Z. Phys. C – Particles and Fields 52, 13–29 (1991)

Zeitschrift für Physik C Particles and Fields

Parton distributions from a global QCD analysis of deep inelastic scattering and lepton-pair production

J.G. Morfín[1⋆,3] and W.K. Tung[1,2⋆⋆]

[1] Fermi National Accelerator Laboratory, Batavia, IL 60510, USA
[2] Illinois Institute of Technology, Chicago, IL 60616, USA
[3] Laboratori de Fisica d'Altes Energies, Universitat Autonoma de Barcelona, Spain

Received 16 July 1990

Abstract. Parton distribution functions consistent with neutrino and muon deep inelastic scattering as well as Drell-Yan pair production results have been extracted. This analysis incorporates experimental systematic errors which are the dominant errors in recent deep inelastic scattering experiments. The dependence of the results on factors such as kinematic cuts in the data, heavy target corrections, and choice of initial functional form are also explored. The form adopted is motivated by perturbative QCD and particularly useful in exploring the small-x extrapolation of the distributions. This is crucial for studying the range of predictions for Collider, HERA, and SSC/LHC cross sections. Representative distribution function sets are presented in a very compact parametrized form both in the DIS and MS-bar renormalization schemes.

1 Introduction

The QCD parton model provides a comprehensive framework for describing general high energy processes in current and planned accelerators and colliders. In this framework, the cross-section $\sigma_{AB\to C}$ for a hadron–hadron collision process $A+B\to C+X$, where C represents a final state of physical interest, is written as the convolution of a set of universal parton distribution functions $f_A^a(x,Q)$ and parton-initiated fundamental hard cross-sections $\sigma_{ab\to C}$

The process-independent parton distributions are the key link between the physically measured cross-sections $\sigma_{AB\to C}$ and the basic processes of the theory $\sigma_{ab\to C}$. The precise determination of these functions are of fundamental importance for the interpretation of experimental results within the Standard Model and in any search for "new physics". Several well-known parametrizations [1] of parton distributions extracted from early experimental data and using leading order QCD formalism have long been in wide use. Analyses based on more current data and incorporating next-to-leading order QCD evolution of the distribution functions have also recently become available [2, 3]. However, most of these analyses use only limited sets of data, some of which have since been significantly revised. Most of these analyses do not include experimental systematic errors or explore the dependence of the results on such factors as kinematic cuts in the analysed data, heavy target corrections, choice of initial functional forms, etc. Since most modern applications of the QCD parton model either require a high degree of accuracy or involve extrapolation of the kinematic variables (x,Q) well beyond the measured range, all these factors can significantly affect the predictions. Thus, it is crucial to incorporate all available experimental information in the analysis and to adopt a procedure which allows one to systematically map out the range of possible behavior of the parton distributions within and beyond the current x and Q domain.

A comprehensive review of the current status of DIS experiments and parton distribution analyses including a plan to compile an extensive database and to investigate all the relevant factors in such analyses was given at the 1988 snowmass workshop [4]. We report here first results of this global analysis and present representative parton distribution sets with a range of different behaviors in a simple and easy-to-use form. Finally, we discuss some of the physical consequences in current collider processes, as well as projections for HERA and SSC energies. Reports on the some detailed results, including specific effects of the various factors mentioned above and updates, will be given in subsequent publications.

2 Parametrization of the parton distributions

One of the goals of the present analysis is to adopt natural functional forms for the distributions which will be appropriate for parton distributions at all Q. The evolved

⋆ Permanent address

⋆⋆ This work is partially supported by the National Science Foundation under Grant No. PHY89-05161. This work was also supported by Argonne National Laboratory during the 1988–89 academic year when the author was on Sabbatical leave at the Laboratory from IIT

distributions can then be given in a simple analytic form as is done with the initial distributions; and the parameters of the distributions become (slowly) varying functions of Q. This will minimize the special role of the arbitrarily chosen initial point of evolution and, more importantly, help us to visualize (hence gain some physical insight) on how the parton distributions actually evolve. It is also important to adopt parametrizations which are guaranteed to be positive definite for all values of x (as parton distribution functions should be) and which vary smoothly over the entire range of x.

Although there is no real theory on the correct functional form of the parton distributions in the framework of QCD, the above considerations plus the natural occurrence of logarithmic factors in perturbative quantum field theory lead us to adopt the ansatz:

$$x f(x, Q) = e^{A_0} x^{A_1} (1-x)^{A_2} \ln^{A_3} x \ln^{A_4}(1-x). \quad (1)$$

Here, for clarity, we have suppressed a parton flavor label. The A-coefficients will be referred to as "shape parameters" in our data-analysis. In addition to the features mentioned earlier, an important advantage of this parametrization is that it provides a simple and versatile way to study the small-x behavior of the parton distributions. By selectively choosing A_1 and/or A_3 as active parameters, we can use existing data to explore the full range of power- and/or logarithmic-law small-x extrapolations of the parton distributions from the current range (cf. Sect. 9).

Existing parton distribution analyses, as a rule, do not discuss the effects of different choices of the functional forms of the initial distributions or the number of free parameters used to characterize these distributions in the data-fitting. This can be, however, an important issue in this type of analysis, especially when the data set or the kinematic range used is relatively limited. Physical parameters such as $\Lambda_{\overline{\text{QCD}}}$ and familiar shape parameters, such as (α, β) in $x^{\alpha}(1-x)^{\beta}$, can depend sensitively on the choice of the functional form if it is not general enough. Conclusions drawn on these parameters or the assignment of errors to these parameters, without investigating this form dependence, do not necessarily reflect real physics. Furthermore, a specific set of ad hoc functions for the initial distributions can become inappropriate when additional data sets from different processes are included in later analyses. In other words, for a meaningful parton distribution analysis, the choice of functional forms and parameters must be general, flexible, and responsively constructed.

We have attempted to address this problem systematically by allowing, in principle, the rather general multi-parameter functional form, (1) for all relevant parton flavors, and adopting the following strategy to adjust the scope of our fitting parameters in the global analysis: at each stage of the parton distribution analysis, only those parameters in the proposed general parametrization form are activated for which the data sets considered have discriminatory power. The number of parameters are increased, when appropriate, as new data sets are added. In this way we maintain flexibility at each stage, do not lose touch with physics, and ensure that the output from the fitting program be reasonably unique throughout. The next section describes how this strategy is applied to the analysis of DIS data. In our work so far, we have found it sufficient to use the functional form (1) without the factor $\ln(1-x)^{A_4}$; hence we shall omit it in subsequent discussions. It is conceivable, however, that such a factor may be required in future analysis of a wider range of experimental data.

3 Procedure for analysing deep inelastic scattering data

The DIS data sets included in this analysis are CDHSW [5] neutrino scattering results in conjunction with EMC [6] and BCDMS [7] muon scattering experiments.* These data sets were used in various combinations to test both the consistency of the experimental results and the stability of the fitting results.

The effect of experimental considerations were examined such as the sensitivity of the results to minimum Q^2- and W-cuts on the data selected, the influence of an "EMC effect" correction when combining results from light and heavy targets, and the inclusion of systematic in addition to statistical errors. We examined the stability of fitting results as the values of Q^2- and W-cuts are varied; and determined that, without a priori knowledge of higher twist contributions, consistent results are obtained with $Q^2 > 10\,\text{GeV}^2$ and $W > 4\,\text{GeV}$. These default cutoff values preserve the bulk of the high statistics data and decrease any possible contamination of higher twist effects by at least a factor of 4 compared to most recent global analyses.

It is, of course, imperative to include the experimental systematic errors, especially when data from several different high statistics experiments are included. This has not been done systematically in previous published parton parametrization analyses. The statistically rigorous way of combining statistical and systematic errors requires that the fit consist of a "loop" where the experimental central values are shifted by their corresponding systematic errors and then these shifted values are used in the minimum χ^2 or maximum likelihood part of the fit which takes account of the statistical error. Considering the large number of data points included in this analysis as well as the various systematic effects often quoted for each point (e.g. four separate systematic errors from the BCDMS data set), such a procedure would require a prohbitively large amount of computing power. For most of our fits, we use the conventional procedure of combining in quadrature the statistical error with a single combined systematic error. We do investigate how the results change if systematic errors are left out (as is done by most of the existing PDF fits) on the one hand, or if the systematic and statistical errors are added linearly on the other. We plan to do a more thorough error analysis on a smaller set of data in the future to assess the importance and usefulness of any more rigorous and elaborate procedure.

* These data sets are the only high statistics deep inelastic scattering ones with information on systematic errors available to us

The influence of theoretical factors on fitting results were also examined. Among these were the use of 1-loop or 2-loop evolution kernels and the considerations concerning the choice of appropriate functional forms for the parton distributions as discussed earlier. In practice, we use the following parametrization of the initial distributions:

$$x f^a(x, Q) = e^{A_0^a} x^{A_1^a} (1-x)^{A_2^a} \ln^{A_3^a}\left(1 + \frac{1}{x}\right) \qquad (2)$$

where a is the parton-flavor label, and the $\ln x$ factor of (1) has been slightly modified to make it positive definite and to avoid a potential unwanted singular behavior near $x = 1$. All calculations reported below use the full 2-loop evolved [8] parton distributions and the appropriate 1-loop Wilson coefficients [9] for the structure functions. The program used for numerical solution of the 2-loop evolution equation has been used previously to study the small-x behavior of parton distribution functions and has provided critical tests of other existing evolution programs (and found to be accurate) [10].

Our procedure consists of the following steps. We begin with the CDHSW neutrino-iron scattering results for xF_3 which depends only on the sum of all valence quark distributions. We fit data using an initial *flavor-blind* valence quark distribution of the form (1) with the shape parameters and $\Lambda_{\overline{\text{QCD}}}$ as the fitting parameters. The normalization parameters A_0 for the valence quarks are determined by the quark-number sum rules. Excellent fits are obtained with only three shape parameters (A_1, A_2, A_3). We note that xF_3 is particularly suited to determine the parameter A_1 for valence quarks (which controls the small-x behavior of the valence), and the QCD parameter $\Lambda_{\overline{\text{QCD}}}$. The resulting value of A_1 is fixed at this point. Since data on xF_3 is not as statistically accurate as those on F_2, we improve the determination of the shape parameters A_2 and A_3 for valence quarks by including the CDHSW F_2 data with $x > 0.3$ in the fit as well. The small contribution of sea-quarks and gluons to F_2 in this region is verified by including a conventional sea/gluon contribution and noting no significant change in the fit results. We again get excellent fits which yield "better" values of A_2 and A_3 for the (flavor-blind) valence quarks*. We now introduce the muon scattering results. We begin with the BCDMS data on hydrogen and deuterium keeping the $x > 0.3$ cut. Since u-quarks and d-quarks contribute unequally to the muon structure functions, distinct shape parameters $A_2^{u,d}$ and $A_3^{u,d}$ are introduced and determined at this stage. In these fits, we introduce an additional parameter – an overall normalization factor between the two experimental data sets – to be determined by the fit. We now perform the previous step again using CDHSW data along with EMC (in place of BCDMS). Finally, both EMC and BCDMS data are combined with the neutrino scattering data to determine whether a reasonable combined fit can be obtained. With these better-determined $A_2^{u,d}$ and $A_3^{u,d}$ parameters, we go back to determine improved values of the A_1-parameters for the valence quarks and $\Lambda_{\overline{\text{QCD}}}$. We then constrain this set of valence quark shape parameters within limited range (3 *s.d.*) in the next step of analysis.

