CTEQ6 parton distributions with heavy quark mass effects

Stefan Kretzer,[1,2,3] H. L. Lai,[4] Fredrick I. Olness,[5] and W. K. Tung[1]
[1]*Department of Physics and Astronomy, Michigan State University, East Lansing, Michigan 48824, USA*
[2]*Physics Department, Brookhaven National Laboratory, Upton, New York 11973, USA*
[3]*RIKEN-BNL Research Center, Building 510a, Brookhaven National Laboratory, Upton, New York 11973-5000, USA*
[4]*Taipei Municipal Teacher's College, Taipei, Taiwan*
[5]*Department of Physics, Southern Methodist University, Dallas, Texas 75275, USA*
(Received 15 July 2003; published 2 June 2004)

Previously published CTEQ6 parton distributions adopt the conventional zero-mass parton scheme since the corresponding hard cross sections are universally available. For precision observables which are sensitive to charm and bottom quark mass effects, we provide in this paper an improved CTEQ6HQ parton distribution set determined in the more general variable flavor number scheme that incorporates heavy flavor mass effects. We describe in detail the QCD scheme and analysis procedure used, examine the predominant features of the new distributions, and compare them with previous distributions.

DOI: 10.1103/PhysRevD.69.114005 PACS number(s): 13.60.Hb

I. INTRODUCTION

Parton distributions provide the essential link between the theoretically calculated partonic cross sections in the standard-model and new-physics worlds, and the experimentally measured cross sections in the hadron world through the factorization property of perturbative QCD. Since perturbative calculations are renormalization scheme dependent,[1] it is important to use matching hard cross sections and parton distribution functions (PDFs) in evaluating factorized cross sections: Evolution kernels and hard cross sections have to match each other strictly in perturbative order and renormalization scheme[2] since otherwise renormalization group invariance is manifestly violated. On top of this, there are more practical matching conditions [1–3] in the generalized parton model with nonzero parton masses. Different prescriptions (e.g., [1,4,5]) for calculating mass-dependent hard cross sections may all be defined consistently with, say, modified minimal subtraction scheme ($\overline{\mathrm{MS}}$) evolution equations. In practice, however, PDFs are determined from a global fit to data so that different prescriptions for calculating hard cross sections will result in different fitting parameters and, accordingly, different values of the PDFs. Parton distributions should, therefore, not be applied in combination with hard cross sections different from those employed in their extraction from data. Calculations that are unmatched in this practical sense are bound to fail to reproduce experimental data, even though they may be theoretically consistent in the above strict sense.[3]

For most physical applications, it is convenient (and a good approximation) to use hard cross sections calculated in the zero-parton-mass limit. In fact, next-to-leading-order (NLO) hard cross sections in the nonzero-quark-mass case are not even available beyond deep inelastic scattering (DIS). Thus the most useful parton distributions for general applications are those determined in global QCD analysis using the zero-quark-mass approximation. This was the choice made for the latest series of CTEQ parton distributions CTEQ6M, 6D, and 6L [6].

This paper extends Ref. [6] by performing a similar global analysis using the generalized (nonzero-quark-mass) perturbative QCD framework of [1,3,7–9]. When matched to the corresponding hard cross sections calculated in the same scheme, these should provide a more accurate description of the precision DIS structure function data, as well as other processes that are sensitive to charm and bottom mass effects. The main result is a new standard set of parton distributions named CTEQ6HQ.[4] At high energies, predictions based on these PDFs are indistinguishable from those based on the zero-mass ones, within existing experimental and theoretical uncertainties.

Perturbative QCD with nonzero quark masses is by now well established [1,3,5,7,12–14]. Its implementation is, however, necessarily more complicated than the corresponding zero-mass case. In particular, in addition to the familiar renormalization scheme dependence, more ambiguities arise from the implementation of the charm and bottom thresholds (in relation to their masses), and the detailed way the three-flavor, four-flavor, and five-flavor renormalization schemes are matched at the transition scales. Several schemes proposed in the recent literature are formally equivalent to each other at high energies. But they can differ near the thresholds as well as in the intermediate energy range, where most ex-

[1]We shall use the term *renormalization scheme* in the general sense, to include both the usual renormalization (ultraviolet subtraction) scheme and the factorization (infrared and collinear subtraction) scheme.

[2]For example, "$\mathrm{PDF}^{\overline{\mathrm{MS}}}_{\mathrm{NLO}} \otimes \hat{\sigma}^{\overline{\mathrm{MS}}}_{\mathrm{NLO}}$," "$\mathrm{PDF}^{\mathrm{DIS}}_{\mathrm{NNLO}} \otimes \widehat{\sigma}^{\mathrm{DIS}}_{\mathrm{NNLO}}$," etc.

[3]As we will discuss below, the prescriptions differ by power suppressed terms. The ambiguity is, therefore, no different in principle from the consistent inclusion (or omission) of higher twist terms in PDF analyses and their application—although quark mass effects may be more important quantitatively.

[4]A fit named CTEQ6F3 employing the three fixed flavor-number scheme (3-FFNS) for DIS (cf. Refs. [10,11]) which is suitable for special applications in that scheme, will be presented separately.

perimental data lie [1,5,9,14–16]. Some of these are more natural and numerically robust, e.g., insensitive to variations of the remaining parameters (such as the factorization scale), than others. We follow the procedure proposed in [16,17] which has been shown to be particularly stable, as it naturally takes into account the dominant (logarithmic *and* nonlogarithmic) mass-dependent kinematic effects with a physically motivated choice of the scaling variable.

