lagrange multiplier technique provides a systematic method to address and remedy both of these problems.

Let the physical quantity X of the last section be the cross section σ_W for $W^{\pm}$ production at the Tevatron. Applying the Lagrange method, we obtain the constrained minimum of χ^2_{global} as a function of σ_W, shown as solid points in Fig. 3. The best estimate value, i.e., the prediction for the standard set S_0, is $\sigma_{W0} = 21.75$ nb. The curve is a polynomial fit to the points to provide a smooth representation of the continu-

ous function $\chi^2_{\text{global}}(X)$. We see that all the sample PDF sets obtained by this method lie on a smooth quasiparabolic curve with the best-fit value at the minimum.

As discussed earlier (in Fig. 1) points on the curve represent our sample of optimal PDF's relevant to the determination of the uncertainty of σ_W. To quantify this uncertainty, we need to reach beyond the effective χ_{global} function, and establish the confidence levels for these "alternative hypotheses" with respect to the experimental data sets used in the global analysis.

IV. QUANTIFYING THE UNCERTAINTY

Consider a series of sample PDF sets along the curve $\chi^2_{\text{global}}(X)$ of Fig. 3 denoted by $\{S_\alpha; \alpha = 0, 1, \ldots, M\}$ where S_0 is the standard set. These represent "alternative hypotheses" for the true PDF's, and we wish to evaluate the likelihoods associated with these alternatives. To do so, we go back to the individual experiments and, in each case, perform as detailed a statistical analysis as is permitted with available information from that experiment. After we have obtained meaningful estimates of the "errors" of these candidate PDF's with respect to the individual experiments, we shall try to combine this information into a global uncertainty measure in the form of ΔX and $\Delta \chi^2_{\text{global}}$.
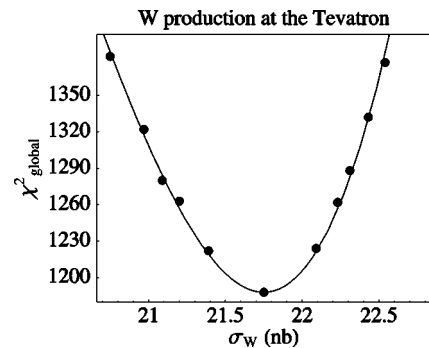

FIG. 3. Minimum χ^2_{global} versus σ_W, the inclusive $W^{\pm}$ production cross section at the Tevatron ($\bar{p}p$ collisions at $\sqrt{s} = 1.8$ TeV) in nb. The points were obtained by the Lagrange multiplier method. The curve is a polynomial fit to the points.

---

[4]These plots show the product of σ_W times a leptonic branching ratio, which is what is measured experimentally. The branching ratio B has some experimental error. For studying the uncertainties of σ_W, we will focus on σ_W itself in the rest of the paper.

The experimental data sets included in our global analysis are listed in Table I. For some of these experiments, information on correlated systematic errors is available (albeit usually in unpublished form). For these, statistical inference should be drawn from a more accurate χ_n^2 function than the simple formula Eq. (3) used for the global fit. In particular, if σ_{nI} is the uncorrelated error and $\{\beta_{kI};k=1,2,\ldots,K\}$ are the coefficients of K distinct correlated errors associated with the data point I, then an appropriate formula for the χ_n^2 function is

$$\chi_n^2 = \sum_I \frac{(D_{nI}-T_{nI})^2}{\sigma_{nI}^2} - \sum_{k=1}^{K} \sum_{k'=1}^{K} B_k (A^{-1})_{kk'} B_{k'}, \quad (5)$$

where B_k is a vector, and $A_{kk'}$ a matrix, in K dimensions:

$$B_k = \sum_I \beta_{kI}(D_{nI}-T_{nI})/\sigma_{nI}^2;$$

$$A_{kk'} = \delta_{kk'} + \sum_I \beta_{kI}\beta_{k'I}/\sigma_{nI}^2. \quad (6)$$

(The sum over I here includes only the data from experiment n.) Traditionally, χ_n^2 is written in other ways, e.g., in terms of the inverse of the $(N \times N)$ variance matrix. For experiments with many data points, the inversion of such large matrices may lead to numerical instabilities, in addition to being time consuming. Our formula (5) has a significant advantage in that all the systematic errors are first combined ("analytically") in the definitions of B_k and $A_{kk'}$. Equation (5) requires only the inverse of the much smaller $(K \times K)$ matrix $A_{kk'}$. (K is the number of distinct systematic errors.) The derivation of these formulas is given in Appendix B. Equation (5) reduces to the minimum of χ_n^2 in Eq. (3) with respect to $\mathcal{N}_n$ if the only correlated error is the overall normalization error for the entire data set; in that case $\beta_I = -\sigma_n^N D_{nI}$.

By using Eq. (5) or Eq. (3) for cases where the correlations of systematic errors are unavailable, we obtain the best estimate on the range of uncertainty permitted by available information on each individual experiment. We should note that the experimental data sets are continuously evolving. Some data sets in Table I will soon be updated (Zeus, H1) or replaced (CCFR).[5] In addition, most information on correlated systematic errors is either unpublished or preliminary. The results presented in the following analysis should, therefore, be considered more as a demonstration of principle—as the first application of our proposed method—rather than the final word on the PDF uncertainty of the W cross section.

A. Uncertainty with respect to individual experiments

As an example, we begin by comparing the $\{S_\alpha\}$ series for σ_W at the Tevatron to the H1 data set [15]. Results on
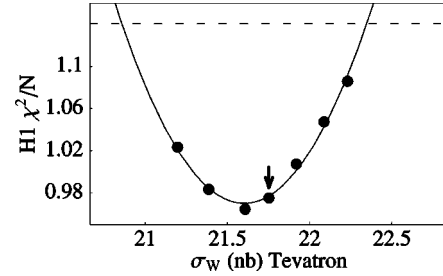
FIG. 4. χ^2/N of the H1 data, including error correlations, for sample PDF's obtained by the Lagrange multiplier method for constrained values of σ_W at the Tevatron. The arrow indicates the global minimum.

correlated systematic errors are available for this data set,[6] and are incorporated in the calculation using Eq. (5). The number of data points in this set is $N_{H1}=172$. The calculated values of χ_{H1}^2/N_{H1} are plotted against σ_W in Fig. 4. The curve is a smooth interpolation of the points. The value of χ_{H1}^2/N_{H1} for the standard set S_0 (indicated by a short arrow on the plot) is 0.975, and it is 0.970 at the minimum of the curve. These values are quite normal for data with accurately determined measurement errors. We can, therefore, apply standard statistics to calculate the 90% confidence level on χ^2/N for $N=172$. The result is shown as the dashed horizontal line in Fig. 4.

We have similarly calculated χ_n^2/N_n including information on the correlations of systematic errors for the BCDMSp data set. The results are similar to the H1 results, except that the absolute values are all larger than 1.12, a large value for $N=168$ data points. This is a familiar problem in data analysis, and it is encountered in several other data sets in this global analysis (cf. below). The χ_n^2/N_n calculation including correlations of the errors, is also done for the D0 and Collider Detector at Fermilab (CDF) jet cross sections.[7] For those experiments that have only provided (effective) uncorrelated errors, we must rely on Eq. (3) for our error calculation, since that represents the best information available.

In order to obtain usable likelihood estimates from all the data sets, one must address the problem mentioned in the previous paragraph: Even in a "best fit," the values of χ^2 per data point, χ_n^2/N_n, for individual experiments vary considerably among the established experiments (labeled by n). Specifically, χ_n^2/N_n ranges from $1.5-1.7$ (for ZEUS and CDF-jet) on the high end to $0.5-0.7$ (for some DY experiments) on the low end in all good fits. Considering the fact that some of these data sets contain close to 200 points, the range of variation is huge from the viewpoint of normal statistics;

---

[5]Compare talks presented by these collaborations at DIS2000 *Workshop on Deep Inelastic Scattering and Related Topics*, Liverpool, England, 2000.

[6]These systematic errors are unpublished results, but are made available to the public on the H1 Web page. For convenience, we have approximated each of the pair of four nonsymmetrical errors by a single symmetric error. The size of the resulting error on σ_W, inferred from this evaluation, is not affected by that approximation.

[7]The measurement errors of the jet cross sections are dominated by systematic errors, so the error correlation matrices are used for χ_n^2 of these experiments even in χ_{global}^2.

experiments with χ_n^2/N_n deviating from 1.0 by a few times $\sqrt{2/N_n}$ in either direction would have to be ruled out as extremely unlikely [25].

