New Generation of Parton Distributions with Uncertainties from Global QCD Analysis

Wu-Ki Tung[a]
Department of Physics and Astronomy
Michigan State University
East Lansing, MI 48824 USA

A new generation of parton distribution functions with increased precision and quantitative estimates of uncertainties is presented. This work includes a full treatment of available experimental correlated systematic errors for both new and old data sets and a systematic and pragmatic treatment of uncertainties of the parton distributions and their physical predictions. The new gluon distribution is considerably harder than that of previous standard fits. Extensive results on the uncertainties of parton distributions at various scales, and on parton luminosity functions at the Tevatron RunII and the LHC, are obtained. The latter provide the means to quickly estimate the uncertainties of a wide range of physical processes at these high-energy hadron colliders, such as the production cross sections of the W, Z at the Tevatron and the LHC, and that of a light Higgs.

1 Introduction

Progress on the determination of the parton distribution functions (PDF's) of the nucleon, from global quantum chromodynamics (QCD) analysis of hard scattering processes, is central to precision standard model (SM) phenomenology, as well as to new physics searches, at lepton-hadron and hadron-hadron colliders. There have been many new developments in recent years, beyond the conventional analyses that underlie the widely used PDF's [1,2]. These developments have been driven by the need to quantify the uncertainties of the PDF's and their physical predictions [3–9]. This report describes a comprehensive new global QCD analysis based on the most current data, and on recently developed techniques of analysis [7–9] that:

---

[a] The work reported in this talk has been done in collaboration with J. Pumplin, D.R. Stump, J. Huston, H.L. Lai, and P. Nadolsky.

- allow the precise treatment of fully correlated experimental systematic errors in the χ^2 minimization process, thereby greatly reducing the degree of difficulty of the problem (to the same as the simple case of no correlated systematic errors). In the global analysis context, typically involving ~ 2000 data points and ~ 100 different sources of systematic errors, the conventionally used numerical methods inevitably become practically untenable and numerically unreliable. (Sec. 2.)
- significantly expand the traditional paradigm (in which specific subjectively chosen "best fits" are produced) to a systematic procedure to characterize the parton parameter space in the neighborhood of the global minimum, thereby enabling the systematic exploration of the uncertainties of parton distributions and their physical predictions due to known experimental errors and uncertainties of input theoretical model parameters. This is made possible by an iterative procedure to reliably calculate the Hessian matrix for error propagation in the global analysis context. (Sec. 3)

Results of this analysis consist of:

- a new generation of parton distributions (CTEQ6), which include the standard sets CTEQ6M (msbar scheme), CTEQ6D (DIS scheme), and CTEQ6L (leading order);
- uncertainties on parton distributions (embodied in 40 sets of eigenvector parton distribution sets) and their physical predictions, e.g. uncertainty ranges of various quark-quark, quark-gluon, and gluon-gluon luminosity functions at the Tevatron and LHC energies—from which the uncertainties of a variety of standard model and new physics processes can be inferred.

Details of this work, as well as extensive references (which do not fit in this short summary) can be found in Ref. [10].

2 Experimental input, new method of χ^2 minimization, and new CTEQ parton distributions

Among the new data sets used in this analysis, the most notable ones are from H1[11], ZEUS[12], and DØ[13]. The greater precision and expanded (x, Q) ranges compared to previous data in both processes provide improved constraints on the parton distributions. For the first time in a full global analysis, the correlated systematic errors for all DIS experiments are taken into account. Since ~ 2000 data points from 15–20 diverse experimental data sets are used in this global analysis, traditional methods of χ^2 minimization and error propagation are inadequate. In the covariance matrix approach, numerical instability arises from the inversion of large dimensional matrices for some data sets. An alternative approach is to add experimental fitting parameters, one for each source of systematic error, to the theory model (PDF) parameters. In this case, the total number of fitting parameters becomes so large (of the order of 100) that general programs of χ^2 minimization (such as MINUIT) do not consistently yield reliable results (with errors).

We overcome this problem[8, 10] by performing the minimization with respect to the experimental parameters analytically, before the numerical minimization. This reduces the latter to the same (manageable) level as the case with only theory parameters. (Cf. Section 2.2 of Ref. [10].) A significant additional advantage in this approach is that the analytic results on the optimal deviations associated with sources of systematic error provide important insight on the quality of the fits. It also provides a way to compare data and theory with the effects of systematic errors explicitly taken into account. These features are explained in Ref. [10] (cf. Appendix B in particular).