Section II summarizes the method of analysis. Because this work represents an extension of the zero-quark-mass CTEQ6 global QCD analysis, we shall focus mostly on the new elements related to nonzero quark mass; details that are in common with the massless case can be found in Ref. [6]. Section III describes the global analysis of data. Section IV compares the new parton distributions with data. Section V makes comparisons between the CTEQ5M, CTEQ6M, and CTEQ6HQ PDFs. Quark mass effects become negligible at high energy scales. One expects the CTEQ6M and CTEQ6HQ PDFs to approach one another in the high Q limit. This comparison illustrates how the differences in initial PDFs fitted at low Q diminish with increasing Q. This section also contains a discussion of applications and issues related to uncertainties of the physics predictions resulting from these parton distributions.

II. THE GENERALIZED $\overline{\text{MS}}$ SCHEME WITH NONZERO MASS HEAVY FLAVOR PARTONS

The main feature of this analysis, compared to the standard zero-mass parton CTEQ6 PDF sets, is the adoption of the generalized perturbative QCD framework incorporating nonzero-mass effects for the charm and bottom quarks.[5] Collins has shown that, to all orders of the perturbation expansion, factorization of the DIS structure functions holds for massive partons to the same degree of rigor as in the familiar zero-mass case [3]. The relatively simple physical ideas underlying the full factorization proof are described in Refs. [9,16,17]. To implement this formalism, which we shall refer to as the generalized $\overline{\text{MS}}$ scheme [or general mass variable flavor number scheme (GM-VFNS), as mentioned earlier], we follow the explicit procedures outlined in the Appendix of Ref. [9], and supplement this with the threshold prescription of Refs. [16,17]. The last reference also contains a review of alternative approaches.

The NLO DIS structure functions are given by the generic formula (suppressing the structure function labels 1,2,3)

$$F(x,Q)=\sum_a \int \frac{dz}{z} f_a(z,\mu)\ \hat{\omega}^a\left(\frac{x}{z},\frac{Q}{\mu},\frac{m_H}{\mu},\alpha_s(\mu)\right) + \mathcal{O}\left(\alpha_s^2,\frac{\Lambda^2}{Q^2},\frac{\Lambda^2}{m_H^2}\right), \quad (1)$$

where a is the initial state parton label, f_a is the parton distribution function, $\hat{\omega}^a$ is the hard-scattering cross section calculated in perturbative QCD (PQCD) to order α_s^1, Λ is the QCD lambda parameter, and m_H (generically) represents heavy quark masses, if present. The prescription dependence allowed by the PQCD formalism is associated with possible implementations of the first term on the right-hand side of Eq. (1), within the accuracy specified by the remainder term (which will be dropped from now on).[6] In the sum over partons in Eq. (1), contributions due to the light quark (u,d,s) are standard; contributions arising from massive charm and bottom quarks, along with that of the gluon, are treated differently, as explained below.

A. Leading contributions and the ACOT(χ) implementation

To be specific, we will consider DIS neutral-current charm production process. The leading-order process is a virtual photon scattering off a charm parton: $\gamma^* c \to c$. The NLO QCD corrections for this process are the boson-gluon fusion process ($\gamma^* g \to c\bar{c}$) and the gluon radiation process ($\gamma^* c \to cg$) as well as the charm-initiated one-loop virtual process ($\gamma^* c \to c$).

First, let us examine the kinematic region near the charm production threshold: $W=\sqrt{Q^2(1/x-1)} \simeq 2m_c$. To make the notation precise, we will use the renormalization scale μ to generically label our characteristic energy scale; separately, we use μ_m to label the *matching point* and μ_t to label the *transition point* (cf. [9] and below). To implement the generalized $\overline{\text{MS}}$ scheme, the first step is to consider a three-flavor scheme (appropriate for energy scales $\mu \lesssim m_c$) and a four-flavor scheme (for scales $\mu \gtrsim m_c$), and choose a matching point μ_m (on the order of m_c) where the two schemes are matched [i.e., where the discontinuities of $\alpha_s(\mu)$ and $f_a(x,\mu)$ are calculated]. We choose the conventional value for the matching point, $\mu_m = m_c$, which yields the simplest matching conditions [7].

Having precisely defined the three-flavor and four-flavor schemes, which coexist in the region of the charm threshold, we can still choose a *transition point* μ_t, where one makes the transition from the three-flavor to the four-flavor scheme in the calculation of physical quantities.[7] The choice of the transition point is arbitrary—its choice is part of the definition of the composite renormalization scheme. While there is much flexibility in the choice of μ_t, clearly it must lie within the overlapping region of applicability of the three-flavor and four-flavor schemes—therefore, we typically choose $\mu_t \sim m_c$. Although a case can be made, in principle, for choosing μ_t far above m_c (since the four-flavor scheme is not a

[5]The top quark can, for all practical purposes, be treated as a heavy particle, not a parton.

[6]Note that according to the factorization proof of Ref. [3], the remainder term contains corrections of order Λ^2/Q^2 and Λ^2/m_H^2, but *no* corrections of order m_H^2/Q^2.

[7]The distinction between the matching and the transition points is not addressed in most papers. This distinction is made in Ref. [3]; it is discussed at length in the Appendix of Ref. [9]. As mentioned above, arguably, there is a good case for choosing the transition point μ_t to be higher than the simple matching point $\mu_m = m_c$.

natural scheme just above m_c), or for considering an x-dependent transition scale $\mu_t(x,m_c)$ (because the threshold $W>2m_c$ depends on x as well as Q), we make the plain choice $\mu_t=m_c$ to simplify the calculation. This choice is reasonable *only* within a prescription for handling x-dependent threshold effects of Eq. (1) which naturally suppresses the four-flavor contribution to the physical structure function when $\mu\sim m_c$ and/or $W\gtrsim 2m_c$. The specific renormalization prescription of Ref. [16] has precisely this feature, as we shall outline below.

