# Global QCD Analysis of Parton Distributions

Wu-Ki Tung
*Michigan State University, E. Lansing, MI*

Recent global QCD analyses of high energy lepton-hadron and hadron-hadron interactions are reviewed. New data on asymmetries in Drell-Yan and W-lepton-asymmetry processes contribute to better determine the $d/u$ ratio; while improved experimental data on inclusive jet production, in conjunction with precise deep inelastic scattering data, place good constraints on the gluon over a wide range of $x$. Relevant theoretical considerations are discussed. Comparisons of the results of the CTEQ5 and MRST analyses are made; and their differences are described. Open issues and the general problem of determining the uncertainties of parton distributions are discussed.

The structure of hadrons represented by parton distributions is an essential part of our knowledge of the elementary particle physics world. The interpretation of existing experimental data in terms of the Standard Model (SM), the precision measurements of SM parameters, as well as the direct search for signals for physics beyond the SM, all rely heavily on calculations based on Quantum Chromodynamics (QCD) and the QCD-parton picture, with the parton distribution (and fragmentation) functions as essential input. The (non-perturbative) parton distribution functions at some given momentum scale are currently determined phenomenologically by a global analysis of a wide range of available hard scattering processes involving initial-state hadrons, using the perturbative QCD-parton framework.

This talk reviews features of recent QCD global analyses by the MRS [1] and CTEQ [2] groups based on new experimental results of the last two years, highlights the relevant physics issues, describes the remaining open problems in the determination of parton distributions, and assesses various sources of uncertainties in the parton distribution parameters and the prospect for quantifying the uncertainties.

## I. NEW EXPERIMENTAL INFORMATION AND THEIR USE IN GLOBAL ANALYSIS

Experimental data relevant for recent global QCD analyses typically consist of DIS data from BCDMS, NMC, H1, ZEUS, CCFR, E665 experiments; Drell-Yan data from E605; W-lepton-asymmetry data from CDF; direct photon production data from WA70, UA6; and inclusive jet data from D0 and CDF. Notable new experimental results which motivated the current round of global analysis consist of:

**Deep inelastic scattering**: The NMC and CCFR collaborations have published final analyses of their respective data on muon-nucleon [3] and neutrino-nucleus [4] scattering. These new results lead to subtle changes in their implications for $\alpha_s$ and parton distribution determination.

**Lepton-pair production ($p/d$) asymmetry**: The E866 collaboration has measured the ratio of lepton-pair production (Drell-Yan process) in $pp$ and $pd$ collisions over the $x$ range 0.03 – 0.35 [5], thus expanding greatly the experimental constraint on the ratio of parton distributions $\bar{d}/\bar{u}$ (compared to the single point of NA51 at $x = 0.18$ [6]). This data set has the most noticeable impact on the new round of global analysis.

**Lepton charge asymmetry in W-production**: The CDF collaboration has improved the accuracy and extended the $y$ range of the measurement of the asymmetry between $W \to \ell^{\pm}\nu$ at the Tevatron [7]. This provides additional constraints on $d/u$.

**Inclusive large $p_T$ jet production**: The D0 collaboration has recently finished the final analysis of their inclusive jet production data, including information on the correlated systematic errors [8]. The CDF collaboration also has presented new results from their RunIB data set [9]. Systematic errors in these data sets dominate the experimental uncertainty over much of the measured $p_T$ range. The correlated systematic errors provide important information on the shape of the differential cross-section, $d\sigma/dp_T$, and constrain the parton distributions accordingly.

**Direct photon production**: The E706 collaboration at Fermilab has published the highest energy fixed-target direct photon production data available to date [10]. The measured cross-sections lie a factor of $2-3$ above the traditional next-to-leading (NLO) QCD calculation, thus posing a real challenge for their theoretical interpretation and their use in global analysis. (Cf. below.)

*Physics issues* which need to be considered in order to incorporate the new experimental data in current global analysis of parton distribution functions include:

**Charge asymmetry data and quark flavor differentiation**:

Most inclusive processes are not sensitive to differences between the quark parton flavors, since contributions from them are summed in the cross-section. In global analysis, these differences represent "fine structure" that can be resolved by including physical quantities asymmetric in the various flavors. In particular, the difference between the $u$ and $d$ quarks is determined by differences between cross-sections with proton/neutron targets in DIS and Drell-Yan processes, or with $W^{\pm}$ final states (manifested by the decay leptons) in $\bar{p}p$ collisions. As mentioned in the previous section, new data from E866 and CDF have an immediate impact on flavor differentiation in current global analyses. These new data are complemented by the final results from the very precise measurement of $F_2^d/F_2^p$ by the NMC experiment.

An important source of uncertainty in the study of quark flavor dependence arises from the necessity of using DIS and Drell-Yan data on a deuteron target, in the absence of a neutron target. Recently, Bodek and Yang [11] advocated an unconventional behavior of the $d/u$ ratio at large $x$, based on a modified version of MRS parton distributions and specific deuteron- and target-mass corrections. They pointed out the importance of studying this issue in a full global QCD analysis. We have found, by such an analysis, that equally consistent descriptions of all current data can be obtained with or without applying the corrections of [11]. A different uncertainty concerns heavy-target correction for neutrino experiments which directly impact on the determination of the charged current (CC) structure function of the nucleon. The origin of the apparent disagreement of the ratio of existing CC and NC structure functions with the classic "charge ratio" (5/18 rule) in the region $x < 0.1$, and the relation of this to the true size of the strange quark fraction as well as the validity of charge symmetry for parton distributions (i.e. $f_p^u = f_n^d$ ) are still not understood. [12] These open questions deserve further study. In the absence of compelling reasons to do otherwise, both the MRS and CTEQ analyses follow the practices of previous work, and avoid applying any model-dependent correction effects.