The next step is to determine the sea and gluon parameters by dropping the x cut in fitting the combined neutrino and muon data. In this exercise, we again use CDHSW data with BCDMS alone, then with EMC alone and finally with both data sets together. An important new feature of our fits compared to those in the literature, made possible by the use of the parametric form (2) for the parton distributions, is that we selectively use the parameters A_1 and A_3 for the gluon and sea-quarks as fitting parameters, thus letting data determine these parameters (either singly or both) which control behavior of the parton distributions when extrapolated to very small values of x.

Even with the results of the very high statistics deep inelastic scattering experiments currently available*, the individual parton distributions cannot be fully differentiated since the data is not directly sensitive to the gluon and the individual sea distributions. For most of our analysis, we take the A_2-parameter for the gluon to be one unit less than that of the sea quarks (reflecting the conventional wisdom that the gluons are the source of the sea quarks hence must have a harder x distribution). We begin with $SU(3)$ symmetric sea quark distributions and obtained excellent fits. We also tried $SU(2)$-symmetric sea together with a strange quark content of the order indicated by DIS dimuon data [11], and unrestricted sea and gluon distributions. Neither the added parameters nor the added input improve the quality of the (already satisfactory) fit; they only make the fitting parameters under-determined. Clearly, additional data from other processes are needed to effect further *flavor differentiation* among the gluon and sea-quark distributions.

As the last step of this analysis of DIS data, we fine-tune all the active parameters – valence, sea and gluon – in a final multi-parameter "best fit" to the full data set chosen for that fit. The fits obtained in this way tend to have rather hard gluon distributions, characterized by A_2^G of the order 3.5 to 4.5. But, as is generally known, DIS data alone do not provide stringent constraints on the shape of the gluon distribution.

4 Inclusion of lepton-pair production data in the analysis

The remaining freedom in the shape parameters of the sea quarks and the gluon can be further constrained with the aid of other physical processes, notably the hadroproduction of real or virtual vector bosons-γ, W, and Z. Among these, the best candidate is lepton-pair production via the virtual γ (the Drell–Yan process) where the theory is relatively clean [12] and where high-statistics

* For us, *better* fits means fits with lower values of χ^2 per degree of freedom. With the crude procedure of combining statistical and systematic errors in quadrature or in linear form, the meaning of the absolute value of χ^2/dof is not clear. However, its relative value does give a good measure of *better* or *worse* fits

* We also included the heavy target data of EMC and BCDMS. However, the larger systematic errors of these data sets led to relatively small impact on the overall results

experimental data exist. For this first phase of our study, we only include results of the Fermilab E288 [13] and E605 [14] experiments of scattering of proton on nuclear target for which information on point-to-point systematic errors have been obtained from the authors.* Other experiments with pion, kaon (as well as proton) beams do exist. However, we are interested only in parton distributions of the nucleon for the moment; and the lack of critical comparative studies of existing experiments in lepton-pair production makes combining data from different sources difficult.

We fit to the complete sets of published data on $d\sigma/dQ^2dy$ from E288 and E605 using the NLO formulas of [12]. (Note that the MRS group [2] only uses the Q^2-distribution.) We find that the inclusion of Drell–Yan data has an important impact on the global analysis. This is because the Drell–Yan cross-section in proton-nucleon scattering is particularly sensitive to the product of the u and $\bar{u}$ distributions. Since certain linear combinations of the u- and $\bar{u}$-quark distributions are relatively well-determined in DIS, the additional handle on the *product* is very useful in helping to differentiate the sea-quark ($\bar{u}$) from the valence quark distribution. In addition, as we are performing a next-to-leading order analysis, the gluon distribution comes in here in a direct way.

Because the strange and charm quarks make only a very small contribution to the cross-section, they are not well constrained. In principle, the Drell–Yan cross-section is sensitive to the normalization of the u- and d-anti-quark distributions; however, the experimental uncertainty on the relative normalization of the D–Y and DIS cross-sections essentially neutralizes this sensitivity. Hence, we gain only information on the shape of these distributions.

Thus, even with the addition of this new data, we still can not avoid making the simplifying assumption $A_2(\text{sea}) = A_2(\text{gluon}) + 1$. Likewise, in most of our fits we choose the same A_1 for the gluon and the sea-quarks; and, unless otherwise stated, assume $SU(3)$ flavor symmetry for the latter. Improved quality data from direct photon production and W- and Z-production as well as semi-inclusive deep inelastic scattering, such as charm-production, will eventually furnish independent information on the gluon and individual sea-quark distributions, and allow the unconstrained determination of these distributions. For the DIS and D–Y data used in our analysis, we consistently get good overall fits with these simplifying assumptions. In comparison to fits to DIS alone, as described before, these combined fits consistently favor a softer gluon-sea quark distribution characterized by $A_2^G (= A_2^{\text{sea}} - 1)$ of the order 6.5–7.5.

It may be tempting to determine the gluon distribution from existing data on direct-photon production in hadron collisions. We decided not to include this process in our global analysis at this stage because of the following considerations: (i) current experimental results are rather limited in accuracy, in statistics, and in x-range coverage compared to DIS and D–Y counterparts; (ii) the application of next-to-leading order QCD formalism to this process involve uncertainties which have yet to be clearly understood. However, noting that the often quoted gluon distribution from direct photon analyses [15] appears to be much harder than that mentioned above, we have studied fits to the DIS and D–Y data with the gluon shape parameter A_2^G decoupled from that of the sea quarks. In particular, we studied fits with this parameter fixed at 4. The resulting fits to the DIS F_2 data are comparable to our standard fits, but the χ^2 for the xF_3 and D–Y data sets are increased by 30–40% (cf. next section). We will comment more about the issue of hard vs. soft gluon distribution in the section on comparisons with other parton distribution sets.

5 Results of global fits

Because there is a wide range of possibilities on data-selection (Q^2-cut, W-cut, ···), error handling, and choice of shape parameters, it is possible to obtain a large number of good fits to the above mentioned experiments. In the course of this on-going study, we try to understand the systematics of these fits, and to identify physically relevant but unresolved features of the parton distributions. A comprehensive discussion of the details of these fits is outside the scope of this paper. We shall concentrate on: (i) the general features, (ii) a small set of representative results, and (iii) the most noteworthy consequences.

With our usual choice of Q^2-cut ($10\,\text{GeV}^2$), W-cut (4 GeV), and error handling (systematic and statistical added in quadrature), the global fits to the BCDMS H2 and D2, CDHSW, and the E288 and E605, data (referred to henceforth as the "B-fits") involve 647 data points. The overall χ^2/dof for these fits is on the order of 0.8 and evenly distributed among the data sets* – indicating a large degree of consistency among these different physical measurements in the QCD framework. Correspondingly, the global fits to the EMC H2 and D2, CDHSW, and the E288 and E605 data with the same choice of kinematic cuts and error handling, (referred to henceforth as the E-fits) involve 472 data points; the overall χ^2 per degree of freedom is typically around 0.93. The χ^2 of the individual data sets varies between 0.65–0.85 for all sets except for the EMC D2 data set where it is around 1.5. Representative of these B and E-fits, Fig. 1 shows the BCDMS hydrogen data with the B1-fit (solid line) as well as the E1-fit (dashed line); while Fig. 2 shows the EMC hydrogen data and Fig. 3 shows the CHHSW F_2 data with corresponding curves obtained in the same fits.

Finally, the global fits to ALL the data combined (referred to as the S-fits) involve 828 data points. The overall χ^2/dof range is 0.94–0.97; the χ^2/dof for the individual data sets are not as consistently distributed as

* We thank Chuck Brown and G.M. Moreno for discussions on this point

* We have added a 5–10% point-to-point systematic error to the published E288 and E605 data points which only show statistical errors. This estimate is adopted after consultation with members of these experiments (cf. previous footnote)

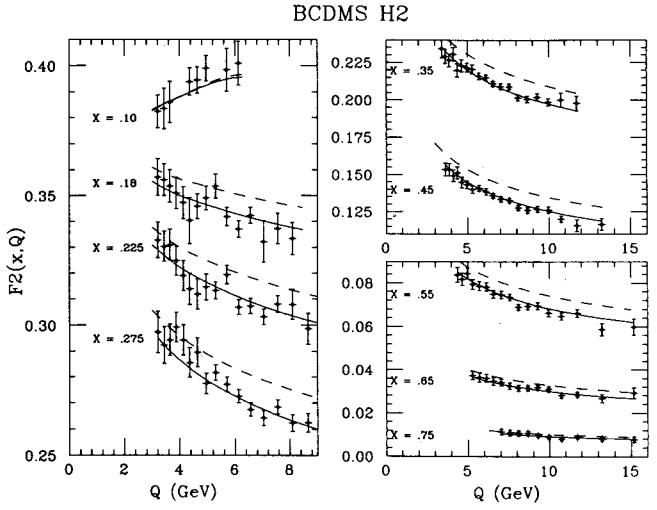

Fig. 1. Results of B1-fit (solid) and E-fit (dashed) compared to BCDMS H2 measurement of $F_2(x, Q)$