Detailed descriptions of the various NLO contributions to the the generalized $\overline{\rm MS}$ scheme were discussed in Ref. [9]. While we refrain from reproducing them here, we do need to specify the precise way the threshold effects are implemented when heavy quarks are involved. Consider first the numerically significant NLO *boson-gluon fusion* contribution $(\gamma^* g\to c\bar{c})$:

$$F^{(1)}(\gamma^* g\to c\bar{c}) = \int \frac{dz}{z} g(z,\mu)\ \hat{\omega}_g\left(\frac{x}{z},\frac{Q}{\mu},\frac{m_c}{\mu}\right)$$
$$= \alpha_s(\mu)\left[\int_\chi^1 \frac{dz}{z} g(z,\mu)\omega_g^1\left(\frac{x}{z},\frac{m_c}{Q}\right) - \ln\left(\frac{\mu}{m_c}\right)\int_\zeta^1 \frac{dz}{z} g(z,\mu) P_{g\to c}\left(\frac{\zeta}{z}\right)\omega_c^0\left(\frac{m_c}{Q}\right)\right], \quad (2)$$

where the first term on the right-hand side corresponds to the unsubtracted boson-gluon fusion diagram contribution (with the α_s factor explicitly taken out),[8] and the second term represents the subtraction term which renders $\hat{\omega}_g$ infrared safe. To simplify the discussion, we restrict ω_c^0 to be a constant, even though it could be a more complicated functional form such as discussed in [4]. The variable χ is dictated by the kinematics of the boson-gluon fusion partonic subprocess $(\gamma^* g\to c\bar{c})$ to be a generalized scaling (or ''rescaling'') variable

$$\chi = x\left(1+\frac{4m_c^2}{Q^2}\right), \quad (3)$$

where x is the conventional Bjorken variable. In contrast, the other scaling variable ζ in Eq. (2) can, in principle, be chosen arbitrarily provided the following two constraints are imposed: (1) $\zeta\to x$ in the high energy limit; and (2) The same ζ variable is used here as in the order α_s^0 (''leading-order'') process $\gamma^* c\to c$.

In the high energy limit $(m_c^2/Q^2\to 0)$, the first constraint ensures that both $\zeta\to x$ and $\chi\to x$ such that the complete boson-gluon fusion contribution is infrared safe. The leading-order process $\gamma^* c\to c$ is, in our approach,

$$F^{(0)}(\gamma^* c\to c) = \int \frac{dz}{z} c(z,\mu)\hat{\omega}_c^0\left(\frac{Q}{\mu},\frac{m_c}{\mu}\right)\ \ \delta\left(1-\frac{\zeta}{z}\right)$$
$$= c(\zeta,\mu)\ \ \hat{\omega}_c^0\left(\frac{Q}{\mu},\frac{m_c}{\mu}\right). \quad (4)$$

For the second constraint, if the same ζ variable is used for both the boson-gluon fusion process of Eq. (2) $(\gamma^* g\to c\bar{c})$ and the leading-order contribution of Eq. (4) $(\gamma^* c\to c)$, then the gluon subtraction term [last term in Eq. (2)] cancels the leading-order contribution in the limit $\mu\to m_c$; this ensures that the four-flavor formula will reduce to the three-flavor formula, and the dominant contribution in this region will be the unsubtracted gluon fusion term, as required by the correct physics to this order of PQCD.

A natural choice of ζ that ensures the best behavior at *both* the high and the low energy limits is proposed by Ref. [16]:

$$\zeta=\chi\equiv x(1+4m_c^2/Q^2). \quad (5)$$

This choice has the property that, as the heavy quark production threshold is approached from above $(W\to 2m_c)$, then $\zeta\to 1$ such that the integration measure in Eq. (2) vanishes.[9] Therefore, with the choice $\zeta=\chi$, the subtraction term in Eq. (2) and the LO quark term Eq. (4), individually vanish at threshold by kinematics. In addition, the cancellation between these two terms due to the dynamics of PQCD, as explained in Ref. [1], still operates.

This results in unequaled stability of the predictions in the threshold region, compared to previous implementations of the GM-VNFS in the threshold region [1,4,5,15]. These usually give rise to rather prescription-sensitive behavior just above threshold. The robust implementation we adopt will be referred to as the ''ACOT(χ)'' scheme, following Refs. [16,17].[10]

B. Other contributions to the full NLO calculation

For the complete order α_s calculation, we must also include the real and virtual NLO quark-initiated contributions to Eq. (1),

$$F^{(1)}(\gamma^* c\to c+X) = \int \frac{dz}{z} c(z,\mu)\hat{\omega}_c^1\left(\frac{x}{z},\frac{Q}{\mu},\frac{m_c}{\mu}\right). \quad (6)$$

The hard-scattering cross section $\hat{\omega}_c^1(x,Q/\mu,m_c/\mu)$, has been computed for general masses [18] in the ACOT scheme [1]. For scales around the canonical DIS choice of $\mu\simeq Q$,

[8] We use the notation that, for each flavor, ω_a denotes the parton-level cross section *before* infrared and collinear subtraction, while $\hat{\omega}_a$ denotes the corresponding infrared safe cross section *after* subtraction.

[9] This is more easily seen from the equivalent formula $\chi=1-(1-x)(1-4m_c^2/W^2)$.