The reasons for χ_n^2/N_n to deviate from 1.0 in real experiments are complex, and vary among experiments. They are, almost by definition, not understood, since otherwise the errors would have been corrected and the resulting χ^2 would become consistent with the expected statistical value. Under these circumstances, a commonly adopted pragmatic approach is to focus on the relative χ^2 values with respect to some appropriate reference χ^2.[8] Accordingly, in the context of performing *global* QCD analysis, we adopt the following procedure. For each experiment (labeled by n):

(i) Let $\chi_{n,0}^2$ denote the value of χ_n^2 for the standard set S_0. We assume S_0 is a viable reference set. Because $\chi_{n,0}^2$ may be far from a likely value for random errors, we *rescale* the values of $\chi_{n,\alpha}^2$ (for $\alpha=0,1,2,\ldots,M$) by a factor C_{n0}, calling the result $\bar{\chi}_{n,\alpha}^2$

$$\chi_{n,\alpha}^2 \rightarrow \bar{\chi}_{n,\alpha}^2 \equiv C_{n0}\chi_{n,\alpha}^2. \tag{7}$$

The constant C_{n0} is chosen such that, for the standard set, $\bar{\chi}_{n,0}^2$ assumes the most probable value for a chi-squared variable: $\bar{\chi}_{n,0}^2 = \xi_{50} \equiv$ the 50th percentile of the chi-squared distribution $P(\chi^2,N_n)$ with N_n degrees of freedom, defined by

$$\int_0^{\xi_{50}} P(\chi^2,N_n)d\chi^2 = 0.50. \tag{8}$$

(If N_n is large then $\xi_{50} \approx N_n$.) The rescaling constant C_{n0} is thus $\xi_{50}/\chi_{n,0}^2$. For random errors the probability that $\chi^2 < \xi_{50}$ (or $> \xi_{50}$) is 50%. For those experiments whose $\chi_{n,0}^2$ deviates significantly from ξ_{50}, this rescaling procedure is meant to provide a simple (but crude) way to correct for the unknown correlations or unusual fluctuations.

(ii) We then examine the values of $\bar{\chi}_{n,\alpha}^2$ for the alternative sets S_α with $\alpha=1,2,\ldots,M$, using $\bar{\chi}_{n,\alpha}^2 - \bar{\chi}_{n,0}^2$ to compute the statistical likelihood of the alternative hypothesis S_α with respect to the data set n, based on the chi-squared distribution with N_n data points.

This procedure does not affect the results presented earlier for the H1 experiment, since $\chi_{n,0}^2/N_n$ is already very close to 1 for that experiment.

Before presenting the results of the likelihood calculation, it is interesting to examine, in Fig. 5, the differences $\Delta\chi_{n,\alpha}^2 = \chi_{n,\alpha}^2 - \chi_{n,0}^2$ (before rescaling) versus σ_W for the 15 data sets. (N.B. The vertical scales of the various plots are not the

---

[8]The alternative is to take the *absolute* values of χ_n^2 seriously, and hence only work with alternative hypotheses and experiments that are both self-consistent (i.e., have $|\chi_n^2/N_n - 1| \lesssim \sqrt{2/N_n}$) and mutually compatible in the strict statistical sense (i.e., have overlapping likelihood functions). Since few of the precision DIS experiments are compatible in this sense, one must then abandon global analysis, and work instead with several distinct (and mutually exclusive) analyses based on different experiments.

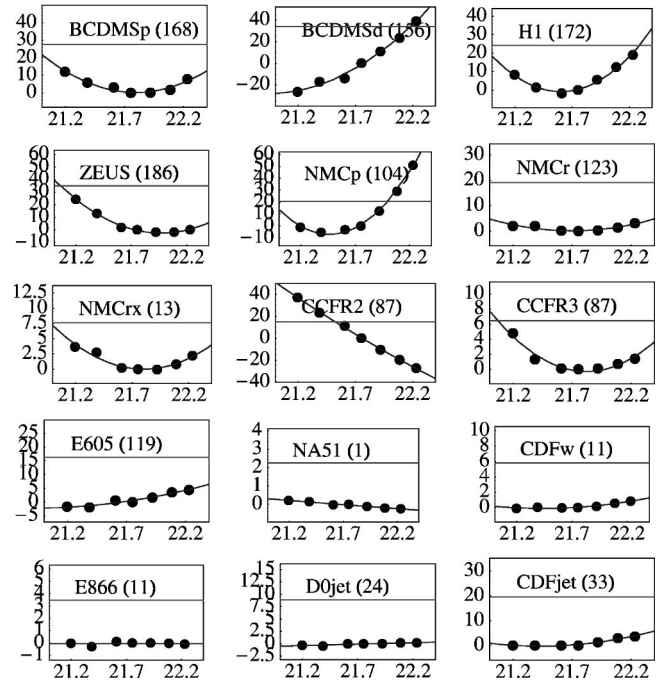$\chi^2-\chi_0^2$ vs σ_W (Tevatron)

FIG. 5. The abscissa is σ_W in nb, at the Tevatron. The ordinate is $\chi_n^2-\chi_{n,0}^2$. The number in parentheses is the number of data points. The horizontal lines are explained in the text.

same, due to the large variations in the value of $\Delta\chi_{n,\alpha}^2$ for different experiments.) The ordering of the experiments in Fig. 5 is the same as in Table I, with experiments ordered by process (DIS, DY, W, and jet production). It is clear from these graphs that the DIS experiments place the strongest constraints on σ_W, because they have the largest $\Delta\chi_n^2$ for the same $\Delta\sigma_W$. This is to be expected since quark-antiquark annihilation makes the dominant contribution to σ_W. We also observe that most experiments place some constraint on σ_W on both sides, but a few bound it on one side only. Globally, as shown in Fig. 3, the combined constraints give rise to a classic parabolic behavior for $\chi_{\text{global}}^2(\sigma_W)$.

To estimate the statistical significance of the individual χ_n^2 increases, we assume that the rescaled variable $\bar{\chi}_n^2$ obeys a chi-squared distribution $P(\chi^2,N_n)$ for N_n data points. Thereby, we estimate the value of $\bar{\chi}_n^2$ that corresponds to the 90% confidence level (C.L.) uncertainty for σ_W (with respect to experiment n) from the formula $\bar{\chi}_n^2 = \xi_{90}$, where ξ_{90} is the 90th percentile defined by

$$\int_0^{\xi_{90}} P(\chi^2,N_n)d\chi^2 = 0.90. \tag{9}$$

For example, Fig. 6 shows the chi-squared distribution $P(\chi^2,N_n)$ for $N_n=172$, the number of data points in the H1 data set. The 50th and 90th percentiles are indicated. We choose a conservative 90% C.L. because there are other theoretical and phenomenological uncertainties not taken into account by this analysis.
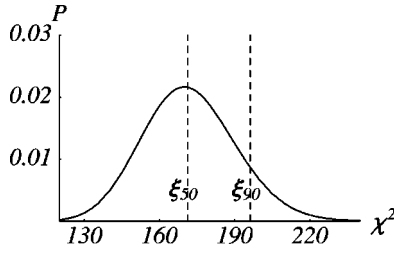
FIG. 6. The chi-square distribution $P(\chi^2, N_n)$ for $N_n = 172$ data points. The dashed lines indicate the 50th and 90th percentiles.

To summarize our procedure, an alternative PDF set S_α lies within the 90% C.L. for experiment n if it has $\bar{\chi}^2_{n,\alpha} < \xi_{90}$; that is, if

$$\frac{\chi^2_{n,\alpha}}{\chi^2_{n,0}} < \frac{\xi_{90}}{\xi_{50}}. \tag{10}$$

We judge the likelihood of S_α from the *ratio* of $\chi^2_{n,\alpha}$ to the reference value $\chi^2_{n,0}$, rather than from the absolute magnitude. The horizontal lines in Fig. 5 correspond to the values of $\Delta\chi^2_n$ obtained in this way. Finally, from the intercepts of the line with the interpolating curve in each plot in Fig. 5, we obtain an estimated uncertainty range of σ_W from each individual experiment. The results are presented collectively in Fig. 7, where, for each experiment, the point ($\bullet$) is the value of σ_W for which χ^2_n is minimum, and the error bar extends across the 90% C.L. based on that data set.

The uncertainty ranges shown in Fig. 7 with respect to individual experiments represent the most definitive results of our study, in the sense that the input and the assumptions can be stated clearly and the analysis is quantitative within the stated framework. It is natural to proceed further and estimate a global measure of $\Delta\sigma_W$ and the corresponding $\Delta\chi^2_{\text{global}}$. This last step is, however, less well defined and requires some subjective judgment.