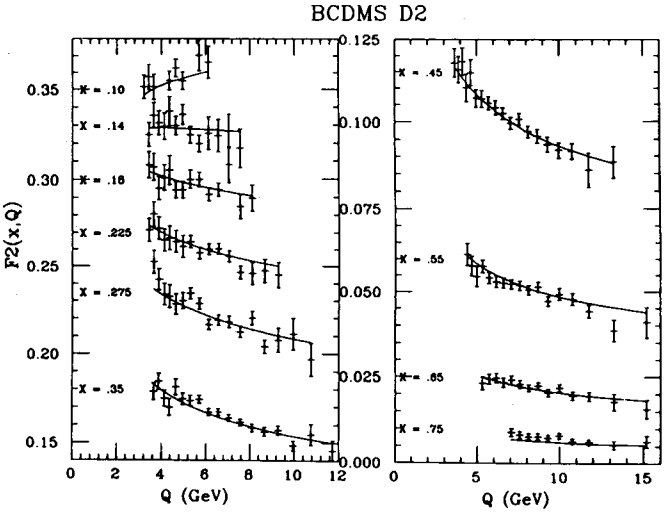

Fig. 4. Results of the S-fit compared to BCDMS D2 measurement of $F_2(x, Q)$

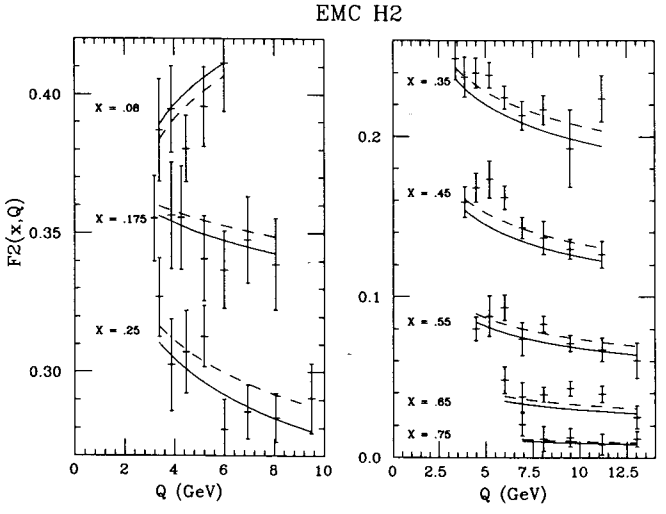

Fig. 2. Results of B1-fit (solid) and E-fit (dashed) compared to EMC H2 measurement of $F_2(x, Q)$

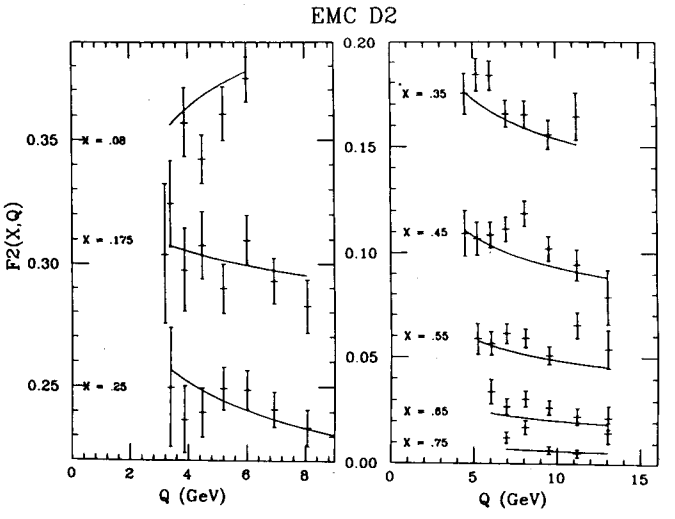

Fig. 5. Results of the S-fit compared to EMC D2 measurement of $F_2(x, Q)$

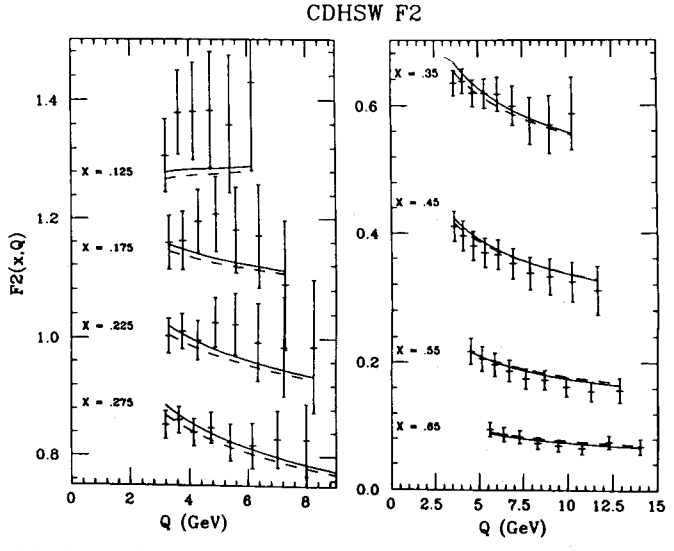

Fig. 3. Results of B1-fit (solid) and E-fit (dashed) compared to CDHSW iron measurement of $F_2(x, Q)$

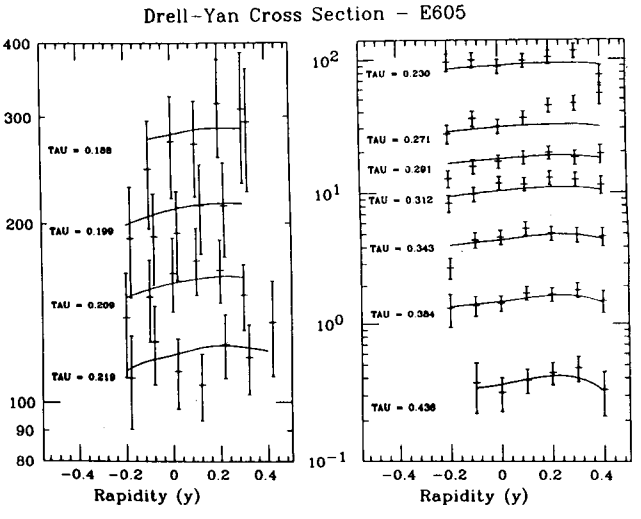

Fig. 6. Results of the S-fit compared to the E605 Drell–Yan cross section measurement

for the B and E fits with χ^2/dof for the EMC data sets about a factor of two higher than the rest. To illustrate the quality of these fits, we show one of them (S1 – solid line) in the comparison plots Fig. 4 (BCDMS-D_2), Fig. 5 (EMC-D_2), and Fig. 6 (E605-Drell–Yan). In all these plots, the experimental error bars represent the combined statistical and systematic errors.

Given the well-publicized "disagreement" between the EMC and BCDMS data sets, the conventional wisdom is that it is not possible to obtain any meaningful combined fit to these two recent muon experiments. In studying this issue we have found it important to consider (i) the explict inclusion of experimental systematic errors, omitted in most comparisons of these experiments, which considerably narrows the gap between the data points; (ii) an overall relative constant normalization between the experiments included in the fitting procedure which helps to bring the data sets in line for smaller values of x (where experimental errors are small) without introducing too much disagreement at larger x (where errors are big); and (iii) reasonable Q^2-cuts, imposed for the QCD fits (in order to exclude non-perturbative effects), which also tend to exclude the region of most severe disagreement between the experiments. The relative normalization factors obtained in these combined fits agree quite well with those put forth independently by recent critical comparisons and reviews of these experiments [16]. The quality of this combined fit can be questioned since, as mention previously, the χ^2 for the EMC data set is relatively high compared to those for the other data sets. However we refer the reader to Figs. 4–6 (especially Fig. 5 which yields the highest χ^2/dof among all data sets) to assess the quality of the data and the quality of this fit.

6 Comparisons with existing parton distributions

To compare our global fits to DIS and D–Y data with previously published sets of parton distribution functions, we have to bear in mind that some crucial data sets used in earlier analyses have been significantly revised (e.g. compare the 1983 CDHS data [17] with the new CDHSW results [5]); and that the very high statistics BCDMS muon data are not used by most existing published parton distribution sets. Thus, such distributions should not be expected to fit current accurate DIS data to within the experimental errors. Figure 7 illustrates this fact by comparing a representative group of BCDMS hydrogen data with the structure function F_2 calculated from the following parton distributions: our B1 set (dark-solid), EHLQ-1 (dashed), Duke–Owens-1 [1] (light-solid), MRSB [2] (dashdotted), and DFLM-NLLA [3] (dotted). Note that, of the last four sets, only MRSB used the BCDMS data in their analysis.* This plot illustrates that for QCD parton model studies requiring accuracy, the earlier well-known parton distribution sets

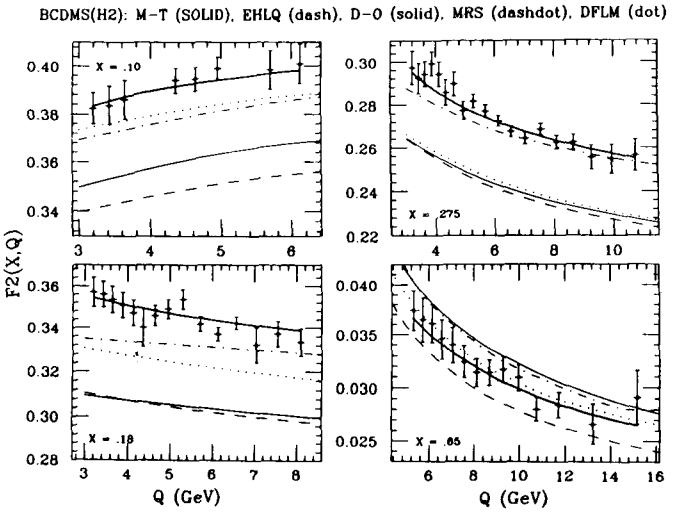

Fig. 7. Comparison of the EHLQ, Duke–Owens-2, MRS-B, DFLM, and present parametrizations of the parton distribution functions to BCDMS H2 measurements of F_2 at four representative x values

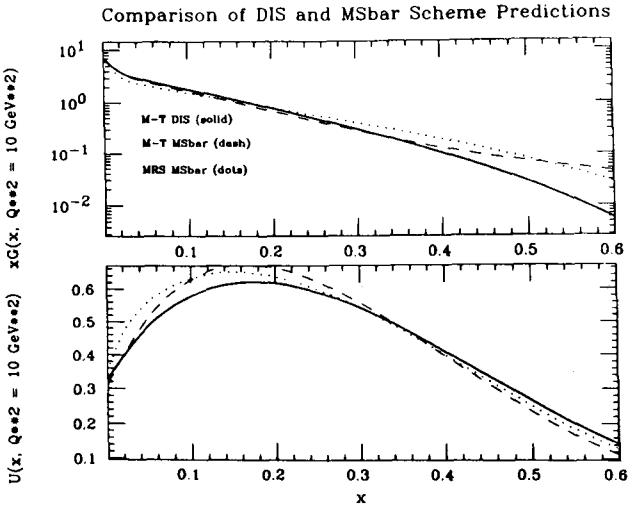

Fig. 8. Comparison of $xG(x)$ and $xu(x)$ as fit in the DIS scheme and as converted to the MS bar scheme

are no longer sufficient. The fact that the DFLM set was obtained without using the muon data also clearly shows in this plot.