[10] So named since it is a variant of the original Aizavis-Collins-Olness-Tung (ACOT) approach [1], replacing the naive x with the kinematically natural rescaling variable χ of Eq. (3).

this term is numerically insignificant compared to the terms considered above. We include it in our calculation to ensure (perturbative) consistency.

It is also worth mentioning that the calculation of this term, which is computation intensive in spite of the smallness of its numerical contribution, can be considerably streamlined by using a simplifying feature of PQCD in the generalized $\overline{\text{MS}}$ scheme: one can set $m_c=0$ in $\hat{\omega}_c^1$ and ω_c^0 in Eq. (2) and Eq. (4) without sacrificing accuracy. This point is discussed in detail in Ref. [2] where the numerical insensitivity of the calculation to the choice of m_c is explicitly demonstrated.[11]

C. Remaining scheme and scale choices

The only remaining arbitrariness of our prescription is now associated with the familiar choice of the factorization (and renormalization) scale μ. The scale dependence of perturbative calculations in general arises from imperfect matching between the terms in the truncated perturbation series. Because of the near-perfect matching in the ACOT(χ) prescription discussed above (in contrast to alternate choices), the dependence of the heavy quark contributions to the physical structure functions on the choice of μ is very small (cf. Ref. [16]) throughout the energy range of our global analysis. For definiteness, we use $\mu^2=Q^2+m_c^2$. In summary, the quark-initiated terms in the simplified ACOT(χ) scheme adopted in this work are

$$\int_\chi^1 \frac{dz}{z} c(z,\mu)\hat{\omega}_c^{0,1}\left(\frac{\chi}{z},\frac{Q}{\mu},\frac{m_c}{\mu}=0\right)\Bigg|_{\mu^2=Q^2+m_c^2}$$
$$=\int_\chi^1 \frac{dz}{z} c(z,\mu)C^{0,1}{}_{\overline{\text{MS}}}\left(\frac{\chi}{z},\frac{Q}{\mu}\right)\Bigg|_{\mu^2=Q^2+m_c^2} \quad (7)$$

with $C^{0,1}{}_{\overline{\text{MS}}}$ representing the standard coefficient functions for massless quarks in LO and NLO, defined in the conventional $\overline{\text{MS}}$ scheme.

The same treatment is applied to the b quark threshold region. Here, the four-flavor scheme is matched onto the five-flavor scheme for scales larger than m_b, and the production of open bottom particles in the final state is added to the inclusive DIS structure functions. In principle, a similar treatment needs to be applied for other processes included in the global analysis (currently, Drell-Yan and jet production) using appropriate hard-scattering cross sections calculated in the massive scheme. In practice, this is not done for two reasons: (i) NLO calculations of the hard-scattering cross section involving massive quarks do not yet exist for these processes; and (ii) unlike the case of DIS, the experimental errors in these processes are relatively large compared to the anticipated differences between the massless and massive scheme calculations.

[11]The simplified scheme [2] can also be recovered by investigation of the resummed perturbation series [17].

III. GLOBAL FITTING AND THE CTEQ6HQ PARTON DISTRIBUTIONS

The new global fitting performed in the generalized $\overline{\text{MS}}$ formalism follows the same procedure as that of the earlier CTEQ6 analysis [6] using the zero-mass-parton formalism. The data sets used before are supplemented by (i) the H1 $F_2^{e^+p}$ set [19] (which was inadvertently left out in Ref. [6]); and (ii) the H1 [20] and ZEUS [21] data sets for the structure function F_2^c with tagged charm particles in the final state. The additional H1 $F_2^{e^+p}$ data set does not have much influence on the new analysis, since one already obtains an excellent fit to these data just by comparing them with predictions of CTEQ6M. The F_2^c data sets are quite relevant for this analysis since F_2^c is sensitive to the charm and gluon distributions, which are tightly coupled in the generalized $\overline{\text{MS}}$ formalism.[12]

The parametrization of the nonperturbative parton distribution functions at $Q_0=m_c=1.3$ GeV is the same as in Ref. [6]. For this study, we assume as usual that charm partons are entirely "radiatively generated" (i.e., through QCD evolution) from the starting scale Q_0 onward. This assumption is somewhat arbitrary, and it is obviously dependent on the choice of Q_0. The possibility of having a small component of nonperturbative charm [22] at low Q and its physical consequence will be examined separately.

As in the previous CTEQ6 analysis, correlated experimental systematic errors are fully incorporated whenever available. The best fit obtained with these inputs will be called CTEQ6HQ, or C6HQ for short. A broad measure of the quality of this fit is provided by the overall χ^2 of 2033 for a total number of 1950 data points (χ^2 per degree of freedom (DOF)$=1.04$). This is to be compared to a χ^2 of 1946 for 1811 points (χ^2 per DOF$=1.07$) in the case of CTEQ6M (abbreviated to C6M in the following) [6]. To gain a better feel of how these fits compare, we show in Table I (the third and fourth columns) the overall χ^2 values and, in parentheses, χ^2 per data point, as well as for the applicable individual data sets. The total number of data points in this head-to-head comparison (not including the charm production data points) is 1925. The new C6HQ fit reduces the overall χ^2 by 29 out of $\sim$2000 as compared to the C6M fit. The improvement of this generalized $\overline{\text{MS}}$ result over the zero-mass $\overline{\text{MS}}$ result is encouraging, since the generalized $\overline{\text{MS}}$ formalism represents a more accurate formulation of PQCD. However, a difference of χ^2 of 29 is within the current estimated range of uncertainty of PDF analysis [6,23,24]. Therefore, the significance of this difference is arguable. We also note that the improvement in χ^2 is spread over most of the data sets: there is no smoking gun for the overall difference.