**Direct photons, inclusive jets, and the gluon distribution**:

Since the recently published E706 direct photon data [10], measured at 530 and 800 GeV, cover a wide range of $x$, and report comparatively small statistical and systematic errors, one might hope to determine the gluon distribution directly from this process over the full range covered by this and earlier experiments. However, the factor of $2-3$ difference between this data and the conventional NLO QCD calculation highlights the need of a clearer theoretical understanding of this process before the data can be utilized in a convincing way to determine parton distributions. Cf. [13–15] for detailed discussions on relevant issues and on-going theoretical efforts. Experimentally, the consistency between the new data and the previous fixed-target data could be open to question [13], the problem being partly dependent on the theoretical framework used to compare experiments at different energies. We note that, direct photon production has also been measured at hadron colliders [16]. The cross-section at high $p_T$ agree rather well with NLO QCD calculations; however, at the low $p_T$ end, one also observes an enhenced cross-section compared to theory. The statistics for these experiments are currently too low to make these data useful for the global analysis of parton distributions.

Inclusive large $p_T$ jet production at the Tevatron, on the other hand, provides a much more reliable experimental constraint on the gluon distribution, since the NLO QCD theory has been shown to be rather stable [17,18] in the region $p_T > 40$ GeV where measurements exist. This energy scale is considerably higher than that of fixed-target direct photon discussed in the previous paragraph. Multi-soft gluon effects are insignificant for data in this range. The recent CTEQ5 analysis made full use of the latest D0 and CDF data, and confirms the previous finding that the combination of jet and DIS data constrain the gluon distribution quite well in the range $0.05 < x < 0.25$.

**Charm production in DIS and QCD formalism for heavy quark production:**

Preliminary results on charm production at HERA [19,20] have highlighted the need for a more careful treatment of heavy quarks in the perturbative QCD (PQCD) formalism. Theories for heavy quark production have existed for some years [21–24], and the CTEQ4 analysis provided several sets of parton distributions which incorporate charm quark mass effects (CTEQ4HQ,4F3,4F4). The recent MRST analysis applied a similar method [25], resulting in distributions which are in a different scheme than previous MRS distributions. It is, however, important to bear in mind the limitations of the current state of the art on this subject. Experimentally, one can only measure the cross-sections for producing $D$ and $D^*$ mesons in certain kinematic ranges. Extracting $F^c_{2,\mathrm{exp}}$ requires: (i) an extrapolation of $D$- and $D^*$-production data to the full phase space to obtain $F_2^{D,D^*}$; and (ii) a procedure to infer $F^c_{2,\mathrm{exp}}$ from $F_2^{D,D^*}$ involving, among other uncertainties, the not so well-known fragmentation functions for $D, D^*$. On the theoretical side one faces a different, but related, dilemma. On one hand, among the existing schemes for treating heavy quarks in PQCD, $F^c_{2,\mathrm{th}}$ is in principle defined only in the fixed-three-flavor scheme (with $u, d, s$ being the only quark partons); but this scheme is not suitable for quantitative treatment of high energy (i.e. collider) inclusive processes which are essential for global QCD analysis. On the other hand, in modern $\overline{MS}$ schemes admitting a non-zero mass charm quark as an active parton at high energies [23–25], which are suitable for global analysis at high energies, "$F_2^c$" is not a well-defined quantity in principle because the naive "$F_2^c$" contains large logarithms of the same type that are resummed into charm parton distributions. Only in $F_2^{\mathrm{tot}}$, and in $F_2^{D,D^*}$, do these logarithms cancel between contributions from all parton flavors to yield infrared safe quantities that are suitable for comparison with experiment. For these experimental and theoretical reasons, the CTEQ group did not use the preliminary charm production data as input to the fitting (as MRST did), but presented comparisons of calculated "$F^c_{2,\mathrm{th}}$" using the new parton distributions with available data on $F^c_{2,\mathrm{exp}}$. [2] This makes little difference in practice, since current data on charm production do not carry enough weight to influence the results, even if they are included in the global fit.

## II. NEW PARTON DISTRIBUTIONS

The recent MRS [1] and CTEQ [2] global analyses made some notable different choices concerning the above issues which led to noticeable differences in results – more so than in precious generations of parton distributions for which the two groups had made rather similar choices. (This only underlines the misleading nature of the common practice of assessing uncertainties of parton distributions by the differences between the results of various groups.) The first table below compares the main features of the two new analyses. The second one lists the parton distribution sets from the CTEQ5 analysis.