Any direct comparison of distinct sets of parton distributions themselves must take into account the precise definition of the distribution function adopted as, in next-to-leading order of QCD, these quantities depend critically on the renormalization scheme used.* Of the two recent published analyses, the DFLM sets are in the so-called DIS scheme (in which the gluon contribution to the *total inclusive* F_2 structure function is, by definition, absorbed into the quark distributions), whereas the MRS

* Since the MRS distributions are given in the $\overline{\text{MS}}$ scheme, they are first converted into the DIS scheme before-substituting into the structure function formula along with the other sets in this comparision

* For recent reviews of this renormalization scheme dependence and related issues, see [4, 18]

sets are in the ("universal") $\overline{\text{MS}}$ scheme. The precise definitions of parton distributions defined in these two schemes are given in Appendix I. In our analysis, we use the DIS scheme distributions in the fitting process for the practical reason that the comparison with F_2 data, which dominate the fit, is made very simple. The results of these analyses, however, can be presented in *any scheme* with the proper transformation applied.

To illustrate the scheme-dependence of NLO parton distributions, we show in Fig. 8a a comparison of the gluon distribution at $Q^2 = 10\,\text{GeV}^2$ from our B1 fit in the DIS scheme (solid line), from the same fit in the $\overline{\text{MS}}$ scheme (dashed), and from the MRSB set (dotted) which is in the $\overline{\text{MS}}$ scheme. This plot illustrates the importance of specifying the scheme in order to make any meaningful discussion about "soft" or "hard" gluons. The conversion of the gluon distribution from one scheme to another necessarily turns a "soft" gluon distribution into a harder distribution, since the redefinition (in particular, the convolution integral of the quark distribution with a splitting function) involves re-interpreting the gluons radiating off the (hard) valence quarks. (Cf. Appendix I for detailed formulas.) Thus, our relatively "soft" gluon distribution in the DIS scheme becomes much harder after conversion into the $\overline{\text{MS}}$ scheme. The apparent discrepancy between our $A_2^G(6\text{–}7)$ parameter in the DIS scheme and the corresponding one adopted by MRS (~ 4) disappears after conversion into the same scheme. In this connection, we note that: it is natural to expect a soft gluon distribution in the DIS scheme, as the contribution from the gluon to F_2 is absorbed into that from the quarks by fiat; whereas in the $\overline{\text{MS}}$ scheme, the gluons radiating off the valence are indeed counted as gluon partons. In Fig. 8b we show the comparison of the corresponding curves for the u-quark distribution. It is evident that because of the re-interpretation of the partons engendered by the transformation between the two schemes, the u-quark distribution becomes slightly softer in the large-x region in the $\overline{\text{MS}}$ scheme. Note that the y-axis in Fig. 8b is in linear scale while that of Fig. 8a is in logarithmic scale. We add that the numerical difference between the distributions in the two renormalization schemes diminishes with increasing Q, becoming insignificant beyond $Q^2 = 100\,\text{GeV}^2$ or so.

Because of the significant difference between the same distribution presented in the two distinct schemes in some kinematic regions, users of these distributions must pay close attention to the scheme in which the distributions are defined and use them accordingly – the $\overline{\text{MS}}$ distributions must be used in conjunction with $\overline{\text{MS}}$ hard matrix elements (Wilson coefficient), and DIS distributions with DIS matrix elements, ··· etc. This important point is often ignored in the literature where comparisons are frequently made between cross-sections calculated with MRS and DFLM distributions (one in $\overline{\text{MS}}$ and the other in DIS scheme) convoluted with the same hard matrix elements. Figure 8 above shows explicitly how misleading such inconsistent use of parton distributions can become. We shall present our results in *both* schemes, so that our distributions can be used with hard matrix elements calculated in either scheme.

7 Behavior of parton distributions – reparametrization

One of the important motivations for adopting the functional form, (1) is that it is naturally suited to represent the parton distributions at any value of Q. Thus, although we must resort to rather involved numerical integration of the evolution equations during the fitting process, it is possible to re-express all the final parton distributions in this simple functional form. The QCD-evolution of the distribution functions then manifests itself in Q-dependent A-coefficients. Because the natural evolution variable is $\ln(\ln(Q))$, we can expect rather weak Q-dependence of these coefficients which are then easily parametrized by simple functions.

At the current level of accuracy, we found it possible to parametrize the parton distribution functions for all flavors and all Q in the same functional form (2) as used for the initial distributions. In Fig. 9 we show the Q-dependence of the A-coefficients for the various flavors from one of our parton distribution sets – the B1-fit. We see that a substantial number of these coefficients are almost linear in the natural evolution variable, whereas the rest can easily be represented by quadratic functions over the Q range from threshold to $10^4\,\text{GeV}$. This plot exhibits clearly the steady increase in the powers of $(1-x)$ and $\ln(1/x)$ as well as the decrease of the power of x and the normalization factor, all manifestations of the well-known softening of the parton distributions with increasing Q. Given the well behaved functional dependence, we represent our parton distributions in the form (2), and parametrize the A-coefficients for each parton flavor as:

$$A^i(Q) = C_0^i + C_1^i\,T(Q) + C_2^i\,T(Q)^2 \tag{3}$$

where $i = 0 - 3$, and

$$T(Q) = \ln\frac{\ln\dfrac{Q}{\Lambda}}{\ln\dfrac{Q_0}{\Lambda}}. \tag{4}$$

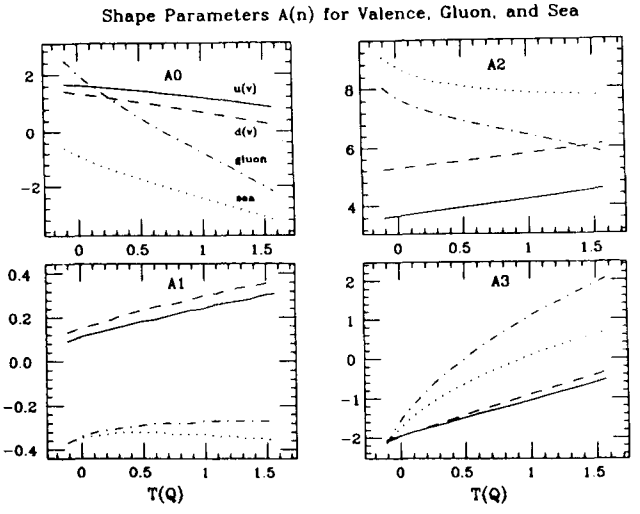

Fig. 9. The shape parameters A_i for the different partons as a function of $T(Q)$

Table 1a. Fit S – DIS scheme

$\Lambda(2,4) = 0.212\,\text{GeV}$					$Q_0^2 = 4\,\text{GeV}^2$				
	d(val)	u(val)	gluon	u(sea)	d(sea)	s	c	b	t
					A_0				
C_0	1.34	1.62	1.88	−0.99	−0.99	−0.99	−3.98	−6.28	−13.08
C_1	−0.57	−0.33	−2.78	−1.54	−1.54	−1.54	0.72	2.62	8.54
C_2	−0.08	−0.10	0.13	0.10	0.10	0.10	−0.63	−1.18	−2.70
					A_1				
C_0	0.15	0.11	−0.33	−0.33	−0.33	−0.33	−0.15	−0.18	−0.40
C_1	0.16	0.14	0.10	0.03	0.03	0.03	−0.06	0.02	0.31
C_2	−0.02	−0.01	−0.04	−0.03	−0.03	−0.03	0.00	−0.03	−0.12
					A_2				
C_0	5.30	3.68	7.52	8.53	8.53	8.53	7.46	6.56	15.35
C_1	0.43	0.53	−1.13	−1.08	−1.08	−1.08	0.96	1.40	−11.83
C_2	0.06	0.03	0.04	0.39	0.39	0.39	−0.30	−0.38	4.16
					A_3				
C_0	−1.96	−1.94	−1.34	−1.55	−1.55	−1.55	0.35	0.65	−0.43
C_1	1.08	0.87	2.92	2.02	2.02	2.02	0.89	1.13	3.18
C_2	−0.03	0.02	−0.49	−0.39	−0.39	−0.39	−0.04	−0.16	−0.82
Equivalent "conventional parametrization" coefficients at $Q_c^2 = 5.0\,\text{GeV}^2$ $f(x,Q_c^2) = e^{B_0} x^{B_1}(1-x)^{B_2}(1+B_3 x)$									
B_0	−0.49	−0.31	0.48	−2.65	−2.65	−2.65	0.00	0.00	0.00
B_1	0.43	0.36	−0.15	−0.14	−0.14	−0.14	0.00	0.00	0.00
B_2	5.36	3.70	8.02	9.58	9.58	9.58	0.00	0.00	0.00
B_3	10.68	11.82	8.20	13.60	13.59	13.59	0.00	0.00	0.00

The constant coefficients are determined by an overall fit to the particular parton distribution function over the range ($10^{-5} < x < 1, 3\,\text{GeV} < Q < 10^4\,\text{GeV}$). The resulting parametrization proves to be accurate to within the same degree as the original fit to data, thus it is a faithful representation of fitting results. This means each set of parton distributions is specified by a compact table of the C-coefficients. A typical table of such coefficients – that corresponding to the S fit – is given as Table 1a. In view of the discussion on the scheme dependence of the parton distributions, we also present the coefficients of the parametrization for the same fit in the $\overline{\text{MS}}$ scheme in Table 1b.*

Tables of coefficients from fits E, B1, and B2, mentioned in the text, are given in Appendix II. For the readers' convenience we present in Appendix III two additional fits which can be useful in various applications. The first one is a next-to-leading order fit of the combined data (SN-fit) which includes a *non-SU(3)-symmetric* sea (as suggested by some neutrino di-muon studies). The second one is a *leading order* fit of the combined data (SL-fit), which should be used in applications where leading order hard scattering matrix elements are employed.

Although our functional form (2) differs from conventional ones only in the $\ln(1/x)$ factor, this difference is quite significant. In particular, for the initial distributions, the values of A_1 and A_2 should not be compared directly with the corresponding parameters in published parton distributions sets. The $\ln(1/x)$ factor has a direct influence on the effective powers of x and $(1-x)$ in the small-x and large-x regions respectively. For the purpose of comparison, we include at the end of each Table a list of "equivalent conventional coefficients" B_0–B_3 which appear in the functional form*

$$x f(x, Q_1^2) = e^{B_0} x^{B_1}(1-x)^{B_2}(1+B_3 x) \qquad (5)$$

at $Q_1^2 = 5.0\,\text{GeV}^2$.