The last two columns of Table I compare the above results with two possible uses of the PDFs that represent a *misuse* of

[12]These data sets were not used in the CTEQ6M analysis because F_2^c is not well defined theoretically in the zero-mass-parton formalism.

TABLE I. Comparison of the χ^2 values of the general-mass CTEQ6HQ fit (third column) with the zero-mass CTEQ6M fit (fourth column). Also included are comparisons with two ''mismatched'' cases when GM and ZM parton distribution functions are convoluted with the other (wrong) hard cross section (fifth and sixth columns). The first number of each entry is the χ^2 value; the number in parentheses is χ^2 per number of points. Correlated systematic errors, if available, are included.

Data set	No. of pts	CTEQ6HQ		CTEQ6M		C6M⊗GM		C6HQ⊗ZM	
Bcdms_p	339	370	(1.09)	370	(1.09)	370	(1.11)	373	(1.10)
Bcdms_d	251	269	(1.07)	279	(1.11)	274	(1.07)	281	(1.12)
H1a	104	94	(0.91)	102	(0.98)	258	(2.84)	387	(3.72)
H1b	126	124	(0.99)	130	(1.03)	135	(1.11)	123	(0.98)
H1c	129	103	(0.80)	111	(0.86)	119	(0.84)	104	(0.80)
Zeus	229	266	(1.16)	261	(1.14)	474	(2.11)	364	(1.59)
Nmc_p	201	304	(1.51)	299	(1.49)	273	(1.35)	366	(1.82)
Nmc_d/p	123	112	(0.91)	111	(0.91)	111	(0.90)	114	(0.92)
Ccfr_F2	69	90	(1.30)	120	(1.74)	116	(1.82)	107	(1.55)
Ccfr_F3	86	35	(0.41)	37	(0.43)	36	(0.40)	36	(0.42)
E605	119	102	(0.86)	103	(0.86)	101	(0.86)	102	(0.86)
Cdf_wasy	11	9	(0.78)	9	(0.83)	9	(0.83)	9	(0.78)
E866	15	5	(0.34)	6	(0.43)	6	(0.43)	5	(0.34)
D0_jet	90	71	(0.79)	49	(0.55)	49	(0.55)	71	(0.79)
Cdf_jet	33	55	(1.66)	50	(1.51)	50	(1.51)	55	(1.66)
All	1925	2008	(1.04)	2037	(1.06)	2431	(1.26)	2496	(1.30)

PQCD in principle, but occur frequently in the literature in practice, perhaps out of necessity. These involve using PDFs obtained in the general-mass (GM) scheme convoluted with hard-scattering cross sections (Wilson coefficients) defined in the zero-mass (ZM) scheme, and vice versa.[13] For the same data sets, these mismatched schemes result in a difference of 420–490 in the overall χ^2. These are quite large differences relative to the tolerances discussed in Refs. [6,25] and result in clear discrepancies with some of the precision DIS data sets, as can be seen in Table I. The lesson is clear: *for quantitative applications, it is imperative to maintain consistency between the PDFs and the hard-scattering cross sections*.

IV. COMPARISON WITH DATA

In Ref. [6], we presented an extensive comparison between the CTEQ6 results and the data sets used in the global analysis, including new ways to explicitly account for the correlated experimental systematic errors. Since the new C6HQ fit is generally similar to the C6M fit, we shall not duplicate the same comparisons for those quantities where the differences are minimal. In Sec. IV A, we compare with the DIS charm production data, which were not used in the previous CTEQ6 analysis. In Sec. IV B, we discuss some implications of the neutrino data.

[13]For instance, the Martin-Roberts-Stirling-Thorne distributions [25] are obtained in the general-mass formalism (using the Robert-Thorne implementation [4]); they are often used in applications convolving with readily available zero-mass hard-scattering cross sections.

A. DIS charm production data from HERA

Because the H1 and ZEUS F_2^c structure functions are sensitive to the charm and gluon distributions, these will play an important role in extracting the C6HQ PDFs. In Fig. 1, we display comparisons of the charm production data from H1 [20] and ZEUS [26] with the C6HQ theory value. For simplicity, we combine all the data points from each experiment in a single plot, and we scale the plot by the theoretical calculation. On the x axis, the data points are ordered using x and Q as the primary and secondary sorting variables, respectively. The error bars represent the statistical and uncorrelated systematic errors added in quadrature. We see that the fits are good with χ^2 per data point of 0.881 (7/8) for H1 and 0.98 (18/18) for ZEUS, respectively.

B. Neutrino data in global parton analyses

The NuTeV measurement of the weak mixing angle $\sin^2\theta_W$ (Ref. [27]) has recently focused considerable attention on the neutrino-induced DIS process [28]. In this measurement, the weak mixing angle $\sin^2\theta_W$ is extracted from the ratio of neutral-current (NC) to charged-current (CC), ν and $\overline{\nu}$, cross sections. Even before this $\sin^2\theta_W$ measurement, there have been other long-standing unresolved issues in comparing CC structure functions F_2^{CC} (measured in ν and $\overline{\nu}$ scattering) with NC structure functions F_2^{NC} (measured in $e^{\pm}$ and $\mu^{\pm}$ scattering) at modestly low $x(\sim 10^{-2})$. One manifestation of this is that the recently measured structure function $\Delta x F_3^{\nu/\overline{\nu}N}$ is not compatible with the QCD predictions [29,30]. As this structure function is particularly sensitive to the heavy quark components $\Delta x F_3^{\nu/\overline{\nu}N} \simeq s(x) - c(x)$, it is certainly important to properly treat the heavy quark mass.