Comparison of the differences in choices made by the MRST and CTEQ5 global analyses.

| | MRST | CTEQ5 |
|---|---|---|
| HQ sch. | on-shell (TR) | 5M,D,L: conventional,<br>5HQ: on-shell (ACOT),<br>5F3,F4: fixed-flavor |
| Dir. Ph. | WA70 + $k_T$ 's<br>(√ E706) | —<br>(√ WA70, E706) |
| Incl.Jet | – | CDF + D0 |
| $\alpha_s(m_Z)$ | fixed: 0.1175 | fixed: 0.118 |

| PDF set | Description |
|---|---|
| | conventional (zero-mass parton) sets |
| CTEQ5M | $\overline{MS}$ scheme |
| CTEQ5D | DIS scheme |
| CTEQ5L | Leading-order |
| CTEQ5HJ | large-$x$ gluon enhanced |
| | on-mass-shell heavy quark sets |
| CTEQ5HQ | $\overline{MS}$ (ACOT) scheme |
| CTEQ5F3 | fixed-flavor-number ($N_f = 3$) scheme |
| CTEQ5F4 | fixed-flavor-number ($N_f = 4$) scheme |

The main result from these analyses is that the wide range of hard processes and experimental data are described extremely well by the PQCD framework. In the CTEQ analysis, the nominal $\chi^2$, as a simple measure of goodness-of-fit, is around 1275 for 1295 data points with 16 parameters for the parton distribution functions at a fixed initial scale (chosen to be 1 GeV for this study), and 7 relative normalization parameters for the experiments. The $\chi^2$'s per point for individual experiments are uniformly close to one, hence there is no inconsistency (except for that between NC and CC DIS in the range $0.01 < x < 0.1$ mentioned earlier). We shall skip familiar theory vs. experiment plots which demonstrate this remarkable agreement; and confine ourselves to a brief description of the motivation and the features of the different sets of CTEQ5 parton distributions, in relation to the new experimental results and future developments, before comparing them with the MSRT distributions in the next section.

The **CTEQ5M** set is defined in the $\overline{MS}$ scheme, matched with conventional NLO hard cross-sections calculated in the zero-quark-mass approximation for all active flavors, including charm and bottom. This set is the most convenient one to use for general calculations, as the vast majority of available hard cross-sections in the literature and in existing programs have been calculated in this limit. It represents an updated version of the CTEQ4M distribution set. Compared to the previous generation of distributions, such as CTEQ4M, the most noticeable changes are in the difference of $\bar{u}$ and $\bar{d}$ quarks, due to the influence of the new data of E866, NMC, and CDF W-lepton asymmetry. Figs. 1a,b show the changes in the combinations $\bar{d}/\bar{u}$ and $\bar{d}-\bar{u}$. Figs. 2a,b,c show comparisons of NLO QCD calculations based on the CTEQ5M parton distributions to the experimental data of CDF in the W-lepton asymmetry, of NMC on the DIS deuteron to proton ratio, and of E866 on the Drell-Yan deuteron to proton ratio respectively. Excellent agreement is observed in all cases. There is no obvious need for a different treatment of the deuteron data as suggested in Ref. [11], although we find that the alternative scenario is also allowed by the global analysis.

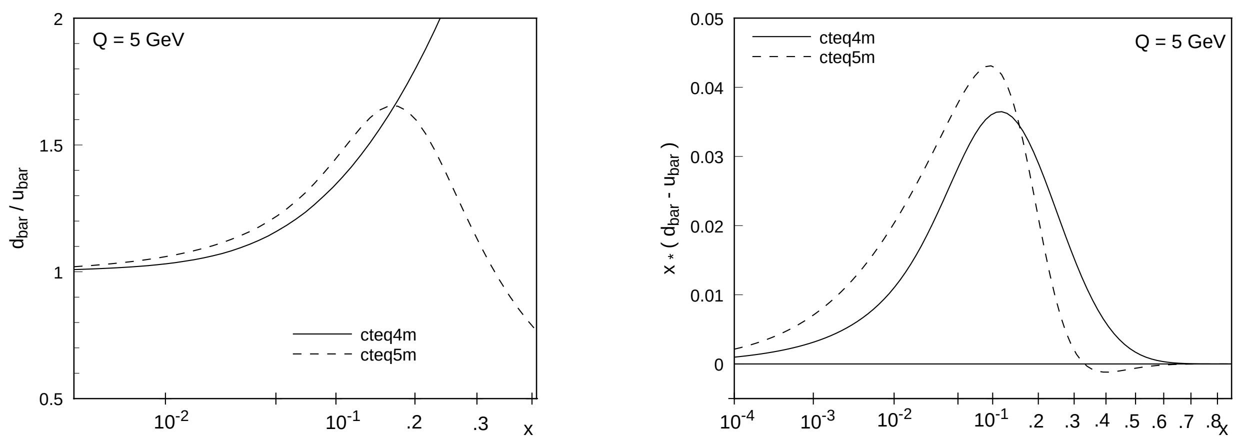


FIG. 1. The change in $\bar{d}$ and $\bar{u}$ distributions from CTEQ4 to CTEQ5 analyses as the result of new experimental input: (a) $\bar{d}/\bar{u}$ ; and (b) $\bar{d}-\bar{u}$.