B. The global uncertainty

It should be emphasized that the ranges shown by the error bars in Fig. 7 are not errors determined independently
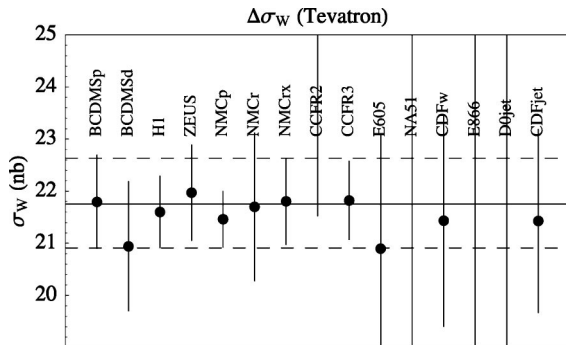

FIG. 7. Ranges of σ_W within the 90% C.L. for the individual experiments. The ordinate is σ_W for the Tevatron process $p\bar{p} \rightarrow W^\pm X$. The solid line is the best estimate according to the standard PDF set S_0. The dashed lines are the bounds described in Sec. IV B.

by each experiment; rather they represent the ranges allowed by the experiments for alternative *global* fits $\{S_\alpha\}$. For this reason, and others related to the rescaling of χ^2 mentioned earlier, as well as approximations inherent in many of the original published errors,[9] it is not obvious how to *combine* these errors. We refer to the ranges in Fig. 7 by the generic term *local* (i.e., single-experiment) *uncertainties*. On a qualitative level, Fig. 7 exhibits the same features seen earlier in Fig. 5: (i) the quark dominated DIS experiments give the smallest error bars, and (ii) a few experiments only set bounds on one side, while the rest limit the range in both directions. In addition, Fig. 7 gives us an overall view which clearly shows that σ_W is well constrained in the global analysis, and the experimental bounds are consistent with each other.

The important question is how to provide a sensible measure of the overall uncertainty in view of the complexity of the problem already described. The situation here is not unlike the problem of assigning an overall systematic error to an experimental measurement. Figure 7 shows a set of 90% C.L. ranges for σ_W from different sources, but these ranges are highly correlated, because the alternative hypotheses being tested come from global fits. The final uncertainty must be a reasonable intersection of these ranges.

We will state an algorithm for obtaining the final uncertainty measure of σ_W based on Fig. 7. The same algorithm can be applied in the future for predictions of other observables. It has the following two parts: (1) determine the central value using all the experiments, that is the solid line in Fig. 7; (2) take the *intersection* of the error ranges as the combined uncertainty. But in calculating the intersection, experiments below the mean are used only for setting the lower bound, and experiments above the mean are used only for setting the upper bound. With this algorithm, experiments that permit a large range of σ_W, i.e., that depend on aspects of the PDF's that are not sensitive to the value of σ_W, will not affect the final uncertainty measure (as they should not). According to this algorithm, the result for the uncertainty of σ_W is 20.9 nb $< \sigma_W <$ 22.6 nb. These bounds are approximately $\pm 4\%$ deviations from the prediction (21.75 nb) and so we quote a $\pm 4\%$ uncertainty in σ_W due to PDF's.

Now we may determine the increase in χ^2_{global} that corresponds to our estimated uncertainty $\Delta\sigma_W$ in the σ_W prediction. Referring to Fig. 3, a deviation of σ_W by $\pm 4\%$ from the minimum corresponds to an increase $\Delta\chi^2_{\text{global}} \approx 180$. That is, $\Delta\chi^2_{\text{global}}$ in Fig. 1 is 180. In other words, along the direction of maximum variation of σ_W a PDF set with $\Delta\chi^2_{\text{global}} \gtrsim 180$ is found to violate some experimental constraints by this analysis.

C. Comments

We should point out that the above uncertainty estimate, $\Delta\sigma_W/\sigma_W \sim 4\%$, represents only a lower bound on the true

---

[9]For instance, the single uncorrelated systematic error associated with each data point, which is the only systematic error given for most experimental data sets, is clearly only an "effective uncorrelated error" which qualitatively represents the effects of the many sources of systematic error, some of which are really correlated.

uncertainty, since many other sources of error have not yet been included in the analysis, theoretical ones such as QCD higher-order and resummation effects, power-law corrections, and nuclear corrections. These need to be taken into consideration in a full investigation of the uncertainties, but that goes beyond the scope of this paper.[10] We shall add only two remarks which are more directly related to our analysis.

The first concerns a technical detail. In the results reported so far, we have fixed the normalization factors $\{\mathcal{N}_n\}$ in the definition of χ^2_{global} [Eq. (2)] at their values determined in the standard fit S_0. If we let these factors float when we perform the Lagrange multiplier analysis, $\Delta\sigma_W$ will increase noticeably compared to Fig. 3 for the same $\Delta\chi^2_{\text{global}}$. However, upon closer examination, this behavior can be easily understood and it does not imply a real increase in the uncertainty of σ_W. The key observation is that the additional increase (or decrease) in σ_W is entirely due to a *uniform* increase (or decrease) of $\{\mathcal{N}_n\}$ for all the DIS experiments. There is a simple reason for this: The values of the q and $\bar{q}$ distributions in the relevant x range (which determine the value of σ_W) are approximately proportional to $\{\mathcal{N}_n\}_{DIS}$. Although every experiment does have a normalization uncertainty, the probability that the normalization factors of all the *independent* DIS experiments would shift in the *same* direction by the *same* amount is certainly unlikely. Hence we avoid this artificial effect by fixing $\{\mathcal{N}_n\}$ at their "best values" for our study. Allowing the factors $\{\mathcal{N}_n\}$ to vary *randomly* (within the published experimental normalization uncertainties) would not change our estimated value of $\Delta\sigma_W$ significantly.

The second remark concerns the choice of parametrization. We have mentioned that even the robust Lagrange multiplier method depends in principle on the choice of the parton parameter space, i.e., on the choice of the functional forms used for the nonperturbative PDF's at the low-momentum scale Q_0. To check how our answers depend on the choice of parametrization in practice, we have done many similar calculations, using different numbers of free parameters within the same functional form (cf. Appendix C) and using different functional forms for the factor multiplying $x^a(1-x)^b$. We have not seen any dependence of the uncertainty estimates on these changes. Although more radical ways of parametrizing the nonperturbative PDF's might affect the result more, there is no known example of such a parametrization, which at the same time still provides an equally good fit to the full data set.

V. FURTHER EXAMPLES

A. $W^\pm$ production at the LHC

A study similar to the last section has been carried out for inclusive $W^\pm$ production at the LHC. Figure 8 shows χ^2_{global} versus σ_W for the process $pp \to W^\pm X$ at $\sqrt{s} = 14$ TeV,
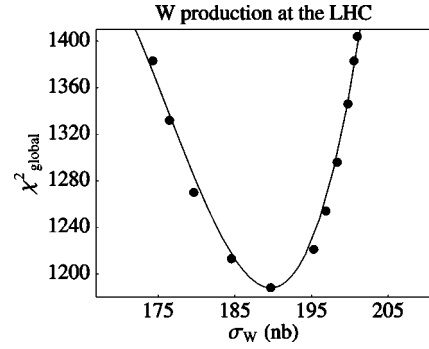
FIG. 8. Minimum χ^2_{global} versus σ_W in nb, for inclusive $W^\pm$ production at the LHC. (Cf. Fig. 3.) The prediction is 189.7 nb. The points are the results of LM calculations. The curve is a polynomial fit to the points.

summed over the two final states. The curve is a smooth interpolation of a series of PDF sets $\{S_\alpha\}$ generated by the Lagrange multiplier method. The best estimate value of the LHC cross section is $\sigma_W = 189.7$ nb.

Comparing Figs. 3 and 8, one immediately notices that the uncertainty of $\sigma_W(\text{LHC})$ is greater than that of $\sigma_W(\text{Tevatron})$ for the same $\Delta\chi^2_{\text{global}}$. This indicates that because W production in pp collisions at the LHC and $\bar{p}p$ collisions at the Tevatron involve different mixtures of parton subprocesses as well as different kinematic ranges, the constraints imposed by current experiments included in the global analysis are also different for the two cases. Referring to the map of the d-dimensional PDF parameter space on the left-hand side of Fig. 1, we are generating sample PDF's along different directions L_X in the two cases. Therefore, it is not surprising that the rate of variation of χ^2_{global} is also different.

To demonstrate this point, and to quantify the uncertainty on the LHC prediction, we have carried out the same error analysis as in Sec. IV, i.e., comparing the alternative PDF's to the individual experiments. Figure 9 gives the final overview of the 90% C.L. ranges of σ_W obtained from these comparisons, analogous to Fig. 7 for the Tevatron cross section. There are some differences compared to the Tevatron case. The LHC prediction is more tightly constrained by experiments that are sensitive to PDF's at small x. This makes sense, because W production at the LHC is not dominated by valence $q\bar{q}$ interactions. We note in particular that the two

FIG. 9. The same as Fig. 7, except for the LHC case.

---

[10]Because there are these additional sources of uncertainty, we have used 90% C.L.'s, rather than 68% C.L.'s, to calculate the error.