We have also tried to use the functional form (5) to parametrize the parton distribution functions for all Q with Q-dependent B-coefficients, as a possible alternative to our approach. We found, however, it is not possible to fit the gluon and sea-quark distributions to this form

* The gluon distribution transformed from the DIS to $\overline{\text{MS}}$ scheme is somewhat more difficult to parametrize in this compact form for all x. The parametrization is obtained by performing a fit for the region $x < 0.35$ which contains most of the gluons. The difference between the extrapolation beyond this region according to our functional form and the perturbative formula lies well within the overall uncertainty of our knowledge of $G(x)$

* The advantage of this form is that the parameters B_1 and B_2 are primarily responsible for the behavior of $f(x,Q)$ in the regions $x \sim 0$ and $x \sim 1$ respectively

Table 1b. Fit S – $\overline{\text{MS}}$ scheme

$\Lambda(2,4) = 0.212\,\text{GeV}$					$Q_0^2 = 4\,\text{GeV}^2$			
d(val)	u(val)	gluon	u(sea)	d(sea)	s	c	b	t
				A_0				
C_0 1.75	2.03	1.09	−0.14	−0.14	−0.15	−2.36	−2.19	−24.77
C_1 −1.02	−0.78	−2.41	−1.98	−1.98	−1.98	−1.42	−3.86	−23.00
C_2 0.05	0.03	−0.12	0.23	0.23	0.23	0.21	0.15	34.44
				A_1				
C_0 0.11	0.06	−0.24	−0.49	−0.49	−0.49	−0.49	−1.07	7.52
C_1 0.26	0.24	0.08	0.02	0.02	0.02	0.44	1.56	0.48
C_2 −0.06	−0.04	0.02	−0.02	−0.02	−0.02	−0.22	−0.73	−6.26
				A_2				
C_0 6.20	4.43	5.97	10.24	10.24	10.23	9.00	11.30	−99.51
C_1 −0.41	−0.18	−0.90	−1.43	−1.44	−1.44	−0.46	−7.20	−16.45
C_2 0.29	0.22	−0.35	0.44	0.45	0.45	0.29	3.85	97.19
				A_3				
C_0 −2.35	−2.35	−0.64	−2.57	−2.57	−2.57	−1.74	−4.85	36.02
C_1 1.68	1.52	2.71	2.32	2.32	2.32	3.93	10.51	16.51
C_2 −0.24	−0.19	−0.20	−0.47	−0.47	−0.47	−1.34	−4.36	−40.40
Equivalent "conventional parametrization" coefficients at $Q_c^2 = 5.0\,\text{GeV}^2$ $f(x,Q_c^2) = e^{B_0} x^{B_1}(1-x)^{B_2}(1+B_3 x)$								
B_0 0.03	0.23	0.68	−2.26	−2.26	−2.26	0.00	0.00	0.00
B_1 0.53	0.46	−0.14	−0.10	−0.10	−0.10	0.00	0.00	0.00
B_2 6.08	4.33	5.88	10.12	10.11	10.11	0.00	0.00	0.00
B_3 7.96	8.93	0.43	11.40	11.39	11.39	0.00	0.00	0.00

with any reasonable degree of accuracy for Q values beyond about 50 GeV. The contrast with the parametrization (2) in this regard clearly support our original expectation that the latter is naturally suited to represent the QCD-evolved parton distributions, at least in the perturbative framework.

We summarize the distinctive advantages of the Q-dependent parametrization of the parton distribution functions (2) compared to conventional ones based on expansions in terms of polynomials (such as the above) or other orthogonal functions (such as Chebyshev polynomials, used by EHLQ): (i) It is *compact* – the maximum size of the table of C-coefficients is a $3 \times 4 \times 9$ matrix, for the case of 6 quark flavors with no symmetry assumed for the sea. The size is smaller for less number of active flavors and/or with any assumptions on symmetry of the sea-quarks. (ii) The parton distributions are always *positive definite* for all values of (x, Q). (iii) These functions are *smoothly varying* in both x and Q, thus never lead to pathological behaviors even when they are used (intentionally or inadvertently) outside the original range – as often happens in applications of parton distributions to very high energy processes over some part of the phase space integration. (iv) The functional form is ideally suited to explore the *small-x behavior of the parton distributions* – an area of central importance for application of the QCD parton model framework to current and future high energy processes.

8 Range of validity of distributions

Strictly speaking, the distribution functions presented in this paper, as with all other published distributions, are valid only within the range of variables (x, Q) covered by the data sets used in the global analysis. This means, for the current analysis, approximately, $0.03 < x < 0.75$ and $5 < Q^2(\text{GeV}^2) < 250$.* However, the most important feature of the QCD parton model is that it allows us to use the parton distribution functions determined at current energy scales to make predictions on all hard processes at much higher energies and shorter distances. In particular, QCD can reliably predict the Q-dependence of these distributions through the renormalization group (or evolution) equation. Thus, based on perturbative QCD, these distribution functions can be trusted to very large values of Q provided the evolution is done correctly.

Since we present our results in the form of a parametrization of the evolved distributations, one should also ask the range of applicability and accuracy of this parameterization. We obtain our parametrizations by fitting the numerically evolved distributions to

* Although our normal cut is $10 < Q^2(\text{GeV}^2)$ in data-fitting, we have verified that the predictions extrapolated to $5 < Q^2(\text{GeV}^2)$ still fit existing data rather well

the adopted functional form over the range $5 < Q^2(\mathrm{GeV}^2) < 10^8$ to an accuracy comparable to that of the experimental data used – 1–2% for the dominant distributions at small and moderate x, increasing to bigger errors toward the large x region (where experimental errors increase) and for the numerically small sea-quarks, especially the heavy flavor ones.

A separate question is the applicability of these distributions at smaller Q values, say 1–2 GeV. Here, the real issue is the applicability of the QCD parton model itself in a region of where its theoretical basis – the factorization theorem (an asymptotic theorem) – is obviously questionable. Because this involves physics at the confinement scale, we cannot make any definitive statement on this issue. Phenomenologically, our parametrization is smooth and well behaved above $Q = 2$ GeV. Our recommendation is that, for these low values of Q, cross-sections calculated with our distributions should be regarded only as *extrapolated twist-2* QCD *contributions* which may or may not require additional terms due to higher twist (or non-perturbative) contributions before they can be directly compared to experimentally measured quantities. These additional contributions have been recently examined phenomenologically [19] down to $Q^2 = 0.5\,\mathrm{GeV}^2$ and appear to be quite small for $0.05 < x < 0.4$, becoming more important as x increases to 1.

Likewise, not much is known definitively about extrapolating parton distributions into the small- and large-x regions. Although there is intense current interest in these issues – particularly, the small-x behavior [20], there is still no conclusive theory comparable to that on Q-dependence. Under this circumstance, we take a pure phenomenological approach to this issue and use the functional form adopted to *explore* the range of predicted small-x behavior of physically interesting process (cf. next section). As explained before, our functional form is consistent with known theoretical understanding (for details, see [21]) and it is more flexible than that used by other published parton distribution sets.

9 The small-x extrapolation

A strong motivation for undertaking the task of this global fit project is to systematically explore our *lack* of knowledge of the parton distributions at low x. Most of published parton distribution sets use some assumed B_1 parameters (cf. (5)) for the gluon and sea-quark distributions at a given Q_0. The small-x behavior of the parton distributions so obtained depend sensitively to the values of B_1 and to Q_0 so chosen (see below). Predictions on processes at present ($S\bar{p}pS$, Tevatron) and future (HERA, SSC) accelerators often rely on parton distributions at small x values much below those currently measured, hence they depend critically on implicit assumptions made about the functional form of these distributions at an arbitrarily chosen Q_0 – a fact not always noted by all users of these distributions.

In this section we explore the small-x behavior of the parton distributions, which are consistent with current data, in two different ways which distinguish our approach from previous efforts. First, we leave the parameter A_1 (cf. (1)) for the gluon and sea-quarks as a free parameter in the data analysis, hence its value (at a fixed value Q_0) is determined by the data rather than by an arbitrary assumption. Since the effective power A_1 changes rapidly with Q in the relatively low Q region where evolution starts, any assumption one makes is highly dependent on the choice of Q_0. Our method does not prejudice this choice. Secondly, by introducing a logarithmic factor $(\ln x)_3^A$ in the functional form (1), we allow for the possibility of logarithmic extrapolation to the small-x region in addition to the traditional power-law extrapolation. This is logical, as the evolution equation naturally introduces logarithmic dependences of the parton distributions even if one starts with a pure power-law function.

For a given selection of data sets we routinely perform fits with the A_1 factor alone, with the A_3 factor alone, and with both as fitting parameters. Since available data in DIS and D–Y processes involve a limited range in x, we are able to get good fits in all three cases. Within the (x, Q) range of current experiments the resulting parton distribution sets yield very similar DIS structure functions and D–Y cross-sections; but they lead to different predictions far away from this range, especially for very small x. In this way, we can study the *range* of small-x behavior of parton distributions allowed by current data in a systematic and quantitative way.

For illustration, in Fig. 10a we plot the structure function F_2 and the gluon distribution at $Q_2 = 10\,\mathrm{GeV}^2$ in the x-range $(10^{-5}, 10^{-1})$. The two representative parton distribution sets "B1" and "B2" both fit the existing data ($x > 0.03$) but they have different $A_1 - A_3$ exponents which give rise to quite different predictions in the $x < 0.03$ range. In Fig. 10b the same quantities are plotted at $Q^2 = 10^4\,\mathrm{GeV}$. As expected, there is a migration of the partons to small x caused by the Q^2 evolution, so that differences are reduced as Q^2 increases. In order to explicitly display the uncertainty on the small-x behavior

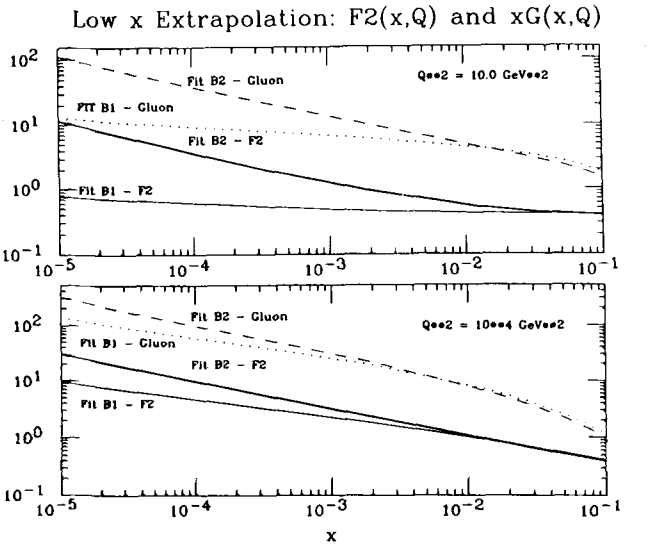

Fig. 10. Predicted values of $F_2(x)$ and $xG(x)$ from fits B1 and B2 at ultra low x for $Q^2 = 10\,\mathrm{GeV}^2$ and $Q^2 = 10^4\,\mathrm{GeV}^2$