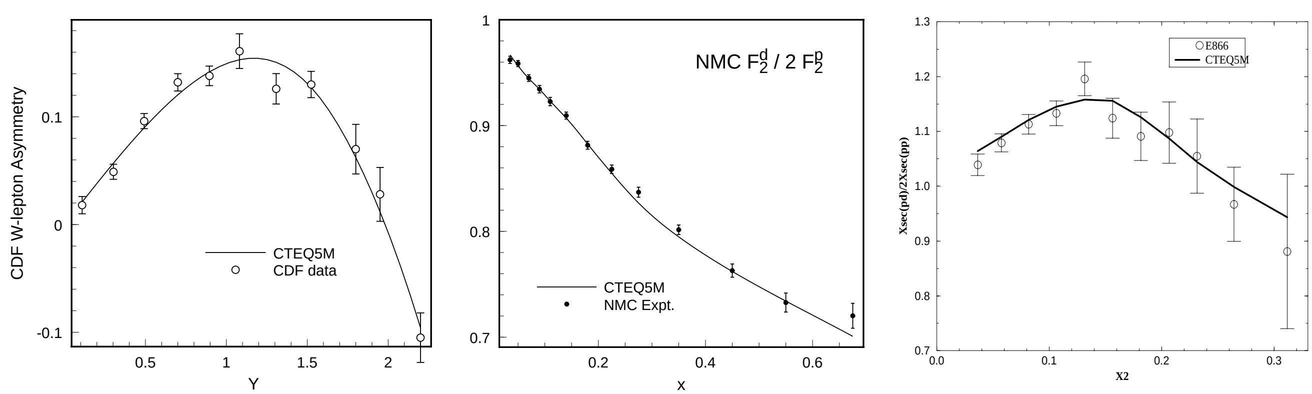


FIG. 2. Comparison of CTEQ5M fit to experimental data for the three experiments which are relevant to the differentiation between $\bar{d}$ and $\bar{u}$ quark flavors: (a) CDF W lepton-asymmetry; (b) NMC DIS deuteron to proton ratio; and (c) E866 Drell-Yan deuteron to proton ratio.

Of much current interest is the degree to which we can determine the **gluon distribution**. In the CTEQ analysis, the gluon distribution is constrained by the $Q$-evolution of the DIS structure functions as well as the inclusive jet production data from CDF and D0. The jet data, available for the range $40 < p_T < 450$ GeV, are systematic error limited in most regions, except at very large $p_T$. The known correlated systematic errors, which constrain the shape of the differential distribution, are incorporated in the global fit. The gluon distribution obtained in this way is close to that of CTEQ4M and the previous MRS parton sets, but differ noticeably from those of the recent MRST. This difference, and the comparison of CTEQ5M with the jet data will be discussed in detail in the next section.

The CDF RunIA inclusive jet production data [26] stimulated much interest in physics at large $x$, in particular the possible range of the gluon distribution in that region. The CTEQ4HJ parton distribution set, proposed two years ago [27], has served as a useful example in investigations of various large $x$ phenomena. In a subsequent systematic study [28], we showed that the range of uncertainty of the gluon distribution is quite significant beyond $x$ 0.2. For currently available jet production data, CDF inclusive jet $p_T$ distribution, as well as the CDF and D0 di-jet mass $m_{jj}$ distributions [29], continue to show a rise of the cross-section above the NLO QCD calculations based on conventional parton distributions, at large $p_T$ and $m_{jj}$ respectively.[1] It is therefore desirable to update the CTEQ4HJ parton distribution set, to complement the new CTEQ5M. This updated set is designated **CTEQ5HJ**. It gives almost as good a global fit as CTEQ5M to the full set of data on DIS and DY processes, with only marginally higher overall $\chi^2$, and has the feature that the gluon distribution is significantly enhanced in the large $x$ region, resulting in improved agreement with the observed trend of jet data at high momentum scales mentioned above. The existence of excellent fits of this kind again serves to illustrate the fact that the large $x$ region remains a fertile ground for further experimental exploration and theoretical development. Fig. 3a shows the comparison between the gluon distributions of CTEQ5HJ and CTEQ5M at 2, 5, and 80 GeV. Due to the feature of QCD evolution mentioned earlier, the large difference of the two distributions at low $Q$ represents the amplified effect of fitting jet data at an energy scale greater than 40 GeV at the Tevatron. In Fig. 3b, we show the ratio of the CDF and D0 data, both normalized to NLO QCD calculation based on CTEQ5HJ. This plot shows that CTEQ5HJ accounts well for both data sets, and that the two data sets are in quite good agreement with each other. Note that experimental systematic errors are not included in this plot; and a relative normalization factor of 4% between the two experiments is applied (this difference in luminosity is due to known sources to the two experiments).