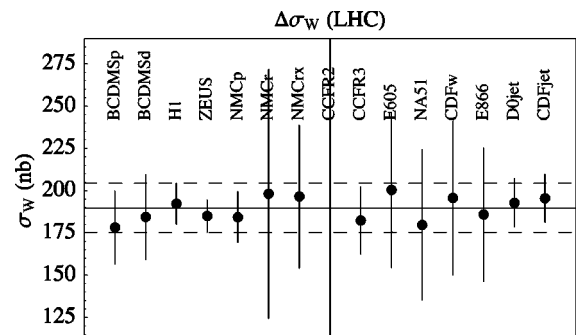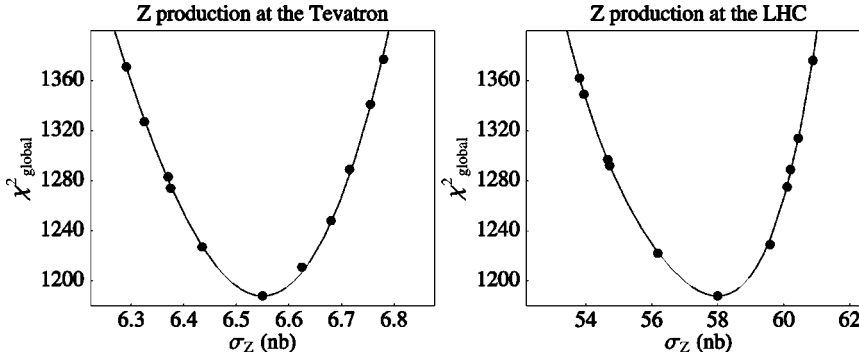FIG. 10. Minimum χ^2_{global} versus σ_Z for (a) the Tevatron, and (b) the LHC. (Cf. Fig. 3.) The predictions are 6.55 nb at the Tevatron and 58.0 nb at the LHC. The points are results of LM calculations. The curves are polynomial fits to the points.

inclusive jet production experiments place significant constraints on σ_W at the LHC.

We can combine the individual error bars in Fig. 9 according to the algorithm proposed in Sec. IV to produce a global uncertainty measure for σ_W at the LHC. The lower bound on σ_W(LHC) obtained by the intersections of the individual ranges is $\sigma_W = 175.3$ nb; the upper bound is $\sigma_W = 204.6$ nb. These bounds, shown as the dashed lines in Fig. 9, correspond to $\pm 8\%$ deviations from the prediction (189.7 nb). The global uncertainty on σ_W(LHC) is thus significantly larger than that on σ_W(Tevatron). Reinforcing this conclusion is the fact that the scatter of the points in Fig. 9 is larger than in Fig. 7.

We can again inspect the increase in χ^2_{global} from the quoted range of alternative fits. A $\pm 8\%$ deviation from the minimum, symmetrized for simplicity, corresponds to the increase $\Delta\chi^2_{\mathrm{global}} \approx 200$. This number is similar to the increase in χ^2_{global} for our estimated uncertainty of σ_W at the Tevatron.

In the companion paper [10], we make some process-independent estimates of $\Delta\chi^2_{\mathrm{global}}$ based on completely different considerations. Those arguments also yield the same order-of-magnitude estimates of $\Delta\chi^2_{\mathrm{global}}$ (in the range from 100 to 200) for acceptable PDF's around the global minimum. Since the effective χ^2_{global}, as a measure of goodness-of-fit, does not have a normal statistical implication, points on a constant χ^2_{global} surface in the PDF parameter space do not necessarily correspond to a constant likelihood. Some variation with the direction in the multidimensional space is to be expected.

B. Uncertainties on Z^0 production

We conclude this section by presenting results from applying the Lagrange multiplier method to Z^0 production at the Tevatron and the LHC.

Figure 10(a) shows the minimum χ^2_{global} as a function of σ_Z at the Tevatron. The global prediction is $\sigma_Z = 6.55$ nb. The experimental measurement by the D0 Collaboration is $\sigma_Z B = 0.221 \pm 0.003 \pm 0.011$ nb; the result from CDF (all data from Run I) is $\sigma_Z B = 0.250 \pm 0.004 \pm 0.010$ nb. [Here B is the branching ratio for $Z^0 \rightarrow e\bar{e}$, which is $(3.367 \pm 0.005) \times 10^{-2}$.] The comparison of the prediction to Tevatron data is discussed below. Analyzing the local χ^2_n in the manner of Sec. IV, in order to quantify the uncertainty of the prediction, we find that the uncertainty of σ_Z(Tevatron) due to PDF's is $\pm 3\%$ of the prediction. The corresponding increase in

χ^2_{global}, symmetrized for simplicity, is approximately 130.

For the LHC process $pp \rightarrow Z^0 X$, Fig. 10(b) shows the minimum χ^2_{global} as a function of σ_Z. The dependence of χ^2_{global} on $\sigma_{W/Z}$(LHC) exhibits a behavior departing from quadratic over the full range of $\sigma_{W/Z}$ under study. This is evidence that the Lagrange multiplier method can go beyond the traditional error matrix approach (which depends on the quadratic approximation) in exploring the neighborhood of the minimum.

The global prediction is $\sigma_Z = 58.0$ nb. Analyzing the local χ^2_n as in the other cases, we find that the uncertainty of σ_Z(LHC) due to PDF's is approximately $\pm 10\%$. As in the case of $W^\pm$ production, the PDF uncertainty for Z^0 production at the LHC is significantly larger than that at the Tevatron. Measurement of $W^\pm$ and Z^0 production at the LHC will, therefore, provide significant information on PDF's.

C. Comparison with existing data

For W and Z production at the Tevatron, we can compare our calculated cross sections σ_W and σ_Z, with their ranges of uncertainty $\pm 4\%$ and $\pm 3\%$ respectively, to the measurements of CDF and D0 from Run I [26]. The comparison is shown in Fig. 11. The two experiments do not measure σ_W and σ_Z *per se*, but rather $\sigma_W B_W$ and $\sigma_Z B_Z$, where B_W is the branching ratio for $W^- \rightarrow e\bar{\nu}$ and B_Z is the branching ratio for $Z^0 \rightarrow e\bar{e}$. We have used the values $B_W = 0.106$ and $B_Z = 0.0337$ for the calculations [27]. The bands in Fig. 11 show the ranges of $\sigma_W B_W$ and $\sigma_Z B_Z$ from our PDF uncertainty study (but no uncertainty included from B_W and B_Z). The two measurements of $\sigma_W B_W$ are consistent with the uncertainty range. The two measurements of $\sigma_Z B_Z$ are not.

It should be noted that CDF and D0 use different normalizations for their luminosity determinations. The CDF Collaboration bases its luminosity purely on its own measure-
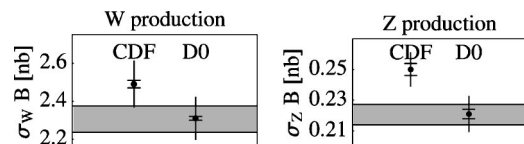

FIG. 11. Experimental measurements of $\sigma_W B_W$ and $\sigma_Z B_Z$ at the Tevatron, compared to the PDF uncertainty band of the theoretical prediction. On the data points, the short error bar is the statistical error, and the long error bar is the combined statistical and systematic error.

ment of the inelastic $\bar{p}p$ cross section [28,29], while D0 uses the world average for this cross section. Thus current luminosities quoted by CDF are 6.2% lower than those quoted by D0. Consequently, all CDF cross section measurements are *ab initio* 6.2% higher than those of D0. If the CDF/D0 measurements of $\sigma_W B_W$ and $\sigma_Z B_Z$ are rescaled by 6.2% with respect to each other, they are in excellent agreement.

Because of the uncertainty in the inelastic $\bar{p}p$ cross section, it has been proposed to normalize future Tevatron (and LHC) physics cross sections to the measured W cross section (or rather $\sigma_W B_W$). This makes the determination of the uncertainty of σ_W due to PDF's even more important.

D. Comparison with previous uncertainty estimates

It is interesting to contrast our results to existing estimates of the uncertainties of σ_W and σ_Z at the Tevatron and LHC colliders based on the traditional method of comparing results obtained from somewhat *ad hoc* PDF's. Some of these previous comparisons for σ_W(Tevatron) between historical PDF's as well as various trial up or down sets obtained by the CTEQ and Martin-Roberts-Sterling-Thorne (MRST) groups, were shown in Fig. 2. We will briefly comment on the results of [13] in the context of this paper.

Reference [13] constructs an extended set of MRST PDF's, of which the most important for σ_W and σ_Z are the standard set MRST99 and three pairs of up and down sets designated

$$\{\alpha_S\uparrow, \alpha_S\downarrow\}, \quad \{g\uparrow, g\downarrow\}, \quad \{q\uparrow, q\downarrow\}$$

in which some aspect of the parton distributions is either raised ($\uparrow$) or lowered ($\downarrow$) by an amount that represents an educated guess of a "standard deviation." The predictions of σ_W and σ_Z are then compared for these alternative PDF sets to get an idea of the uncertainty due to PDF's.