A primary result of the analysis is a *standard set* of parton distributions (the nominal "best fit") in the $\overline{MS}$ scheme, referred to as CTEQ6M. It provides an excellent global fit to the data sets used. The overall χ^2 for the CTEQ6M fit is 1954 for 1811 data points. Figure 1 shows an overview of the comparison between the new PDF's and the previous generation of CTEQ

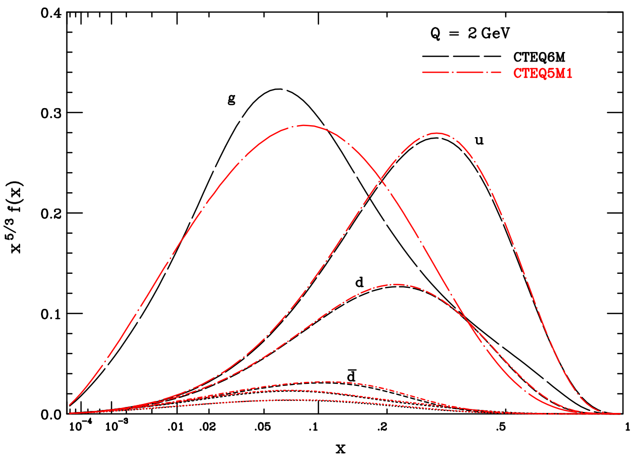

Figure 1: Comparison of CTEQ6M to CTEQ5M1 PDFs at $Q = 2$ GeV.

PDF's, the CTEQ5M1 set, at $Q = 2$ GeV. In order to exhibit the behavior of the PDF's clearly for both large and small x in one single plot, we choose the abscissa to be scaled according to $x^{1/3}$. Correspondingly, we multiply the ordinate by the factor $x^{5/3}$, so that the area under each curve is proportional to the momentum fraction carried by that flavor in the relevant x range. We see that the most noticeable change occurs in the gluon distribution. It has become significantly harder than in CTEQ5M1 and all MRST PDF sets at all Q scales. This behavior is mainly dictated by the DØ inclusive jet data, which lie in the range $50 < Q < 500\,\mathrm{GeV}$ and $0.01 < x < 0.5$. (The higher η bins of this measurement allow a higher x reach than the central jet data from previous measurements.) The hard gluon distribution becomes amplified at lower Q scales, due to the nature of QCD evolution and the fact that there is no direct experimental handle on the gluon at large x and low Q. The enhanced gluon at large x is similar to the CTEQ4HJ and CTEQ5HJ distributions. More detailed comparisons between the new fits, the data sets used, and other existing fits can be found in Ref. [10].

3 New method of error propagation and uncertainties of parton distributions and their physical predictions

There are formidable complications when standard statistical methods are applied to global QCD analysis to make error estimates based on quantitative analysis of the behavior of the χ^2 (or, more generally, likelihood) function in the PDF parameter space. The basic problem is that a large body of data from many diverse experiments, which are not necessarily compatible in a strict statistical sense, is being compared to a theoretical model with many parameters, which has its own inherent theoretical uncertainties as well as numerical instabilities associated with multi-dimensional integrations. In recent papers[7−9], we have proposed and applied a powerful iterative procedure to reliably calculate the behavior of the χ^2 function in the *neighborhood* of the global minimum, overcoming the long-standing difficulties known to plague standard general programs (such as MINUIT) when applied to this type of problems. This method generates eigenvalues and eigenvectors of the Hessian matrix iteratively, as it seeks the right (i.e. physical) step sizes for finite-difference calculations in the multi-dimensional PDF parameter space. In the end, the behavior of the global χ^2 function in the neighborhood of the minimum is encapsulated in $2N_p + 1$ sets of orthonormal eigenvector PDF's, where $N_p \sim 20$ is the number of free PDF parameters. Details are given in Ref. [7,9].

From these PDF sets we can calculate the best estimate, and the range of uncertainty, for the PDF's themselves and for any physical quantity that depends on them. The uncertainty can