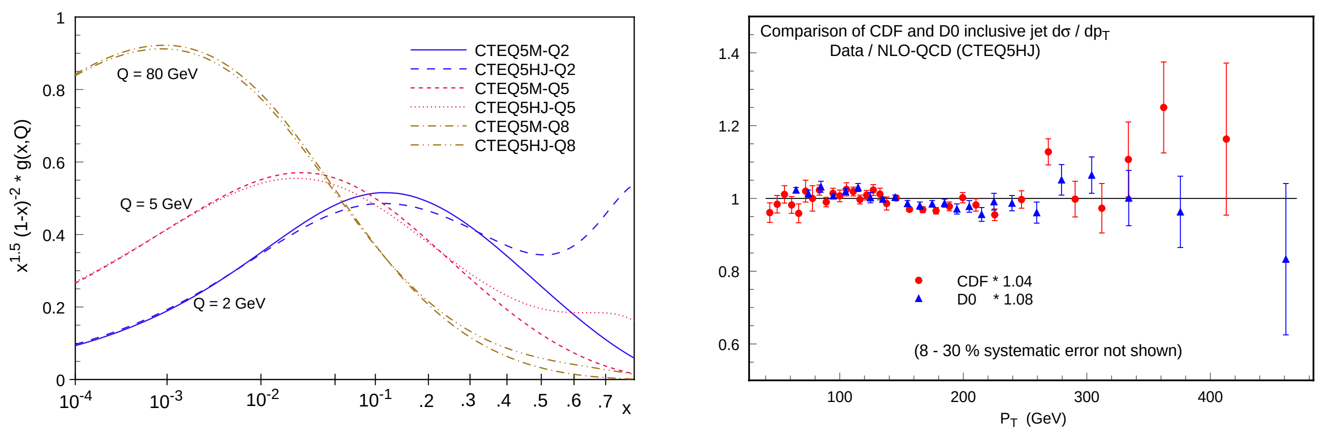


FIG. 3. Gluon determination and inclusive jet data: (a) the CTEQ5M and CTEQ5HJ gluons at three energy scales; and (b) comparison of the CTEQ5HJ fit to the CDF and D0 inclusive jet data.

The other CTEQ parton distribution sets listed in the above table are: CTEQ5D (DIS scheme), CTEQ5L (leading order), CTEQ5HQ (on-shell ACOT [23,24] heavy quark formalism), and CTEQ5F3/4 (fixed-3/4-flavor heavy quark scheme). Space limitation prevents any discussion of these alternative parton sets. (Cf. [2].)

[1] Due to the size and interpretation of current experimental errors, whether this observed trend in each of the two experiments is statistically significant may be open to question.

## III. COMPARISON OF CTEQ AND MRST PARTON DISTRIBUTIONS

As indicated in the first table, there are two main differences in how the most recent CTEQ5 and MRST global analyses were conducted.

First, the MRST group adopted a new procedure for treating charm quark mass effects in DIS processes, applying the method of Ref. [25]. This procedure is similar in principle to that used for CTEQ4HQ and CTEQ5HQ, although the method of [25] does differ from that of [23,24] in the specifics of how the mass effects are treated. The general CTEQ5 distributions, CTEQ5M,D,L, on the other hand, continue to use the conventional zero-quark-mass formalism since the vast majority of contemporary application programs use that formalism. The difference due to these two choices is discussed in [2].

Of more phenomenological interest is the comparison of the gluon distribution in the CTEQ and MRST analyses, because of its implications for future high energy processes. On this issue, the difference due to the choice of scheme is completely overshadowed by that due to the choice of experimental input: to complement the DIS constraints in determining $G(x,Q)$, CTEQ used the inclusive jet data of CDF and D0, as discussed above; whereas MRST relied on direct photon production results of WA70, applying a range of $k_T$ broadening corrections using the E706 data as a constraint. Fig. 4a shows the comparison of $G(x,Q)$ from CTEQ5M and CTEQ5HJ with those of the three MRST sets at $Q = 5$ GeV. The significant difference observed can be readily understood in terms of the inputs.

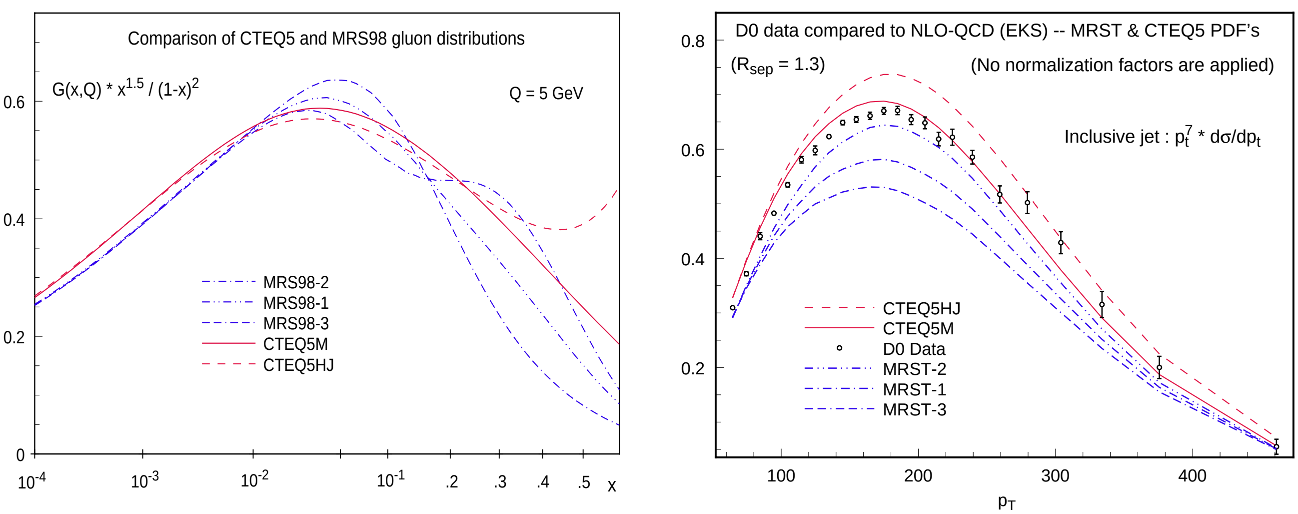


FIG. 4. (a) Comparison of the gluon distributions of CTEQ5 and MRST analyses (cf. detailed discussion in the text); and (b) Comparison of calculated inclusive jet cross-section with D0 data. (Comparison with CDF data shows the same features.)