In the case of the Tevatron processes, the deviations of σ_W or σ_Z from the value for MRST99 for sets $\{\alpha_S\uparrow, \alpha_S\downarrow\}$, $\{g\uparrow, g\downarrow\}$, and $\{q\uparrow, q\downarrow\}$ were found to be $\pm 2\%$, $\pm 1\%$, and $\pm 3\%$, respectively. From these results, the authors of [13] concluded that the uncertainties of σ_W and σ_Z at the Tevatron are no more than about $\pm 4\%$, and mainly attributable to the normalization uncertainty in the input u and d distributions. This conclusion appears to be quite consistent with the results of the previous sections based on exploring the variation of the cross section over the entire PDF parameter space. (This range of uncertainty also happens to coincide with what one would get by comparing historical PDF sets, as shown in the right-hand plot of Fig. 2.)

For the LHC, the MRST study found that the uncertainty of σ_W and σ_Z at the LHC is only slightly larger than at the Tevatron; the uncertainty was estimated to be $\pm 5\%$. The largest observed variations came from the sets $\alpha_S\uparrow$ and $\alpha_S\downarrow$, differing from the standard prediction by ± 4–5 %. This estimate is *considerably lower* than the ± 8–10% result obtained by the detailed analysis of the previous sections. We have verified that the PDF sets that give ± 8–10% deviations of σ_W(LHC) and σ_Z(LHC) from the standard prediction (represented by the points at the outer edges of the cor-

responding plots in Figs. 8 and 10) provide equally good or better fits to the global data sets compared to the fits of $\{\alpha_S\uparrow, \alpha_S\downarrow\}$.[11,12] Thus, it is clear that the Lagrange multiplier method can generate *optimal* PDF's, i.e., having the *largest* excursion of the variable X of interest, which are difficult to discover by *ad hoc* trial-and-error methods used in the past.

VI. PDF SETS FOR EXPLORING *W* AND *Z* PHYSICS

The parton distribution sets used in the above calculations are useful for exploring some aspects of W and Z physics at the Tevatron and LHC, since they provide much more reliable estimates on the PDF uncertainties than existing ones in the literature, which are not designed to probe the full range of possibilities in the parton parameter space. With this in mind, we present in this section some representative PDF sets for applications to the rate of W and Z production at the Tevatron and LHC. These PDF's are relevant to the total cross sections σ_W and σ_Z, and each corresponds to a particular direction (L_X) in the PDF parameter space (see Fig. 1). Therefore they are not suitable for estimating the PDF uncertainties of other observables that are sensitive to other aspects of the PDF's. Other PDF sets can be obtained, using the method introduced in this paper, to probe the range of other variables, such as rapidity (y) or transverse momentum (p_T) distributions (hence relevant to the measurement of W mass). These will be investigated in subsequent work. Also, the companion paper [10] supplies information from the Hessian method that can be used to construct the optimal PDF's for any observable X.

The PDF set that yields the "best estimate" for all of the physical cross sections covered in this paper is our standard set S_0. The parametrization of the initial distribution is given in Appendix C. In the following, we present two sets of PDF's that bound the likely range for each of the cross sections.

To exemplify the PDF's that characterize the range of uncertainty of W production at the Tevatron, we use two representative sets, labeled $S_{W,\mathrm{TeV}}^{\pm}$, which correspond to $\sigma_W = \sigma_W(S_0) \pm \Delta\sigma_W$ (with $\Delta\sigma_W/\sigma_W \sim 0.04$), respectively. These two sets are extreme fits obtained by the Lagrange multiplier method. The parameters $\{a\}$ for these sets are given in Appendix C. We now compare some of the parton distributions from the three sets ($S_{W,\mathrm{TeV}}^-, S_0, S_{W,\mathrm{TeV}}^+$), to examine the ranges of variation of the PDF's themselves.

---

[11]The values of χ^2_{global} for $\{\alpha_S\uparrow, \alpha_S\downarrow\}$ are $\{1531, 1356\}$ compared to ~ 1400 for the outermost LHC sets shown in Figs. 8 and 10 and presented in the next section.

[12]The global fits used at first in our exploration of PDF uncertainty were conducted with fixed α_s. To make sure that this restriction does not result in an underestimate of the uncertainties of σ_W and σ_Z, we have examined the effect of freeing α_s in the analysis (but imposing the known constraints from the world average of α_s). The results on the size of the uncertainties are not changed noticeably. This is because the full variations in the PDF's (particularly the gluon) allowed in the Lagrange approach, can absorb the added degree of freedom.
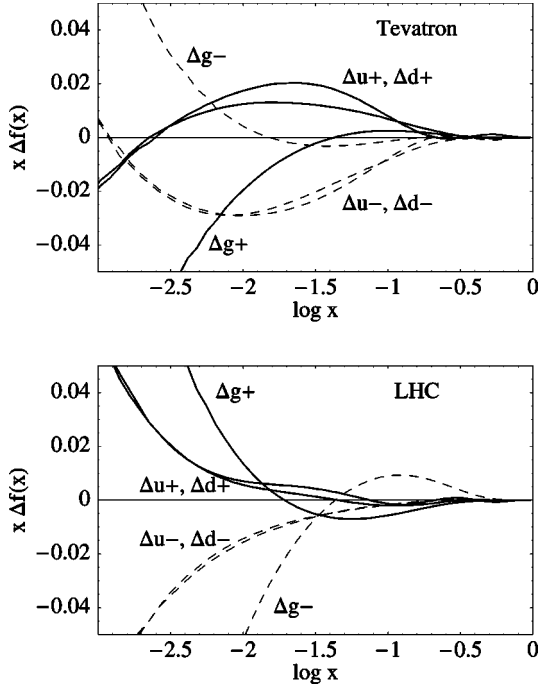
FIG. 12. Comparison of PDF sets $S_W^\pm$ to the standard set S_0 for the Tevatron (upper) and LHC (lower) cases, respectively. $\Delta u_\pm$ is the difference between $u(x)$ with $\sigma_W = \sigma_{W0} \pm \Delta \sigma_W$ and $u(x)$ with $\sigma_W = \sigma_{W0}$; $Q = 80$ GeV. The solid curves are $x\Delta f_+(x)$ and the dashed curves are $x\Delta f_-(x)$. The abscissa is $\log_{10}x$.

Figure 12(a) shows $u(x,Q)$, $d(x,Q)$ and $g(x,Q)$ for $S_{W,\text{TeV}}^\pm$, compared to the standard set S_0 at $Q = 80$ GeV. The function $x\Delta f(x)$ is plotted for each parton flavor, where Δf is $f - f_0$. The gluon function has been divided by 10 to fit on the same graph. The solid curves (Δu_+, Δd_+, Δg_+) correspond to $S_{W,\text{TeV}}^+$, and the dashed curves to $S_{W,\text{TeV}}^-$. For $S_{W,\text{TeV}}^+$, requiring σ_W to be larger than $\sigma_W^{(0)}$, makes the u and d distributions larger than for the standard fit (u_0 and d_0) so Δu_+ and Δd_+ are positive. Then the gluon distribution must be smaller than the standard because of the momentum sum rule. In the case of $S_{W,\text{TeV}}^-$, the reverse is true, resulting in almost a mirror behavior. At the Tevatron, a typical x for the parton-level process $q_1\bar{q}_2 \to W^\pm$ is $M_W/\sqrt{s} = 0.04$. The differences Δu and Δd are significant in the range $0.01 \lesssim x \lesssim 0.04$. The magnitude of $\Delta f(x)$ in this range is a few percent of the standard $f_0(x)$, which makes sense since $\Delta \sigma_W$ is a 4% shift of σ_W for these PDF sets.

We can carry out the same comparison for W production at the LHC. The PDF's that bound the range of uncertainty are designated as $S_{W,\text{LHC}}^\pm$, which correspond to $\sigma_W = \sigma_W(S_0) \pm \Delta \sigma_W$ with $\Delta \sigma_W/\sigma_W \sim 0.08$. The PDF parameters are given in Appendix C. Figure 12(b) shows parton distributions from $S_{W,\text{LHC}}^\pm$. (Again, the gluon has been divided by 10.) In the LHC case, the typical x for the process $q_1\bar{q}_2 \to W^\pm$ is $M_W/\sqrt{s} = 0.006$. The region where Δu and Δd are significant is seen accordingly lower in x than for the Tevatron case.

In Appendix C we also present PDF sets $S_{Z,\text{TeV}}^\pm$ and $S_{Z,\text{LHC}}^\pm$ that characterize the range of uncertainties of Z pro-

duction at the Tevatron and LHC. These correspond to the outlying points on Figs. 10(a) and 10(b). They are similar to $S_{W,\text{TeV}}^\pm$ and $S_{W,\text{LHC}}^\pm$, with small differences in the flavor dependence.