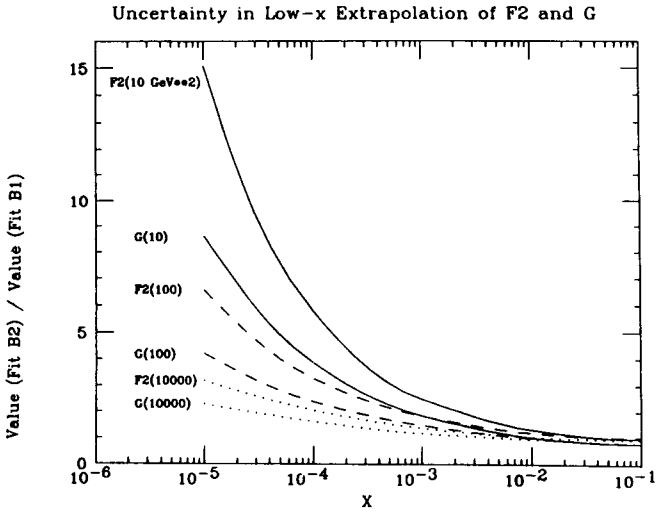

Fig. 11. The ratio of predictions (the uncertainty) for $F_2(x)$ and $xG(x)$ as a function of x for three typical values of Q^2

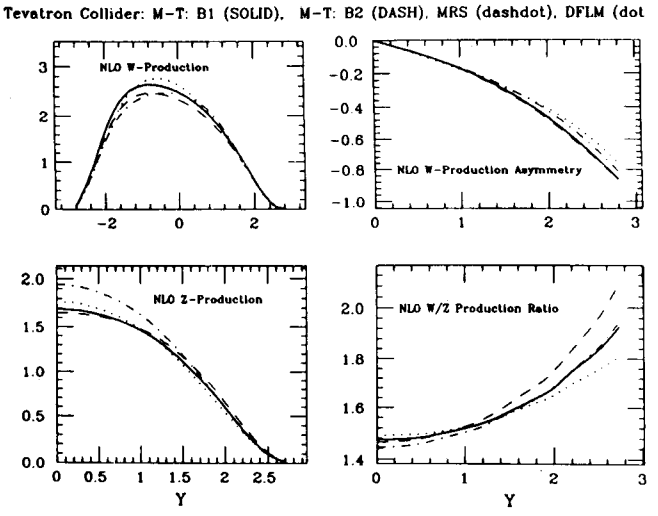

Fig. 12. The predictions for W-production, Z-production, the W-production asymmetry, and the W/Z production ratio at the Tevatron Collider using the parton distribution function of MRS-B, DFLM, and the B1- and B2-fits from the present analysis

associated with these two equally acceptable fits, we plot in Fig. 11 the ratios of corresponding $F_2(x, Q)$ and $G(x, Q)$ obtained from the two fits at three values of Q over the x range as in the previous figure. We see that parton densities, and physical cross-sections derived from them, at $x = 10^{-5}$ can differ by factors of 2–3 at the highest Q^2 and by an order-of-magnitude at more moderate Q^2. Since the HERA experiments are expected to measure the structure functions down to $x = 10^{-4}$, Figs. 10 and 11 illustrate how these experiments can contribute to narrow the uncertainties as they exist now. Before these distributions are measured at HERA, "predictions" on cross-sections for various processes at SSC and LHC which depend on parton distributions at small x have to be considered in the context of the uncertainties described here.

We note that the two fits used above are chosen for illustrative purposes only. They do not necessarily represent the full range of behavior allowed by current data. Detailed study focusing on this question will be pursued and reported elsewhere [21].

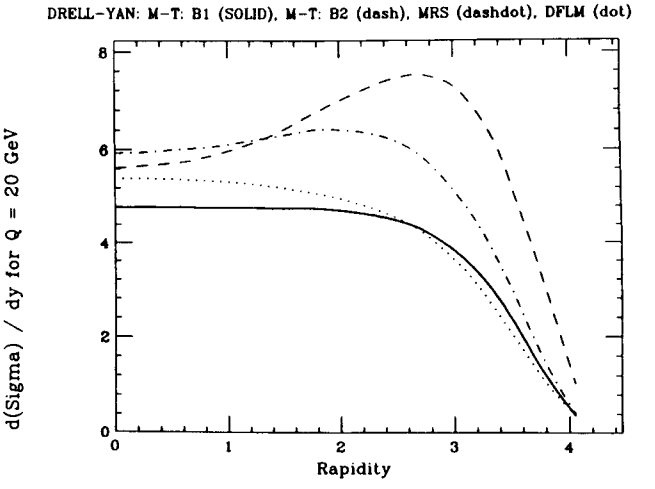

Fig. 13. Prediction for low mass ($Q = 20$ GeV) Drell–Yan pair production at the Tevatron collider for parton distribution functions as in Fig. 12

10 *W*-, *Z*-, and lepton-pair production cross-sections and ratios at collider energies

The real value of parton distributions lies in its universal applicability to all high energy processes in the QCD framework. Of immediate interest is the use of parton distributions extracted from fixed-target DIS and D–Y experiments to "predict" cross-sections of physical processes at current colliders. Since collider processes usually depend on combinations of parton distributions which are different from the fixed-target processes, and since the x-values involved may be beyond the original range, these new cross-sections can also be used to provide important constraints on the x- and the flavor-dependence of the parton distributions not otherwise available. We illustrate this point by presenting W-, Z-, and lepton-pair production cross-sections at current collider energies based on our parton distributions. All results are obtained in next-to-leading order QCD calculation.

Figure 12a shows the differential cross-section in rapidity for W^+ production, Fig. 12b shows that for Z production, Fig. 12c shows the y-asymmetry in W^+ production, and Fig. 12d shows the W/Z cross-section ratio in $p - \bar{p}$ collision at 1.8 GeV. The curves are obtained using the following parton distributions: B1-fit (solid line), B2-fit (dashed), MRSB (dot-dashed)*, and

* The next-to-leading order hard matrix element available is in the DIS scheme, whereas the MRS distributions are in the $\overline{\text{MS}}$ scheme. Unlike in the comparison section for DIS, we did not make the necessary conversion here because of the large amount of computing time required

DFLM-NLLA (dotted). We see that the spread of these curves is not large, except for the W/Z ratio at high values of y. This is understandable since the cross-section for W- and Z-production involve parton momentum fractions x well within the range measured by the fixed target DIS and D–Y experiments. This spread is, however, bigger than the anticipated experimental uncertainty on these quantities in current and near-future runs. Thus, these precise measurements will contribute very useful independent information on the flavor-differentiation of parton distributions described in earlier sections.

In Fig. 13 we show next-to-leading order calculation of the y-distribution of lepton-pairs (D–Y) at the Tevatron energy for dimuon mass $Q = 20\,\text{GeV}$ using the same parton distributions as above. Here we see a dramatic difference, especially between the prediction of the B2-fit distributions and the rest at high y values. This sensitivity is due to the contribution of the small-x parton distributions to the D–Y cross-section – especially in the forward–backward directions. This striking effect has been known for some time, based on crude inputs [22]. The current calculation, using parton distributions known to be consistent with all current experiments, underlines the importance of the collider lepton-pair measurements in probing parton distributions at small-x.

Acknowledgements. The authors would like to express their sincere gratitude to members of the 1988 snowmass working group on structure functions and parton distributions for stimulating discussions and, in particular, to Heidi Schellman and Shuichi Kunori for valuable assistance in assembling the DIS data base.

Appendix I

Definition of parton distributions in the $\overline{MS}$ and DIS schemes

We have mentioned in the text that parton distribution functions $f^a(x,\mu)$ are renormalization scheme dependent beyond the leading order. In applications to various physical processes, the scheme chosen for the parton distributions must match that for the hard cross-section in the QCD parton model formula [18]. The same parton distribution in two different schemes differ by a well-defined expression which is *nominally* of one order higher in α_s. (Exceptions to this rule of thumb do exist, as will be mentioned later.) In this appendix, we define parton distributions used in the two schemes discussed in the text.

The $\overline{\text{MS}}$ scheme is defined by an "universal" prescription to facilitate perturbative calculations independent of any physical process. It is used by most theorists in the calculation of hard matrix elements. The $\overline{\text{MS}}$ parton distributions are guaranteed by their definition to satisfy the momentum sum rule. In this scheme, the NLO formula for the F_2 structure function* of virtual γ deep inelastic scattering reads:

$$F_2^{\gamma}(x,Q) = f_{\overline{\text{MS}}}^{q} \otimes [C_{2,q}^{(0)} + \alpha_s C_{2,q}^{(1)\overline{\text{MS}}}] + f_{\overline{\text{MS}}}^{G} \otimes \alpha_s C_{2,G}^{(1)\overline{\text{MS}}} + O(\alpha_s^2) \quad (1)$$

where $C^{(i)}$, $i = 0, 1$ are the standard hard matrix elements in LO and NLO often called the Wilson coefficients [9, 23]. On the right-hand side of this equation a sum over the quark flavor index is understood.

The "DIS" scheme [24], on the other hand, was defined *specifically* to make the relation between the parton distributions and F_2^{γ} as simple as possible. This is obtained from the above equation by absorbing all the NLO terms into the *definition* of f_{DIS}^{q}:

$$F_2^{\gamma}(x,Q) = f_{\text{DIS}}^{q} \otimes C_{2,q}^{(0)} + O(\alpha_s^2). \quad (2)$$

Thus the difference between the quark distributions in the DIS scheme and the $\overline{\text{MS}}$ scheme is:

$$f_{\text{DIS}}^{q}(x,Q) - f_{\overline{\text{MS}}}^{q}(x,Q) = \alpha_s (f^q \otimes C_{2,q}^{(1)\overline{\text{MS}}} + f^G \otimes C_{2,G}^{(1)\overline{\text{MS}}}). \quad (3)$$

No explicit label is given to the parton distributions on the right-hand side since these terms are of one order higher in α_s, thus either scheme will do.

Equation (2) does not yet define the gluon distribution in the DIS scheme. It is conventional to require that the momentum sum rule be preserved in the DIS scheme as well. This requirement fixes the second moment of the gluon distribution only. To complete the definition of the gluon distribution, it is convenient to generalize the condition on the second moment to all moments [3]. This is the definition we adopt. We then obtain:

$$f_{\overline{\text{MS}}}^{G}(x,Q) - f_{\text{DIS}}^{G}(x,Q) = \alpha_s (f^{q_s} \otimes C_{2,q}^{(1)\overline{\text{MS}}} + f^G \otimes C_{2,G}^{(1)\overline{\text{MS}}}) \quad (4)$$

where q_s denotes the singlet quark distribution and, again, the scheme label is dropped on the right-hand side.