The large range of variation between the MRST sets in the region around $x \sim 0.25$ reflects the freedom of choice of the $k_T$-broadening parameter $\langle k_T \rangle$ which produces a very significant correction factor to the theoretical cross-section (recall this factor needs to be of the order of $2 \sim 3$ for E706 to agree with data), in addition to the well-known large scale dependence for NLO QCD predictions [14,30,31]. For a detailed discussion of the choices made to obtain this range, see Ref. [1]. The much narrower band among the CTEQ5 sets in this $x$ range results from the rather tight constraints on the shape of $G(x,Q)$ imposed by the inclusive jet cross-section (which has rather stable NLO QCD theory predictions) and the stringent criteria we adopted for "good fits" in this particular study. The MRST-G↑ (MRS98-2 in the figure) set uses WA70 data with zero $k_T$ broadening. Its $G(x,Q)$ is closest to that of CTEQ5M, as can be seen in the plot. For the $x > 0.5$ region, the wide range of variation of the CTEQ5 sets reflects the lack of experimental constraints on $G(x,Q)$ at large $x$. The convergence of the MRST gluons in this region is likely due to choosing the same functional form at large $x$ for all these sets. Finally, the differences between the two series in the range $0.01 < x < 0.1$ is most likely correlated to the differences in $0.1 < x < 0.6$ as the result of the momentum sum rule constraint.

Fig. 4b, shows the comparison of the D0 inclusive jet production data with NLO QCD calculations using the CTEQ5 and MRST parton distribution sets. (Comparison of CDF data with these sets shows the same features.) The calculation is performed using the Ellis-Kunzst-Soper program [17] with the scale parameter $\mu = E_T/2$ and the jet-separation parameter $R_{sep} = 1.3$ (which is the current value favored by both CDF and D0). For this comparison, the experimental normalization is not floated, as done in fitting the parton distributions, for the obvious reason that the same experimental data points cannot have many different normalizations. The MRST curves lie considerably lower than the CTEQ5 ones, because their $G(x, Q)$ is much lower in the relevant $x$ range, as already seen in Fig. 4a. The significance of the observed differences must be assessed within the context of relevant theoretical and experimental considerations, some of which have been discussed above.

## IV. CONCLUSIONS AND COMMENTS ON UNCERTAINTIES OF PARTON DISTRIBUTIONS

As both theory and experiment improve steadily, global QCD analyses continue to show a remarkable agreement of perturbative QCD with available data on the wide range of hard-scattering processes and allow us to extract the non-perturbative parton distributions with increasing accuracy. There are, however, still many areas where more detailed theoretical and experimental work will help to clear up current uncertainties, and allow more precise determination of the parton structure of the nucleon. We devote this concluding section to discussions of these areas of uncertainty.

On the theory side, the most desirable advance would be a **reliable calculation of direct photon production** (especially in the $p_T$ range of fixed-target experiments), which could elevate the phenomenology of this process to the same level of confidence as for DIS, DY, and jet processes, and thereby lead to a definitive determination of the gluon distribution. Many theorists are working on the soft-gluon resummation corrections to the NLO QCD calculation to see if this can lead to a quantitative theory [15], accounting for the factor of 2~3 or more difference between the NLO theory and experiment beyond E706 energies. However, this explanation of the discrepancy is not yet universally accepted [14].

Considerable progress has been made on the **differentiation between $u$ and $d$ quarks** in the last year, as the result of complementary information provided by several different DIS and DY measurements, as discussed in Sec. I. However, this analysis relies heavily on: (i) the assumption of charge symmetry (i.e. $f_p^{u(d)} = f_n^{d(u)}$) (which has been questioned in recent literature [12]; and (ii) the extraction of neutron cross-sections from actually measured deuteron cross-sections. The size of nuclear corrections needed to extract the neutron cross-section is still a subject of some controversy. These corrections could affect the determination of $d/u$, especially at large $x$ [11]. We found that, in the global analysis context, all current data can be consistently described within the PQCD formalism with or without applying a deuteron correction; and chose to take the simple option of not applying any such correction. A detailed study is underway to probe this issue more thoroughly. Such studies will clearly benefit from a better theory for nuclear corrections. Conversely, better phenomenological analyses of the existing abundant data could provide useful input to the study of the nuclear effects.

There has been little advance in the unambiguous determination of the **strange distribution**. The long-standing dilemma associated with the discrepancy of the strange quark distribution inferred from the di-muon neutrino data and that from the difference of neutral and charged current structure functions [32] remains unresolved. This problem may be related to that of charge symmetry. [12] To make real progress, the most useful development would be measurements of physical cross-sections (or structure functions) for charm production in neutrino-nucleon scattering, which can then be incorporated in the global analysis. If this cannot be done for existing measurements, one hopes it will be achievable in the analysis of the NuTeV experiment.