VII. SUMMARY

We have developed the Lagrange multiplier method to calculate uncertainties of physical observables due to PDF's, and we have used the method to determine the uncertainty of the total cross sections for $W^\pm$ production and Z^0 production at the Tevatron and LHC. The method is more reliable than past estimates because (i) it explores all the possibilities in the parameter space of the input PDF's, independent of other assumptions, and (ii) it produces the maximum allowed range for the specified physical variables. This is in contrast to previous attempts which relied on varying certain features of the parton distributions chosen in some *ad hoc* way.

From this analysis, we find that the uncertainty of the prediction for σ_W or σ_Z at the Tevatron with current experimental constraints is approximately ± 3–4 %, and at the LHC the uncertainties are approximately ± 8–10 %. These numbers do not include other uncertainties associated with theoretical approximations, nuclear corrections, and other unexpected sources. We have explored, to some extent, the possible effects due to the choice of parametrization of the nonperturbative input PDF's, and found them to be small. The current work should be considered exploratory in nature, as a first application of this improved approach to error estimates. A more comprehensive study, based on soon to be improved data sets, and including other sources of uncertainties, will produce better overall estimates of the physical predictions.

This study should be regarded as the precursor for many interesting applications to come, on physical processes of interest to the precision study of the SM, and on predictions for new physics at future colliders. Some examples are rapidity distributions of $W^\pm$ and Z^0 production, which contain a wealth of information on parton structure of the nucleon, the W mass measurement, top and Higgs cross sections, etc.

There are other approaches to error estimates in global QCD analysis [1–4]. In general, if greater emphasis is placed on the "rigor" of the statistical method, then the range of experiments that can be included in the analysis is narrower. We have chosen to emphasize the inclusion of the full range of experimental constraints, and adapt the statistical analysis to deal with the practical problems that one faces in such a system. Within our general framework, there is an alternative, complementary approach based on the conventional error matrix method [8]. We explore this latter method, as applied to global QCD analysis of PDF's, in a companion paper [10]. We mention briefly the contrasting features and relative merits of the two approaches here.

The Lagrange multiplier method focuses on a given physical observable X (or a set of observables $\{X_k\}$) and determines the uncertainty ΔX allowed by the global data set within a specified tolerance for the global fit. The error matrix approach, using the Hessian matrix, focuses instead on the uncertainties of the PDF's as represented by the param-

eters $\{a_i; i=1, \ldots, d\}$. It is, in principle, universal because, once determined, these errors can be propagated to *any* physical variable X. However, the results are reliable only if the function $\chi^2_{\text{global}}(a)$ and the observable $X(a)$ can be approximated by lowest-order expansions in the parameters $\{a\}$, and if the numerical computation of the derivatives (the Hessian matrix) is under control. The latter problem is surprisingly difficult for global QCD analysis, because the eigenvalues of the error matrix vary by many orders of magnitude. This problem has been solved [8], and the error matrix results are consistent with the constrained fitting results [10]. Thus, at present, both methods appear to be applicable to the study of uncertainties in global QCD analysis.

In Figs. 8 and 10(b) there is a significant cubic term in the dependence of χ^2_{global} on $\sigma_W(\text{LHC})$ and $\sigma_Z(\text{LHC})$, respectively. To calculate χ^2_{global} versus X accurately in such cases, the Lagrange multiplier method is necessary. Traditional linear error analysis based on the Hessian matrix can only produce a quadratic approximation to the dependence.

When both methods are applicable, the Hessian method is more flexible and easier to apply computationally. But generally the Lagrange method is more robust and reliable. As we expand the investigation to other physical processes of interest, we will continue to test the efficacy of both methods and cross check the results.

ACKNOWLEDGMENTS

APPENDIX A: THE EFFECT OF CORRELATED ERRORS ON $\Delta\chi^2$

The global fitting function χ^2_{global} defined in Eq. (2) resembles the standard statistical variable χ^2, so 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 correlations between measurement errors. The purpose of this appendix is to explore this issue. The effect of correlated errors is potentially a source of confusion.

For simplicity we describe the simplest case: the measurement of a single observable. The arguments can be 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 $m_1, m_2, m_3, \ldots, m_N$. The deviations from the true value are $\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_N$, 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

$$dP = \mathcal{N} \exp\left\{ -\frac{1}{2} \sum_{i,j=1}^{N} \alpha_i C_{ij} \alpha_j \right\} d^N\alpha. \quad (A1)$$

Here C_{ij} is a real symmetric matrix, and $\mathcal{N} = \sqrt{\text{Det} \, C}/(2\pi)^{N/2}$ ensures the normalization condition $\int dP = 1$.

We will need the variance matrix $\langle \alpha_i \alpha_j \rangle$, where the notation $\langle Q \rangle$ means the average of Q in the probability distribution (A1). For this Gaussian distribution,

$$\langle \alpha_i \alpha_j \rangle = (C^{-1})_{ij}. \quad (A2)$$

The mean-square fluctuation E_i of the ith measurement m_i is

$$E_i \equiv \langle \alpha_i^2 \rangle = (C^{-1})_{ii}. \quad (A3)$$

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-squared function $\chi^2_u(m)$ by

$$\chi^2_u(m) = \sum_{i=1}^{N} \frac{(m_i - m)^2}{E_i}. \quad (A4)$$

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

$$\bar{m} = \frac{\displaystyle\sum_{i=1}^{N} m_i/E_i}{\displaystyle\sum_{i=1}^{N} 1/E_i}. \quad (A5)$$

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

$$\frac{dP}{d\bar{m}} = \frac{1}{\sqrt{2\pi\Sigma^2}} \exp[-(\bar{m} - m_0)^2/(2\Sigma^2)]. \quad (A6)$$

The standard deviation Σ of $\bar{m}$ is the rms fluctuation; that is,

$$\Sigma^2 = \int (\bar{m} - m)^2 dP = \frac{1}{D^2} \sum_{ij} \frac{(C^{-1})_{ij}}{E_i E_j}, \quad (A7)$$

where

$$D = \sum_i \frac{1}{E_i}. \quad (A8)$$

The question we wish to answer is this: *How much does $\chi^2_u(m)$ 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*

$$\Delta\chi_u^2 = \Sigma^2 D. \tag{A9}$$

This result follows easily from the definition (A4), because

$$\chi_u^2(\bar{m}+\Sigma) - \chi_u^2(\bar{m}) = -2\Sigma \sum_i \frac{m_i - \bar{m}}{E_i} + \Sigma^2 \sum_i \frac{1}{E_i}, \tag{A10}$$

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

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

$$C_{ij} = \delta_{ij}/E_i. \tag{A11}$$

Then the standard deviation of the mean $\bar{m}$ is $\Sigma = 1/\sqrt{D}$. Thus 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 1σ range corresponds to an increase of χ^2 by 1.

An even more special case is when the errors are uncorrelated and constant: $E_i = \sigma^2$ 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

$$C_{ij} = \begin{cases} 1/\sigma^2 & \text{for } i=j, \\ c/\sigma^2 & \text{for } ij=12 \text{ or } 21, 34 \text{ or } 43, \text{ etc.}, \\ 0 & \text{otherwise}, \end{cases} \tag{A12}$$

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

$$\Sigma = \frac{\sigma}{\sqrt{N}\sqrt{1+c}} \quad \text{and} \quad \Delta\chi_u^2 = 1-c. \tag{A13}$$

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 $i = 1,2,3,\ldots,N$,

$$m_i = m_0 + y_i + \beta \tag{A14}$$

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

$$C_{ij} = \frac{1}{\sigma^2} \left(\delta_{ij} - \frac{s^2}{Ns^2 + \sigma^2} \right), \tag{A15}$$

$$(C^{-1})_{ij} = \sigma^2 \delta_{ij} + s^2. \tag{A16}$$

The variance of the individual measurements (m_i) is

$$\langle m^2 \rangle - \langle m \rangle^2 = \sigma^2 + s^2. \tag{A17}$$

Therefore, our uncorrelated chi-squared variable $\chi_u^2(m)$, defined by ignoring the correlations, is

$$\chi_u^2(m) = \sum_{i=1}^{N} \frac{(m - m_i)^2}{\sigma^2 + s^2}. \tag{A18}$$

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

$$\Sigma^2 = \langle \bar{m}^2 \rangle - \langle \bar{m} \rangle^2 = s^2 + \frac{\sigma^2}{N}. \tag{A19}$$

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

$$\Delta\chi_u^2 \equiv \chi_u^2(\bar{m}+\Sigma) - \chi_u^2(\bar{m}) = \frac{\sigma^2 + Ns^2}{\sigma^2 + s^2}. \tag{A20}$$

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.