These equations allow us to convert parton distributions from one scheme to the other. Thus, in principle, one can perform the calculation in either scheme – consistently – and then convert the results to the other scheme if necessary. We followed this procedure in the text of this paper. It is worth pointing out, however, that caution must be exercised under certain conditions in practical applications of this formalism. We encountered one such circumstance in Sect. 6 when we compared the gluon distributions from different parton distribution sets and from the same set in two different schemes. Figure 8 showed a significant difference between the same distribution in the two different schemes at the high end of x. Let us see how this can be understood from (4). *Nominally*, both terms on the right-hand side of the equation are of *one higher order* in α_s then the individual terms on the left-hand side – only the *difference* between the two is expected to be small. *However in reality*, one expects the (valence) quark distribution to be much harder than the gluon distribution. Thus, at large x, the quark term on the right-hand side can become just as

* The precise definition of F_2 has evolved with time, causing confusion sometimes. We use the definition of [23] (cf. also [4])

big or even bigger than the softer of the two gluon distributions on the left-hand side in spite of the extra α_s. When this happens, the equality forces the other term on the left-hand side to be relatively hard! This is precisely what we found in Fig. 8.

Appendix II

Following are the parton distribution function sets E, B1, and B2 discussed in the main body of this report. The parton distributions are given both in the DIS and

Table 3. Fit E – DIS scheme

$\Lambda(2,4) = 0.155\,\text{GeV}$					$Q_0^2 = 4\,\text{GeV}^2$				
	d(val)	u(val)	gluon	u(sea)	d(sea)	s	c	b	t
					A_0				
C_0	1.43	1.69	2.11	−0.84	−0.84	−0.84	−3.87	−6.09	−12.56
C_1	−0.65	−0.33	−3.01	−1.65	−1.65	−1.65	0.85	2.81	8.69
C_2	−0.08	−0.11	0.18	0.12	0.12	0.12	−0.73	−1.34	−2.93
					A_1				
C_0	0.16	0.11	−0.33	−0.32	−0.32	−0.32	−0.05	−0.17	−0.38
C_1	0.16	0.14	0.10	0.02	0.02	0.02	−0.07	0.01	0.30
C_2	−0.02	−0.01	−0.04	−0.03	−0.03	−0.03	0.00	−0.03	−0.12
					A_2				
C_0	6.17	3.69	7.93	8.96	8.96	8.96	7.83	6.75	14.62
C_1	0.43	0.54	−1.40	−1.24	−1.24	−1.24	1.00	1.74	−11.27
C_2	0.06	0.03	0.09	0.45	0.45	0.45	−0.36	−0.56	4.29
					A_3				
C_0	−1.94	−1.99	−1.51	−1.70	−1.70	−1.70	0.21	0.54	−0.41
C_1	1.12	0.90	3.14	2.15	2.15	2.15	0.93	1.15	3.19
C_2	−0.02	0.02	−0.55	−0.43	−0.43	−0.43	−0.03	−0.16	−0.87

Table 4. Fit E – $\overline{\text{MS}}$ scheme

$\Lambda(2,4) = 0.155\,\text{GeV}$					$Q_0^2 = 4\,\text{GeV}^2$				
	d(val)	u(val)	gluon	u(sea)	d(sea)	s	c	b	t
					A_0				
C_0	1.79	2.12	1.58	−0.10	−0.10	−0.11	−2.53	−3.91	−6.57
C_1	−1.05	−0.85	−2.68	−2.29	−2.29	−2.29	−1.16	−0.19	1.15
C_2	0.03	0.07	0.01	0.35	0.35	0.35	0.12	−0.24	−0.48
					A_1				
C_0	0.12	0.02	−0.28	−0.43	−0.43	−0.43	−0.35	−0.44	−0.90
C_1	0.24	0.32	0.05	0.09	0.09	0.09	0.26	0.38	0.95
C_2	−0.04	−0.08	0.00	−0.06	−0.06	−0.06	−0.15	−0.17	−0.33
					A_2				
C_0	7.03	4.46	6.84	10.43	10.43	10.43	8.67	6.85	7.27
C_1	−0.38	−0.28	−0.93	−2.14	−2.14	−2.14	−0.10	2.15	−0.28
C_2	0.27	0.29	−0.26	0.73	0.73	0.73	0.27	−0.74	0.28
					A_3				
C_0	−2.29	−2.57	−1.08	−2.49	−2.49	−2.48	−1.24	−1.56	−5.07
C_1	1.63	1.82	2.76	2.80	2.80	2.80	3.26	4.07	9.02
C_2	−0.18	−0.33	−0.32	−0.67	−0.67	−0.67	−1.06	−1.24	−2.75

$\overline{\text{MS}}$ renormalization schemes. Note that Fit B1 and B2 are representative of the variation in low-x extrapolation allowed by the currently available data.

As a reminder, the general expression for each parton flavor is:

$$x f^a(x,Q) = e^{A^a_0} x^{A^a_1} (1-x)^{A^a_2} \ln^{A^a_3}\left(1+\frac{1}{x}\right) \tag{1}$$

where the shape parameters are defined as:

$$A^i(Q) = C^i_0 + C^i_1\, T(Q) + C^i_2\, T(Q)^2 \tag{2}$$

with $i = 0$–3, and

$$T(Q) = \ln\frac{\ln\frac{Q}{\Lambda}}{\ln\frac{Q_0}{\Lambda}}. \tag{3}$$

Table 5. Fit B1 – DIS scheme

$\Lambda(2,4) = 0.194\,\text{GeV}$					$Q_0^2 = 4\,\text{GeV}^2$				
	d(val)	u(val)	gluon	u(sea)	d(sea)	s	c	b	t
					A_0				
C_0	1.30	1.59	1.48	−1.08	−1.08	−1.08	−4.22	−6.42	−12.92
C_1	−0.57	−0.34	−2.49	−1.33	−1.33	−1.33	0.88	2.67	8.33
C_2	−0.09	−0.10	0.04	−0.03	−0.03	−0.03	−0.69	−1.21	−2.68
					A_1				
C_0	0.19	0.14	−0.14	−0.13	−0.13	−0.13	−0.02	−0.09	−0.36
C_1	0.15	0.13	−0.11	−0.21	−0.21	−0.21	−0.17	−0.03	0.32
C_2	−0.02	−0.01	0.03	0.06	0.06	0.06	0.03	−0.02	−0.13
					A_2				
C_0	5.24	3.65	6.75	8.40	8.39	8.39	7.29	6.47	15.74
C_1	0.44	0.53	−0.54	−0.51	−0.50	−0.50	1.08	1.39	−12.73
C_2	0.05	0.03	−0.15	0.07	0.07	0.07	−0.39	−0.42	4.51
					A_3				
C_0	−1.81	−1.81	−0.50	−0.88	−0.88	−0.88	0.90	1.03	−0.30
C_1	1.06	0.86	2.13	1.18	1.18	1.18	0.50	1.00	3.35
C_2	−0.02	0.02	−0.24	−0.05	−0.05	−0.05	0.08	−0.14	−0.91

Table 6. Fit B1 – $\overline{\text{MS}}$ scheme

$\Lambda(2,4) = 0.194\,\text{GeV}$					$Q_0^2 = 4\,\text{GeV}^2$				
	d(val)	u(val)	gluon	u(sea)	d(sea)	s	c	b	t
					A_0				
C_0	1.66	2.00	0.92	−0.60	−0.60	−0.60	−2.94	−2.95	−3.88
C_1	−0.94	−0.81	−2.28	−1.76	−1.76	−1.76	−1.12	−3.21	−1.59
C_2	0.03	0.05	−0.07	0.13	0.13	0.14	0.15	1.38	−0.05
					A_1				
C_0	0.18	0.09	−0.07	−0.13	−0.13	−0.13	−0.19	−0.62	−0.78
C_1	0.18	0.24	−0.16	−0.27	−0.27	−0.27	0.16	0.99	−0.07
C_2	−0.03	−0.05	0.06	0.09	0.09	0.09	−0.13	−0.51	0.40
					A_2				
C_0	6.04	4.40	5.79	9.31	9.31	9.31	7.94	9.97	3.80
C_1	−0.25	−0.20	−0.68	−0.94	−0.94	−0.94	−0.05	−6.33	2.13
C_2	0.23	0.25	−0.23	0.21	0.21	0.21	0.27	3.71	0.96
					A_3				
C_0	−2.09	−2.24	−0.01	−1.18	−1.18	−1.18	−0.46	−3.00	−2.37
C_1	1.42	1.53	1.93	1.31	1.31	1.31	2.93	8.42	0.48
C_2	−0.14	−0.23	−0.11	0.10	−0.10	−0.10	−1.05	−3.61	2.30

Table 7. Fit B2 – DIS scheme

$\Lambda(2,4) = 0.191\,\text{GeV}$					$Q_0^2 = 4\,\text{GeV}^2$				
	d(val)	u(val)	gluon	u(sea)	d(sea)	s	c	b	t
					A_0				
C_0	1.38	1.64	1.52	−0.85	−0.85	−0.85	−3.74	−6.07	−12.08
C_1	−0.59	−0.33	−2.71	−1.43	−1.43	−1.43	0.21	2.33	7.31
C_2	−0.08	−0.10	0.15	−0.03	−0.03	−0.03	−0.50	−1.15	−2.35
					A_1				
C_0	0.18	0.09	−0.72	−0.82	−0.82	−0.82	−0.58	−0.52	−0.73
C_1	0.16	0.14	0.45	0.35	0.35	0.35	0.24	0.22	0.54
C_2	−0.02	−0.01	−0.15	−0.09	−0.10	−0.10	−0.07	−0.07	−0.18
					A_2				
C_0	5.40	3.74	7.75	9.19	9.19	9.19	9.63	8.33	21.14
C_1	0.42	0.54	−1.56	−0.92	−0.92	−0.92	−1.13	0.28	−19.17
C_2	0.06	0.03	0.16	0.12	0.12	0.12	0.25	−0.28	6.64
					A_3				
C_0	−1.91	−2.02	−2.18	−2.76	−2.76	−2.76	−1.09	−0.52	−1.92
C_1	1.11	0.88	3.75	2.56	2.56	2.56	2.10	1.91	4.59
C_2	−0.03	0.02	−0.76	−0.40	−0.40	−0.40	−0.33	−0.31	−1.25