The **charm distribution** has entered the arena of global QCD analysis with the availability of charm production data in neutral current interactions, particularly at HERA. This has directed attention to more precise formulations of QCD theory including massive quarks, which have been actively pursued over the last ten years. Unfortunately, more precise formulations necessarily lead to additional scheme dependence of the PQCD calculations, thereby complicate the application of the parton formalism for users of parton distributions. We briefly described some of the pertinent issues in Sec. I. An interesting related question is: *is there a non-perturbative component of charm inside the nucleon?* [33,34] This question has not yet been addressed by any of the existing global analysis efforts – all assume a purely radiatively generated charm distribution which vanishes at the threshold scale. Since the charm mass is only slightly above the nucleon mass, there is no strong argument against the existence of an additional non-perturbative component of charm. This issue can be studied once more abundant precision data become available.

It is universally recognized that for a wide range of theoretical and experimental applications, it is extremely important to know the **range of uncertainties of the parton distributions**. The ultimate goal would be to have parton distribution sets with a well-defined correlation matrix for their parameters [35]. To see what needs to be done toward achieving this goal, it is first necessary to recognize the major sources of uncertainties in global QCD analysis and address them systematically.

The most obvious uncertainties are the reported experimental errors. The non-trivial aspect of these are the correlated systematic errors. In principle, there are standard methods to incorporate these errors, often represented as covariance matrices, in data-fitting. Several recent attempts and proposals have been made to pursue this approach [36]. In practice, since only a limited number of experiments present information on correlated errors, the input data sets for the global analysis are much more restricted than required to determine the different parton flavors. In addition, this task is much more complex than appears on the surface, because: (i) it is known that the standard covariance matrix method is not robust under certain conditions [37] and can lead to pathological results, and (ii) the diversity of experiments involved in a global analysis, and the non-uniform information they provide, can easily vitiate some of the essential assumptions underlying the statistical analysis method.

Theoretical uncertainties that affect the global analysis are much less obvious and much harder to quantify than the experimental errors. The magnitudes of the uncertainties due to higher-order effects, scale-dependence, soft-gluon resummation, higher-twist effects, nuclear (deuteron) corrections, etc., vary widely from process to process, and from one kinematic region to another. Thus, while the uncertainties of NLO calculations of DIS and DY processes are known to be under control (except near the boundaries of the kinematic region), and those of inclusive jet cross-section are also stable, the same is far from true for direct photon production (at $p_T$ values of most available data) and for heavy quark production in hadron collisions. These uncertainties have to be dealt with on a case by case basis, using the most up-to-date knowledge of the specific process.

Last, but by no means least, there are hidden uncertainties associated with the choice of functional forms for the non-perturbative initial parton distributions. Although the parameters in these functions are determined by comparison with experiment, the choice of functional form introduces implicit correlations between the parton distributions at different $x$ ranges. We have encountered this hidden correlation often in our investigation of the range of variations of the gluon distribution in previous and current CTEQ analyses. The simpler the functional form (or the more economical the parametrization), the more rigid is the implied correlation. To reduce this undesirable correlation, one cannot, however, indiscriminately increase the degrees of freedom of the parametrization. If there are not enough experimental constraints to determine the parameters, one will get unpredictable artificial behavior of the parton distributions that is not related to the experimental input. We have also encountered examples of this kind in the course of our analyses. Only as more precise experimental data become available for more processes, does it become possible to refine the parametrization in a progressive manner.

The presence of uncertainties of the second and third kind has important implications on efforts to quantify the implications of experimental systematic error on parton distribution analysis, because both uncertainties are of a highly correlated nature and all three are inextricably intertwined.

In the CTEQ series of global QCD analyses [32,38,39,28,40], we try to assess the current knowledge of the parton distributions keeping all the above sources of uncertainties in perspective, and make the best educated estimates on the uncertainties as possible. The global analysis of parton distributions is yet far from being an exact science, due to its complexity and comprehensive scope. However, the steady progress that has been achieved clearly demonstrates that vigorous pursuit of the open problems summarized above will continue to improve our knowledge of the parton structure of hadrons, and pave the way for advances in all fronts in elementary particle physics.