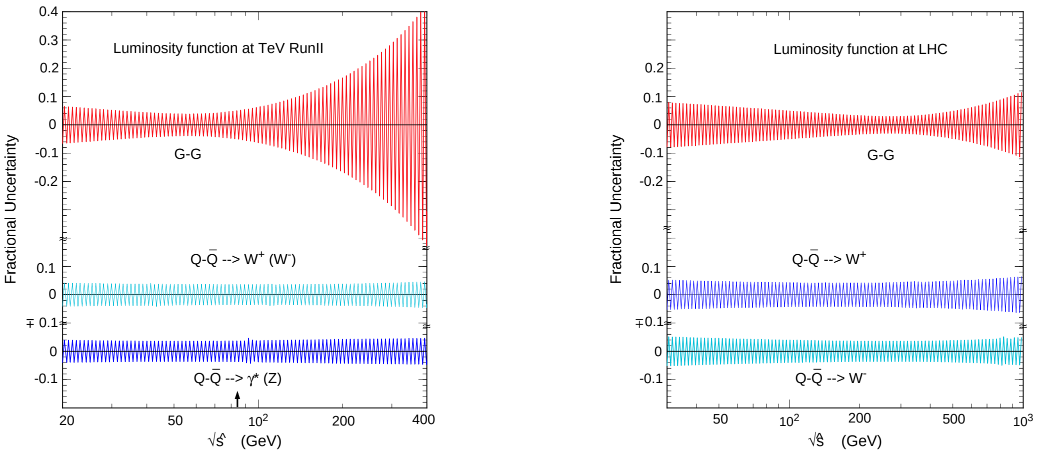

Figure 2: Uncertainties of the luminosity functions at the Tevatron and at the LHC.

be computed from the simple master formula $\Delta X = \frac{1}{2}\left(\sum_{i=1}^{N_p}\left[X(S_i^+) - X(S_i^-)\right]^2\right)^{1/2}$, where X is the observable and $X(S_i^\pm)$ are the predictions for X based on the PDF sets $S_i^\pm$ from the eigenvector basis. As an illustration, Fig. 2 shows the percentage error bands of the quark–anti-quark and gluon-gluon luminosity functions at the Tevatron and the LHC as functions of the parton subprocess CM energy. The uncertainties of the quark and gluon distributions themselves, as well as the quark-gluon luminosity functions, are given in Ref. [10].

4 Concluding Remarks

There are many complex issues involved in a comprehensive global parton distribution analysis. Foremost among these on the experimental side is the "imperfection" of real experimental data compared to textbook behavior. For instance, some experimental measurements appear to be statistically improbable because $\chi^2 \,/\, N$ deviates from 1 substantially more than the expected $\pm\sqrt{2/N}$; or different precision experimental measurements of the same physical quantities appear to be statistically incompatible in all regions of the model parameter space. The methods of Ref. [7-9] cannot resolve these problems completely—no global analysis method can—but the tools developed in this formalism have brought much progress to the global analysis endeavor, and have made possible a deeper look into some of these problems. They allow us to assess the acceptability and compatibility of the affected data sets in more practical terms, and to suggest pragmatic ways to deal with the apparent difficulties. These detailed studies were not possible in the conventional approach. On the theoretical side, the uncertainties on the perturbative QCD (PQCD) calculations of the various physical processes included in the global analysis are not easily quantified in a uniform way. Obviously, much work lies ahead for continued progress in our effort to pin down the parton structure of the nucleon, and to test the limits of perturbative QCD.

References

1. MRST Analyses: A. D. Martin, R. G. Roberts, W. J. Stirling and R. S. Thorne, Eur. Phys. J. **C 4** (1998) 463; Eur. Phys. J. **C 14** (2000) 133; Eur. Phys. J. C **23** (2002) 73.
2. Previous CTEQ Analyses: H. L. Lai *et al.*, Phys. Rev. **D 51** (1995) 4763; Phys. Rev. **D 55** (1997) 1280; Eur. Phys. J. **C 12** (2000) 375.
3. S. I. Alekhin, Eur. Phys. J. **C 10** (1999) 395; and Phys. Rev. **D 63** (2001) 094022.

4. M. Botje, Eur. Phys. J. **C 14** (2000) 285.
5. V. Barone, C. Pascaud and F. Zomer, Eur. Phys. J. **C 12** (2000) 243; C. Pascaud and F. Zomer, Tech. Note LAL-95-05.
6. W. T. Giele and S. Keller, Phys. Rev. **D 58** (1998) 094023; W. T. Giele, S. A. Keller and D. A. Kosower, [hep-ph/0104052].
7. J. Pumplin, D. R. Stump and W. K. Tung, Phys. Rev. **D 65** (2002) 014011.
8. J. Pumplin *et al.*, Phys. Rev. **D 65** (2002) 014013.
9. D. R. Stump *et al.*, Phys. Rev. **D 65** (2002) 014012.
10. J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, P. Nadolsky and W. K. Tung, arXiv:hep-ph/0201195.
11. H1 Collaboration: C. Adloff *et al.*, Eur. Phys. J. **C 13** (2000) 609; Eur. Phys. J. **C 19** (2001) 269; Eur. Phys. J. **C 21** (2001) 33.
12. ZEUS Collaboration: S. Chekanov *et al.*, Eur. Phys. J. **C 21** (2001) 443.
13. DØ Collaboration: B. Abbott *et al.*, Phys. Rev. Lett. **86** (2001) 1707; and Phys. Rev. **D 64** (2001) 032003.