Thus for Example 3—a systematic error with 100% correlation between measurements—the increase of χ_u^2 for a standard deviation of $\bar{m}$ is much larger than 1. If s and σ are comparable, then $\Delta\chi_u^2$ is of order N.

If the correlation matrix C_{ij} is known accurately, then the correlation information can be incorporated into the definition of the χ^2 function, in the manner of Appendix B. For the full list of experiments in the global analysis of parton distribution functions, however, the correlations of systematic errors have not been published, so the fitting function χ_{global}^2 only has uncorrelated systematic errors.

We described above the measurement of a single quantity. The determination of parton distribution functions seeks to measure *many* quantities, i.e., the 16 parameters $\{a\}$. The above arguments can be extended to measurements of multiple quantities. If the measurement errors are uncorrelated, then the increase of χ_u^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 X. The prediction of X is at the absolute minimum of χ_u^2. 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 absolute minimum. But if the errors are correlated then $\Delta\chi_u^2 = 1$ is not the correct criterion for one standard deviation of X.

One reason for describing this familiar, even elementary, statistic is to avoid certain misconceptions. Our standard PDF set S_0 is a parametrized fit to 1295 data points with 16 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 $\Delta\chi^2 \sim \sqrt{1295}$ rather than 1. However, this is a misconception. If the errors are uncorrelated (or if the correlations are incorporated into χ^2) then indeed $\Delta\chi^2 = 1$ would represent a standard deviation. But this theorem is irrelevant to our problem, because the large correlations of systematic errors are not taken into account in χ_{global}^2.

APPENDIX B: χ^2 FUNCTION INCLUDING CORRELATED SYSTEMATIC ERRORS

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-squared distribution.

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

$$m_i = t_i + \text{errors} = t_i + \sigma_i r_i + \sum_{j=1}^{K} \beta_{ji} r_j', \quad \text{(B1)}$$

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

$$p(r) = \frac{e^{-r^2/2}}{\sqrt{2\pi}}. \quad \text{(B2)}$$

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

The probability distribution of the measurements is

$$dP = \int \prod_{i=1}^{N} p(r_i) dr_i \prod_{j=1}^{K} p(r_j') dr_j'$$

$$\times \prod_{i=1}^{N} \delta\left(m_i - t_i - \sigma_i r_i - \sum_{j=1}^{K} \beta_{ji} r_j'\right) d^N m. \quad \text{(B3)}$$

Now we will evaluate the integrals over r_i and r_j' in two steps. First evaluate the r_i integrals using the delta functions,

$$dP = \int \prod_{j=1}^{K} dr_j' \mathcal{C}_1 e^{-\chi_1^2/2} d^N m, \quad \text{(B4)}$$

where $\mathcal{C}_1$ is a normalization constant and

$$\chi_1^2 = \sum_{i=1}^{N} \left(\frac{m_i - t_i - \sum_j \beta_{ji} r_j'}{\sigma_i}\right)^2 + \sum_{j=1}^{K} r_j'^2. \quad \text{(B5)}$$

Note that χ_1^2 is a function of $r_1', \ldots, r_K'$. These variables $\{r_j'\}$ could be used as fitting parameters to account for the systematic errors: minimizing χ_1^2 with respect to r_j' would provide the best model to correct for the systematic error of type j. Because χ_1^2 is only a quadratic polynomial in the r_j' variables, the minimization can be done analytically.

To continue evaluating Eq. (B3) we now do the integration over $\{r_j'\}$. We write χ_1^2 in the form

$$\chi_1^2 = \sum_{i=1}^{N} \frac{(m_i - t_i)^2}{\sigma_i^2} - \sum_{j=1}^{K} 2B_j r_j' + \sum_{j,j'=1}^{K} A_{jj'} r_j' r_{j'}', \quad \text{(B6)}$$

where B_j is a vector with K components

$$B_j = \sum_{i=1}^{N} \beta_{ji}(m_i - t_i)/\sigma_i^2, \quad \text{(B7)}$$

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

$$A_{jj'} = \delta_{jj'} + \sum_{i=1}^{N} \beta_{ji} \beta_{j'i}/\sigma_i^2. \quad \text{(B8)}$$

Then the integration over $d^K r'$ is an exercise in Gaussian integration, with the result

$$dP = \mathcal{C} \exp\left[-\frac{1}{2}\chi^2\right] d^N m, \quad \text{(B9)}$$

where $\mathcal{C}$ is a normalization constant and

$$\chi^2 = \sum_{i=1}^{N} \frac{(m_i - t_i)^2}{\sigma_i^2} - \sum_{j=1}^{K} \sum_{j'=1}^{K} B_j (A^{-1})_{jj'} B_{j'}. \quad \text{(B10)}$$

This equation is the appropriate definition of χ^2 for data with correlated systematic errors. The correlated errors are defined by the coefficients β_{ji} in Eq. (B1), which determine the vec-

TABLE II. Table of coefficients.

		A_0	A_1	A_2	A_3	A_4
S_0:	d_v	0.5959	0.4942	4.2785	8.4187	0.7867
	u_v	0.9783	0.4942	3.3705	10.0012	0.8571
	g	3.3862	0.2610	3.4795	-0.9653	1.0
	$\bar{d}/\bar{u}$	3.051E4	5.4143	15.0	9.8535	4.3558
	$\bar{u}+\bar{d}$	0.5089	0.0877	7.7482	3.3890	1.0
	s	0.1018	0.0877	7.7482	3.3890	1.0
$S^{\pm}_{W,\text{TeV}}$:	d_v	0.2891	0.5141	3.8555	10.9580	0.4128
		0.2184	0.2958	4.6267	35.7229	1.0958
	u_v	1.0142	0.5141	3.3614	9.2995	0.8053
		0.2979	0.2958	3.3279	32.8453	0.9427
	g	4.6245	0.4354	3.4795	-0.9728	1.0
		1.8080	0.0458	3.4795	-0.0519	1.0
	$\bar{d}/\bar{u}$	5.908E4	5.6673	15.0	9.8535	4.7458
		2.041E4	5.1506	15.0	9.8535	4.8320
	$\bar{u}+\bar{d}$	0.4615	0.0108	6.6145	0.92784	1.0
		1.2515	0.3338	7.5216	-0.0570	1.0
	s	0.0923	0.0108	6.6145	0.9278	1.0
		0.2503	0.3338	7.5216	-0.0570	1.0
$S^{\pm}_{Z,\text{TeV}}$:	d_v	0.6061	0.5502	4.0017	5.8346	0.5343
		0.3427	0.3728	4.5166	19.8510	0.9966
	u_v	1.2159	0.5502	3.3347	7.3386	0.7711
		0.5247	0.3728	3.3905	20.1006	0.9556
	g	4.4962	0.4321	3.4795	-0.9023	1.0
		2.3113	0.1032	3.4795	-0.6349	1.0
	$\bar{d}/\bar{u}$	4.321E4	5.4724	15.0	9.8535	4.6298
		2.818E4	5.4540	15.0	9.8535	4.4376
	$\bar{u}+\bar{d}$	0.4609	0.0103	6.6671	0.9822	1.0
		0.9900	0.2926	8.3205	2.1648	1.0
	s	0.0921	0.0103	6.6671	0.9823	1.0
		0.1980	0.2926	8.3205	2.1648	1.0
$S^{\pm}_{W,\text{LHC}}$:	d_v	0.7326	0.5008	4.6393	10.8532	1.0595
		0.5671	0.4771	4.2615	8.8355	0.8130
	u_v	1.0608	0.5008	3.4023	9.6622	0.8968
		0.9142	0.4771	3.3761	10.9138	0.8809
	g	2.2379	0.0733	3.4795	-0.9860	1.0
		2.5021	0.3981	3.4795	1.6229	1.0
	$\bar{d}/\bar{u}$	2.178E4	5.2576	15.0	9.8535	4.4810
		4.531E4	5.4979	15.0	9.8535	4.6585
	$\bar{u}+\bar{d}$	1.1980	0.2952	6.9475	-0.5442	1.0
		0.2759	-0.0918	8.2045	6.3950	1.0
	s	0.2396	0.2952	6.9475	-0.5442	1.0
		0.0552	-0.0918	8.2045	6.3950	1.0
$S^{\pm}_{Z,\text{LHC}}$:	d_v	0.5659	0.4616	4.5297	12.3685	0.9836
		0.4585	0.4496	4.2122	10.3850	0.7760
	u_v	0.8344	0.4616	3.3847	12.1129	0.8872
		0.7640	0.4496	3.3566	12.8253	0.8701
	g	2.3282	0.0918	3.4795	-0.9837	1.0
		2.9475	0.4219	3.4795	0.9447	1.0
	$\bar{d}/\bar{u}$	2.421E4	5.3032	15.0	9.8535	4.5341
		4.416E4	5.4708	15.0	9.8535	4.7925
	$\bar{u}+\bar{d}$	1.1130	0.2698	6.8490	-0.5330	1.0
		0.2719	-0.0899	8.1492	6.5300	1.0
	s	0.2226	0.2698	6.8490	-0.5330	1.0
		0.0544	-0.0899	8.1492	6.5300	1.0

tor B_j and matrix $A_{jj'}$. An interesting relation is that the χ^2 quantity in Eq. (B10) is the minimum of χ_1^2 with respect to the parameters $r_1', \ldots, r_K'$.