[1] A. D. Martin and R. G. Roberts and W. J. Stirling and R. S. Thorne, *Eur. Phys. J.* **C4**, (1998) 463, hep-ph/9803445.
[2] CTEQ Coll.: H.L. Lai, et al, (1999) hep-ph/9903282.
[3] NMC Collaboration: M. Arneodo et al., *Nucl. Phys.* **B483** (1997) 3.
[4] CCFR collaboration: W.G. Seligman et al., *Phys. Rev. Lett.* **79** (1997) 1213.
[5] E866 collaboration: E.A. Hawker et al., *Phys. Rev. Lett.* **80** (1998) 3715, hep-ex/9803011.
[6] NA51 collaboration: A. Baldit et al., *Phys. Lett.* **B332** (1994) 244.
[7] CDF collaboration: F. Abe *et al. Phys. Rev. Lett.*, **81**, (1998) 5754 hep-ex/9809001.
[8] D0 Collaboration: B. Abbott et al., FERMILAB-PUB-98-207-E, e-Print Archive: hep-ex/9807018,
[9] F. Bedeschi, talk at 1999 Hadron Collider Physics Conference, Bombay, January, 1999.
[10] Fermilab E706 Collaboration: L. Apanasevich et al., *Phys. Rev. Lett.* **81**, (1998) 2642, hep-ex/9711017
[11] U. K. Yang and A. Bodek, (1998) hep-ph/9809480.
[12] B. Ma, Phys. Lett. **B274**, (1992) 111; B. Ma, A. Schafer and W. Greiner, Phys. Rev. **D47**, (1993) 51; hep-ph/9211202. C. Boros, J.T. Londergan and A.W. Thomas, hep-ph/9810220; C. Boros, J.T. Londergan and A.W. Thomas, Phys. Rev. Lett. **81**, 4075 (1998) hep-ph/9806249.
[13] L. Apanasevich, et al, to be published in Phys. Rev. (1998) hep-ph/9808467.
[14] P. Aurenche, et al, (1998) hep-ph/9811382. and references therein for earlier papers by the same authors.
[15] See, for instance, E. Laenen, G. Oderda, G. Sterman, hep-ph/9806467; S. Catani, M. Mangano, P. Nason, hep-ph/9806484; H-n. Li, hep-ph/9811340.
[16] UA2 Collaboration: J. Alitti etal., *Phys. Lett.* **B263**, (1991) 544; CDF collaboration: F. Abe *et al. Phys. Rev. Lett.*, **73** (1994) 2662 (1994) ; D0 Collaboration: S. Abachi *etal, Phys. Rev. Letters*, **77**, (1996) 5011.
[17] S. Ellis, Z. Kunszt, and D. Soper, *Phys. Rev. Lett.* **64** 2121 (1990); S. Ellis, Z. Kunszt, and D. Soper, *Phys. Rev. Lett.* **69**, 3615 (1992).
[18] F. Aversa et al., *Phys. Rev. Lett.* **65**, 401 (1990). W. Giele et al., *Nucl. Phys.* **B403**, 2121 (1993).
[19] H1 collaboration: C. Adloff et al., *Zeit. Phys.* **C72** (1996) 593.
[20] ZEUS collaboration: J. Breitweg et al., *Phys. Lett.* **B407** (1997) 402; Paper N-645 presented at International Europhysics Conference on High Energy Physics, HEP97, Jerusalem 1997.
[21] M. Glück, E. Reya and M. Stratmann, *Nucl. Phys.* **B422** (1994) 37.
[22] E. Laenen, S. Riemersma, J. Smith and W.L. van Neerven, *Nucl. Phys.* **B392** (1993) 162.
[23] M. A. G. Aivazis and John C. Collins and Fredrick I. Olness and Wu-Ki Tung, *Phys. Rev.* **D50**, (1994) 3102, hep-ph/9312319.
[24] J. C. Collins, *Phys. Rev.* **D58**, (1998) 094002, hep-ph/9806259.
[25] R. S. Thorne and R. G. Roberts, *Phys. Lett.* **B421**, (1998) 303, hep-ph/9711223.
R. S. Thorne and R. G. Roberts, *Phys. Rev.* **D57**, (1998) 6871, hep-ph/9709442.
[26] CDF Collaboration (Abe et al.), *Phys. Rev. Lett.* **77**, (1996) 439.
[27] J. Huston, et al, *Phys. Rev. Lett.* **77**, (1996) 444, hep-ph/9511386.
[28] J. Huston, et al, *Phys. Rev.* **D58**, (1998) 114034, hep-ph/9801444.
[29] D0 Collaboration (B. Abbott et al.). FERMILAB-PUB-98-220-E, Submitted to Phys.Rev.Lett., e-Print Archive: hep-ex/9807014
[30] P. Aurenche, R. Baier, M. Fontannaz and D.Schiff, *Nucl. Phys.* **B297** (1988) 661; P. Aurenche, P. Chiappetta, M. Fontannaz, J.Ph. Guillet, and E. Pilon, *Nucl. Phys.* **B399** (1993) 34.
[31] J. Huston, et al, *Phys. Rev.* **D51**, (1995) 6139, hep-ph/9501230.
[32] CTEQ Coll.: James Botts, et al, *Phys. Lett.* **B304**, (1993) 159, hep-ph/9303255.
H. L. Lai, et al, *Phys. Rev.* **D51**, (1995) 4763, hep-ph/9410404.
[33] S. Brodsky, *et.al.*, Phys. Lett. 93B (1980) 451; *Phys. Rev.* **D23** (1981) 2745.
[34] B.W. Harris, J. Smith and R. Vogt, *Nucl. Phys.* **B461**, 181 (1996) hep-ph/9508403.
[35] Davison E. Soper, John C. Collins. CTEQ-NOTE-94-01, (1994) hep-ph/9411214.
[36] S. Alekhin, hep-ph/9611213. W.T. Giele and S. Keller, *Phys. Rev.* **D58**, 094023 (1998) hep-ph/9803393.
[37] David Seibert, *Phys. Rev.* **D49**, 6240 (1994) hep-lat/9305014.
G. D'Agostini, *Probability and Measurement Uncertainty in Physics – A Bayesian Primer*, (1995) hep-ph/9512295.
[38] CTEQ collaboration: H.-L. Lai et al., *Phys. Rev.* **D55** (1997) 1280.
[39] H. L. Lai and W. K. Tung, *Z. Phys.* **C74**, (1997) 463, hep-ph/9701256.
[40] S. Kuhlmann, H. L. Lai and W. K. Tung, (1997) *Phys. Lett.* **B409, (1997)** 271 hep-ph/9704338.