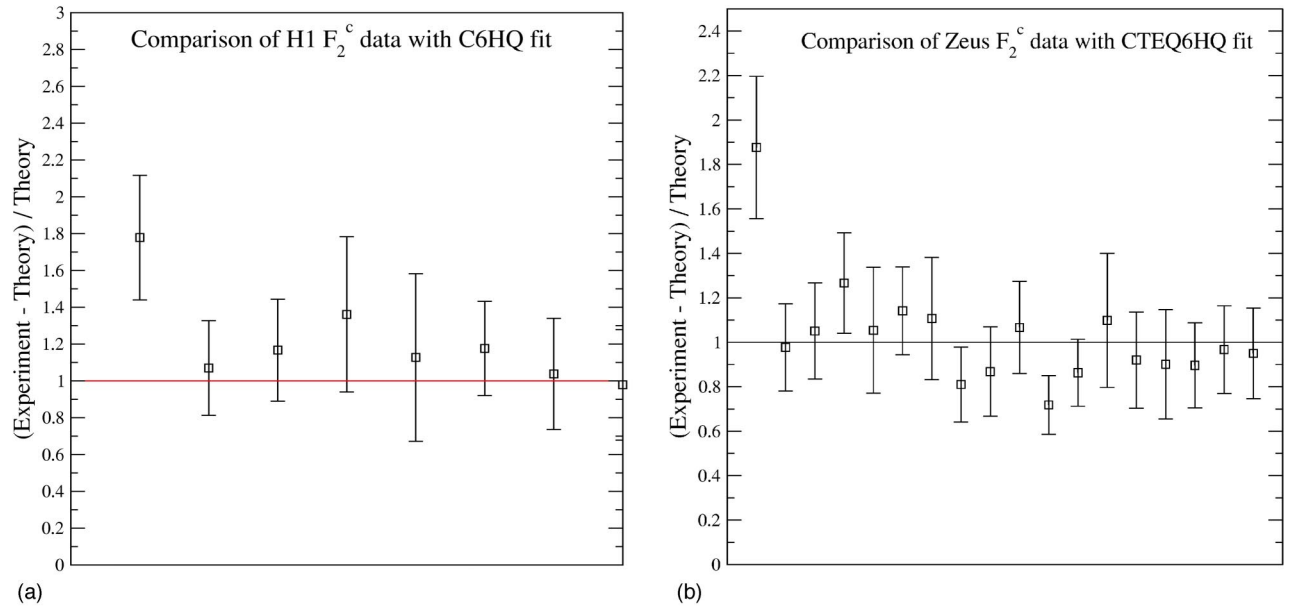


FIG. 1. Comparisons of the charm production structure function F_2^c with the data from H1 [20] (a) and ZEUS [26] (b). The y axis is (experiment−theory)/theory; along the x axis, the data points are ordered using x and Q as the primary and secondary sorting variables, respectively. The error bars represent the statistical and uncorrelated systematic errors added in quadrature.

While the deviations of $\Delta x F_3^{\nu/\bar{\nu}N}$ are at a nondramatic $\sim 1\sigma$ level, the pattern has been systematic and persistent.

Therefore, we reexamine the influence of the neutrino structure functions in the current global analysis keeping in mind the following issues. First, the high statistics neutrino structure function data of the CCFR Collaboration used here are from the newer ''physics-model-independent'' data analysis [31], rather than the earlier ones which contained model-dependent corrections [32]. Secondly, meaningful (i.e., quantitative) comparison of NC and CC structure functions (which represent different hard-scattering processes) can be made only within a common underlying theoretical framework with an accuracy comparable to the experimental uncertainties.[14] The proper treatment of charm mass effects, such as those described in the current study, is an important part of that theoretical formalism; cf. [33].

First, we ask whether the neutrino structure function $F_2^{\nu/\bar{\nu}N}$ obtained by the newer model-independent analysis is consistent with the muon-induced NC fixed-target structure functions $F_2^{\mu^{\pm}N}$ and the HERA collider structure functions $F_2^{e^{\pm}N}$ in a consistent NLO analysis.[15] Figure 2 shows the comparison between the data with the CTEQ6HQ fit (using $m_c = 1.3$ GeV). We see that, even if this data set is included in the global fit (cf. Table I), the measured values in the low x bins (left figure) consistently lie above the theory curves. The systematic deviation of data points from the global fit is reflected in the relatively large χ^2's associated with this data set, as seen in Table I. The fit is dominated by the much more extensive NC data (from the BCDMS, NMC, H1, and ZEUS experiments); hence one can regard the theory curves as representing the NC data properly ''corrected'' for NLO QCD effects. The pattern of deviation of the neutrino data from the average NC predictions at low x is similar to earlier comparisons. In other words, the discrepancy between the CCFR F_2^{ν} and the other NC $F_2^{e,\mu}$ measurements at small x persists in the context of contemporary global QCD analyses, even when charm mass effects are properly treated as in the present analysis. Other possible theoretical sources for this difference (in addition to the NLO contributions included in this comparison) have been discussed in Refs. [28,29,34],

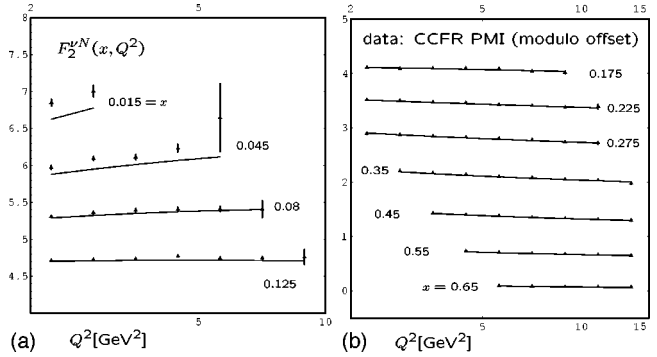


FIG. 2. CCFR $F_2^{\nu N}$ structure function (from the *physics-model-independent* analysis, Ref. [31]) compared to CTEQ6HQ fit: (a) low x bins (left panel) show a systematic disagreement; (b) medium to high x bins (right panel) show good agreement.