Another expression for χ^2, which may be derived from Eq. (B3) by Gaussian integration, is [2]

$$\chi^2 = \sum_{i=1}^{N} \sum_{i'=1}^{N} (m_i - t_i)(V^{-1})_{ii'}(m_{i'} - t_{i'}), \quad \text{(B11)}$$

where V_{ij} is the variance matrix

$$V_{ii'} = \sigma_i^2 \delta_{ii'} + \sum_{j=1}^{K} \beta_{ji} \beta_{ji'}. \quad \text{(B12)}$$

It can be shown that the inverse of the variance matrix is

$$(V^{-1})_{ii'} = \frac{\delta_{ii'}}{\sigma_i^2} - \sum_{j,j'=1}^{K} \frac{\beta_{ji} \beta_{j'i'}}{\sigma_i^2 \sigma_{i'}^2}(A^{-1})_{jj'}. \quad \text{(B13)}$$

Therefore Eqs. (B10) and (B11) are equivalent. However, there is a real computational advantage in the use of Eq. (B10) because it does not require the numerical inversion of the $N \times N$ variance matrix.

To check that Eq. (B10) 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

$$\beta_{ji} = \delta_{ji} b_i. \quad \text{(B14)}$$

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

$$B_j = \frac{b_j(m_j - t_j)}{\sigma_j^2} \quad \text{(B15)}$$

and the matrix $A_{jj'}$ is

$$A_{jj'} = \delta_{jj'}\left[1 + \frac{b_j^2}{\sigma_j^2}\right]. \quad \text{(B16)}$$

Substituting these results into Eq. (B10) we find

$$\chi^2 = \sum_i \frac{(m_i - t_i)^2}{\sigma_i^2 + b_i^2}, \quad \text{(B17)}$$

which makes sense; the uncorrelated errors just combine in quadrature.

The statistical quantity χ^2 has a chi-squared 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 $i = 1, 2, \ldots, N$ and $j = 1, 2, \ldots, K$, must be known from the experiment. A chi-squared 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-squared distribution.

We use Eq. (B10) in Sec. IV to compare the constrained fits produced by the Lagrange multiplier method to data from the H1 and BCDMS experiments. Correlated systematic errors are also used to calculate χ^2 for the CDF and D0 jet experiments.

APPENDIX C: PARTON DISTRIBUTION SETS

We give here the PDF's described in Sec. VI. S_0 is the standard set, defined by the absolute minimum of χ^2_{global}. $S^{\pm}_{W,\text{Tev}}$ are fits to the global data sets with extreme values of σ_W(Tevatron), i.e., the outermost points on Fig. 3, generated by the Lagrange multiplier method. $S^{\pm}_{Z,\text{Tev}}$, $S^{\pm}_{W,\text{LHC}}$, and $S^{\pm}_{Z,\text{LHC}}$ are analogous for Z production and W and Z production at the LHC.

The functional form of the initial parton distributions and the definitions of the PDF parameters at the low-energy scale $Q_0 = 1$ GeV are

$$f(x,Q_0^2) = A_0 \, x^{A_1}(1-x)^{A_2}(1+A_3 \, x^{A_4})$$

for $u_v, d_v, g, \bar{u}+\bar{d}, s(=\bar{s})$, and for the ratio

$$\frac{\bar{d}(x,Q_0^2)}{\bar{u}(x,Q_0^2)} = A_0 \, x^{A_1}(1-x)^{A_2} + (1+A_3 \, x)(1-x)^{A_4}.$$

In Table II, the coefficients are given. With a program to solve the PDF evolution equations, the PDF's for an arbitrary momentum scale Q can be generated.

[1] V. Barone, C. Pascaud, and F. Zomer, Eur. Phys. J. C **12**, 243 (2000); C. Pascaud and F. Zomer, Tech. Note LAL-95-05.

[2] S. Alekhin, Eur. Phys. J. C **10**, 395 (1999), contribution to Proceedings of Standard Model Physics (and more) at the LHC, 1999; S.I. Alekhin, Phys. Rev. D **63**, 094022 (2001).

[3] W.T. Giele and S. Keller, Phys. Rev. D **58**, 094023 (1998); W.T. Giele, S. Keller, and D. Kosower, Proceedings of the Fermilab Run II Workshop, 2000.

[4] M. Botje, Eur. Phys. J. C **14**, 285 (2000).

[5] W. Bialek, C.G. Callan, and S.P. Strong, Phys. Rev. Lett. **77**, 4693 (1996); V. Periwal, Phys. Rev. D **59**, 094006 (1999).

[6] R.D Ball, in Proceedings of the XXXIVth Rencontres de Moriond, 1999.

[7] R. Brock, D. Casey, J. Huston, J. Kalk, J. Pumplin, D. Stump, and W.K. Tung, *Proceedings of the Fermilab Workshop, Run II and Beyond*, 2000, hep-ph/0006148.

[8] J. Pumplin, D.R. Stump, and W.K. Tung, preceding paper, Phys. Rev. D **65**, 014011 (2002).

[9] D.E. Soper and J.C. Collins, "Issues in the Determination of Parton Distribution Functions," CTEQ Note 94/01, hep-ph/9411214.

[10] J. Pumplin, D. Stump, R. Brock, D. Casey, J. Huston, J. Kalk, H.L. Lai, and W.K. Tung, this issue, Phys. Rev. D **65**, 014013 (2002).

[11] H.L. Lai, J. Huston, S. Kuhlmann, J. Morfin, F. Olness, J.F. Owens, J. Pumplin, and W.K. Tung, Eur. Phys. J. C **12**, 354 (2000), and earlier references cited therein.

[12] A.D. Martin, R.G. Roberts, W.J. Stirling, and R.S. Thorne, Eur. Phys. J. C **4**, 463 (1998), and earlier references cited therein.

[13] A.D. Martin, R.G. Roberts, W.J. Stirling, and R.S. Thorne, Eur. Phys. J. C **14**, 133 (2000).

[14] BCDMS Collaboration, A.C. Benvenuti et al., Phys. Lett. B **223**, 485 (1989); **237**, 592 (1990).

[15] H1 Collaboration, S. Aid et al., Nucl. Phys. **B439**, 471 (1995); "1994 data," DESY-96-039, hep-ex/9603004, and H1 Webpage.

[16] ZEUS Collaboration, M. Derrick et al., Z. Phys. C **65**, 379 (1995); "1994 data," DESY-96-076, 1996.

[17] NMC Collaboration, M. Arneodo et al., Phys. Lett. B **364**, 107 (1995).

[18] CCFR Collaboration, W.C. Leung et al., Phys. Lett. B **317**, 655 (1993); P.Z. Quintas et al., Phys. Rev. Lett. **71**, 1307 (1993).

[19] E605 Collaboration, G. Moreno et al., Phys. Rev. D **43**, 2815 (1991).

[20] NA51 Collaboration, A. Baldit et al., Phys. Lett. B **332**, 244 (1994).

[21] E866 Collaboration, E.A. Hawker et al., Phys. Rev. Lett. **80**, 3715 (1998).

[22] CDF Collaboration, F. Abe et al., Phys. Rev. Lett. **74**, 850 (1995).

[23] D0 Collaboration, B. Abbott et al., Phys. Rev. Lett. **82**, 2451 (1999).

[24] CDF Collaboration, F. Abe et al., Phys. Rev. Lett. **77**, 438 (1996); F. Bedeschi, talk at 1999 Hadron Collider Physics Conference, Bombay, 1999.

[25] W.T. Giele, S. Keller, and D.A. Kosower, "Parton distributions with errors," in La Thuile 1999, Results and Perspectives in Particle Physics, Proceedings of "Les Rencontres de Physique de la Valle d'Aoste."

[26] F. Lehner, "Some Aspects of W/Z boson physics at the Tevatron," in Proceedings of the 4th Rencontres du Vietnam: International Conference on Physics at Extreme Energies, Hanoi, 2000; FERMILAB-Conf-00/273-E, 2000.

[27] Particle Data Group, D.E. Groom et al., Eur. Phys. J. C **15**, 1 (2000).

[28] CDF Collaboration, F. Abe et al., Phys. Rev. D **50**, 5550 (1994).

[29] M. Albrow, A. Beretras, P. Giromini, and L. Nodulman, FERMILAB-TM-2071, 1999.