Table 8. Fit B2 – $\overline{\text{MS}}$ scheme

$\Lambda(2,4) = 0.191\,\text{GeV}$					$Q_0^2 = 4\,\text{GeV}^2$				
	d(val)	u(val)	gluon	u(sea)	d(sea)	s	c	b	t
					A_0				
C_0	1.77	2.04	0.74	−0.43	−0.43	−0.43	−3.07	−4.44	−7.03
C_1	−0.98	−0.75	−2.44	−1.96	−1.96	−1.96	−1.03	−0.13	1.10
C_2	0.03	0.02	0.07	0.20	0.20	0.20	0.04	0.23	−0.41
					A_1				
C_0	0.13	0.03	−0.59	−0.86	−0.86	−0.86	−0.66	−0.68	−1.13
C_1	0.23	0.26	0.42	0.43	0.43	0.43	0.45	0.50	1.07
C_2	−0.04	−0.05	−0.15	−0.14	−0.14	−0.14	−0.17	−0.18	−0.35
					A_2				
C_0	6.28	4.48	6.31	10.16	10.16	10.16	8.57	6.90	8.56
C_1	−0.34	−0.15	−1.62	−1.91	−1.91	−1.91	−0.32	1.46	−2.33
C_2	0.26	0.21	0.18	0.53	0.53	0.53	0.17	−0.53	0.87
					A_3				
C_0	−2.30	−2.47	−1.37	−3.14	−3.14	−3.14	−1.68	−1.82	−5.47
C_1	1.60	1.52	3.56	3.14	3.14	3.14	3.48	4.11	9.08
C_2	−0.18	−0.19	−0.77	−0.68	−0.68	−0.68	−0.98	−1.16	−2.66

Appendix III

Table 9 (DIS scheme) and 10 ($\overline{\text{MS}}$ scheme) represent a next-to-leading order fit (SN-fit) of the combined data which assumes a *non-SU(3)-symmetric sea* as suggested by some neutrino di-muon studies. The ratio of $2s/(u+d)$ is set at 0.50 for the input distributions. Table 11 represents a *leading order* fit of the combined data (*SL*-fit) which should be used in applications where leading order hard scattering matrix elements are employed.

Table 9. Fit SN – non-symmetric sea – DIS

$\Lambda(2,4) = 0.237\,\text{GeV}$					$Q_0^2 = 4\,\text{GeV}^2$			
d(val)	u(val)	gluon	u(sea)	d(sea)	s	c	b	t
				A_0				
C_0 1.42	1.68	0.90	−1.48	−1.48	−2.26	−4.68	−6.83	−14.41
C_1 −0.59	−0.33	−1.86	−0.89	−0.89	−0.90	0.92	2.68	9.65
C_2 −0.08	−0.10	−0.09	−0.12	−0.13	−0.06	−0.62	−1.13	−2.98
				A_1				
C_0 0.16	0.08	−0.17	−0.13	−0.13	−0.15	−0.06	−0.12	−0.28
C_1 0.17	0.15	−0.10	−0.19	−0.19	−0.10	−0.12	−0.01	0.15
C_2 −0.02	−0.01	0.02	0.04	0.04	0.01	0.01	−0.03	−0.06
				A_2				
C_0 5.40	3.75	5.27	7.83	7.83	7.47	5.55	5.24	11.48
C_1 0.41	0.53	0.43	−0.06	−0.05	−0.61	1.16	1.14	−7.50
C_2 0.06	0.03	−0.26	0.01	0.00	0.28	−0.26	−0.24	2.54
				A_3				
C_0 −1.99	−2.09	−0.20	−0.38	−0.38	−0.23	1.13	1.19	0.65
C_1 1.12	0.89	1.67	0.68	0.68	1.22	0.50	0.93	1.99
C_2 −0.03	0.02	−0.14	0.05	0.05	−0.16	0.03	−0.13	−0.43

Table 10. Fit SN – non symmetric sea

$\Lambda(2,4) = 0.237\,\text{GeV}$					$Q_0^2 = 4\,\text{GeV}^2$			
d(val)	u(val)	gluon	u(sea)	d(sea)	s	c	b	t
				A_0				
C_0 1.84	2.08	0.31	−1.13	−1.13	−1.82	−3.69	−5.06	−9.92
C_1 −0.97	−0.66	−1.84	−1.26	−1.26	−1.40	−0.47	0.39	4.60
C_2 0.03	−0.02	−0.06	−0.01	−0.01	0.09	−0.10	−0.35	−1.53
				A_1				
C_0 0.12	0.02	−0.10	−0.15	−0.15	−0.18	−0.15	−0.25	−0.38
C_1 0.22	0.19	−0.10	−0.16	−0.16	−0.06	0.04	0.16	0.24
C_2 −0.04	−0.01	0.01	0.03	0.03	−0.01	−0.05	−0.08	−0.08
				A_2				
C_0 6.34	4.53	4.18	8.43	8.43	7.94	5.72	4.42	−1.27
C_1 −0.34	−0.04	0.05	−0.39	−0.39	−0.82	0.93	2.38	9.17
C_2 0.25	0.15	−0.12	0.30	0.30	0.05	−0.11	−0.63	−2.88
				A_3				
C_0 −2.40	−2.51	0.34	−0.64	−0.64	−0.56	0.26	−0.14	−1.60
C_1 1.53	1.24	1.64	1.01	1.01	1.65	1.85	2.72	4.40
C_2 −0.16	−0.05	−0.16	−0.06	−0.06	−0.31	−0.50	−0.75	−1.08

Table 11. Fit SL – leading order

$\Lambda(1,4) = 0.144\,\text{GeV}$					$Q_0^2 = 4\,\text{GeV}^2$				
	d(val)	u(val)	gluon	u(sea)	d(sea)	s	c	b	t
					A_0				
C_0	1.38	1.67	1.52	−0.81	−0.81	−0.81	−3.62	−6.16	−12.68
C_1	−0.62	−0.33	−3.17	−1.13	−1.13	−1.13	0.03	2.37	8.36
C_2	−0.10	−0.13	0.25	−0.26	−0.26	−0.26	−0.48	−1.24	−2.89
					A_1				
C_0	0.16	0.08	−0.25	−0.07	−0.07	−0.07	−0.06	−0.11	−0.35
C_1	0.19	0.17	−0.01	−0.46	−0.46	−0.46	−0.21	−0.05	0.28
C_2	−0.02	−0.01	0.00	0.16	0.16	0.16	0.05	−0.02	−0.12
					A_2				
C_0	5.40	3.75	7.01	9.19	9.19	9.19	8.30	6.49	14.87
C_1	0.59	0.70	−0.90	0.35	0.35	0.35	−0.60	1.28	−12.56
C_2	0.03	0.00	−0.08	−0.49	−0.49	−0.49	0.25	−0.41	4.75
					A_3				
C_0	−1.97	−2.09	−0.79	−0.89	−0.89	−0.89	0.16	0.71	−0.17
C_1	1.24	0.98	2.90	0.33	0.33	0.33	1.26	1.37	3.39
C_2	−0.05	0.02	−0.54	0.40	0.40	0.40	−0.15	−0.26	−0.96

References

1. M. Glueck et al.: Z. Phys. C – Particles and Fields 13 (1982) 119; D. Duke, J. Owens: Phys. Rev. D30 (1984) 49; E. Eichten et al.: Rev. Mod. Phys. 56 (1984) 579 and Erratum 58 (1986) 1065
2. A.D. Martin, R.G. Roberts, W.J. Stirling: Phys. Rev. D37 (1988) 1161; Mod. Phys. Lett. A4 (1989) 1135; P.N. Harriman, A.D. Martin, R.G. Roberts, W.J. Stirling; Phys. Rev. D42 (1990) 798
3. M. Diemoz et al.: Z. Phys. C – Particles and Fields (1988) 21
4. Wu-Ki. Tung et al.: Proceedings of the 1988 summer study on high energy physics in the 1990's, World Scientific 1990
5. J.P. Berge et al.: Preprint CERN-EP/89-103 (1989)
6. J.J. Aubert et al.: Nucl. Phys. B293 (1987) 740
7. A.C. Benvenuti et al.: Phys. Lett. B223 (1989) 485, CERN-EP/89-170, 171 (1989)
8. G. Curci, W. Furmanski, R. Petronzio: Nucl. Phys. B175 (1980) 27; W. Furmanski, R. Petronzio: Phys. Lett. 97B (1980) 437; E.G. Floratos et al.: Nucl. Phys. B192 (1981) 417 and references cited therein
9. Cf. E.G. Floratos et al.: Nucl. Phys. B192 (1981) 417 and references cited therein
10. Wu-Ki Tung: Nucl. Phys. B315 (1989) 378
11. K. Lang et al.: Z. Phys. C – Particles and Fields 33 (1987) 483; S.R. Mishra et al.: in: Proceeding of 14th rencontres de Moriond, Mar. 1889
12. J. Kubar, M. Le Bellac, J.L. Meunier, G. Plaut: Nucl. Phys. B175 (1980) 251
13. A.S. Ito et al.: Phys. Rev. D23 (1981) 604
14. C.N. Brown et al.: Phys. Rev. Lett. 63 (1988) 371
15. P. Aurenche et al.: Phys. Rev. D39 (1989) 3275
16. J. Feltesse, in: Proceedings of the XIV international symposium on lepton and photon interactions, Stanford, August 1989, Singapore: World Scientific
17. Abramowicz et al.: Z. Phys. C – Particles and Fields 17 (1984) 283; Z. Phys. C – Particles and Fields 25 (1984) 29; Z. Phys. C – Particles and Fields 35 (1984) 443
18. Wu-Ki Tung: Overview of parton distribution and the QCD framework, Fermilab preprint (to appear in: Proceedings of workshop on hadron structure functions and parton distributions, Fermilab, April 1990, Singapore: World Scientific.
19. See, for instance, A. Milsztajn, contribution to Proceedings of workshop on hadron structure functions and parton distributions, Fermilab, April, 1990, World Scientific Pub. (to be published)
20. See: Proceedings of workshop on parton distribution functions at small-x, DESY, May 1990, Amsterdam: North Holland (to be published)
21. Wu-Ki Tung, preprint Fermilab-Conf-90/200-T (to appear in: Proceedings of workshop on parton distribution functions at small-x, DESY, May 1990, Amsterdam: North Holland
22. F. Olness, Wu-Ki Tung: Small x physics at the SSC and the Tevatron in: From colliders to super colliders, V. Varger and F. Halzen (eds) Singapore: World Scientific 1987
23. W. Furmanski, R. Petronzio: Z. Phys. C – Particles and Fields 11, (1982) 293
24. G. Altarelli, R.K. Ellis, G. Martinelli: Nucl. Phys. B143 (1978) 521; and ibid. B157 (1979) 461