[14]For example, direct comparisons of the ratio F_2^{ν}/F_2^{μ} to theory (suggested by the ''5/18 rule'') are meaningful only at the LO parton model level, which is not appropriate for the precision data from current experiments.

[15]Reference [31] suggests these data are consistent by examining the ratio of the structure functions. However, as explained in the text, it is not appropriate to compare the ratio of two different structure functions with theory in QCD beyond LO; a complete analysis involving a NLO fit is required.

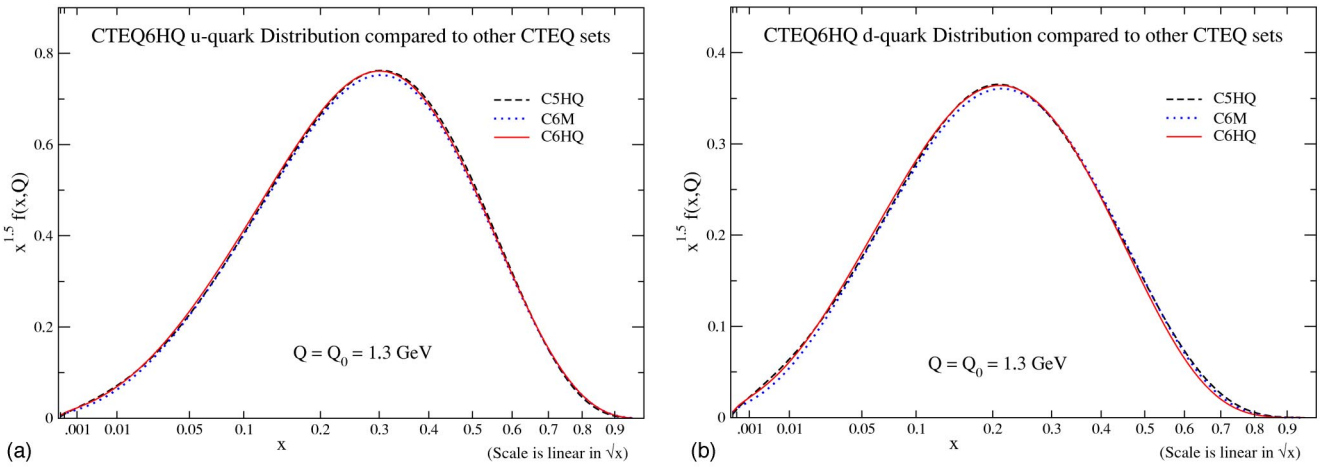


FIG. 3. Comparison of CTEQ5HQ, CTEQ6M, and CTEQ6HQ parton distributions at $Q=m_c=1.3$ GeV: (a) u quark and (b) d quark. The axes are scaled to highlight the valence components of these distributions.

and in the literature quoted therein. On the experimental side, preliminary NuTeV measurements [35] seem to yield improved agreements with the NLO QCD predictions at low x. It remains to be seen whether the final NuTeV results will resolve this problem.

In our analysis, the strangeness distribution is constrained to be proportional to $[\overline{u}(x)+\overline{d}(x)]$, and its normalization at the scale Q_0 is constrained to be

$$\kappa=\frac{\int dx x[s(x)+\overline{s}(x)]}{\int dx x[\overline{u}(x)+\overline{d}(x)]}\simeq\frac{1}{2}, \tag{8}$$

as in most current global analysis. In the future, a NLO analysis of the CC charm production data ($\nu s\to c\mu\to\mu^{\mp}\mu^{\mp}X$) from CCFR and NuTeV [36] has the potential to determine $\{s(x),\overline{s}(x)\}$ with much more precision. Then, a reassessment of the apparent discrepancy between the measured $\Delta xF_3^{\nu/\overline{\nu}N}$ and the low lying QCD predictions [29] will be warranted. An accurate determination of the strangeness of the proton, as well as the strangeness asymmetry $[s(x)-\overline{s}(x)]$, will have important implications for these measurements, including the NuTeV anomaly; cf., Refs. [27,28].

Clearly, the interpretation of the neutrino DIS data requires a detailed NLO analysis. Global QCD analyses of parton distributions, such as the one performed in this paper, are essential to unraveling the underlying physics. Progress on both the experimental and the theoretical fronts is needed; this advancement is ongoing.

V. COMPARISON WITH RELATED PDFS: CTEQ6HQ, CTEQ6M, CTEQ5HQ

The CTEQ6HQ and CTEQ6M fits provide comparable descriptions of the global QCD data in two *different* implementations of the VNFS. Since both are fitted to the same data, the differences in hard cross sections result in somewhat different PDFs. We are particularly interested in the differences for the charm distribution, and the closely correlated gluon distribution, due to the improved treatment of

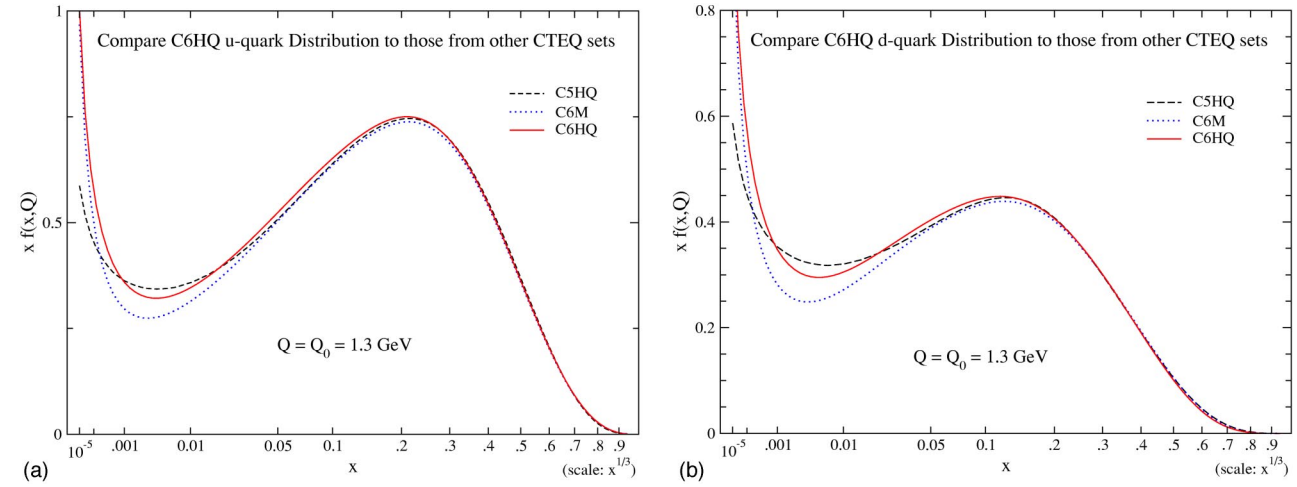


FIG. 4. Same as Fig. 3, except the axes are scaled to highlight the sea components of the PDFs: (a) u quark and (b) d quark.

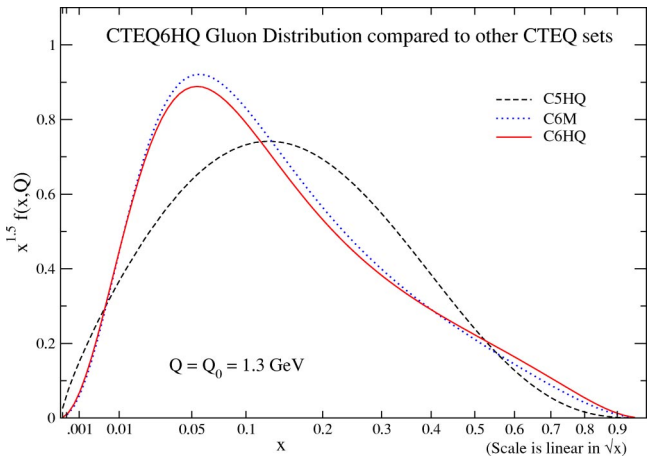


FIG. 5. Comparison of the gluon distributions at $Q_0 = 1.3$ GeV.

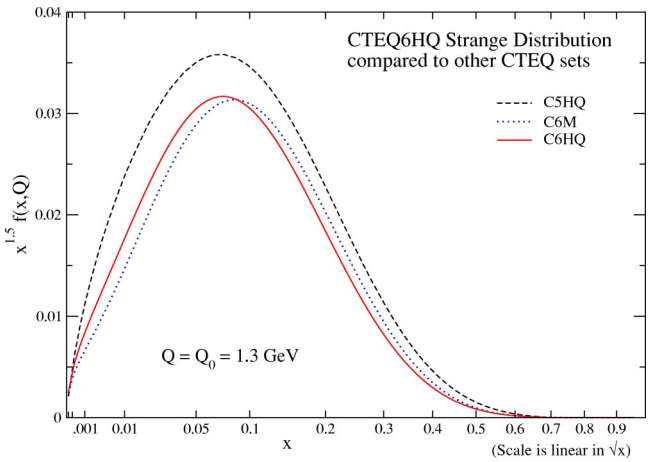


FIG. 6. Comparison of the strange distributions at $Q_0 = 1.3$ GeV.

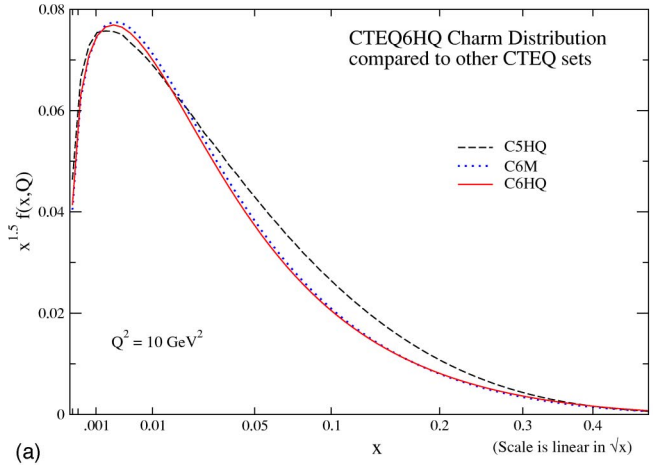
heavy quark effects in the generalized $\overline{\text{MS}}$ scheme. It is also interesting to compare the differences between the earlier (previous generation) CTEQ5HQ (C5HQ) distributions and the new CTEQ6HQ distributions; differences between these PDFs are attributable both to new data and to minor differences in the way the theoretical inputs are implemented. We combine these comparisons in the figures presented below.

A. Light partons at the input scale

Figures 3a and 3b show the *u* quark and *d* quark distributions from C5HQ, C6M, and C6HQ at the initial Q scale of 1.3 GeV (the charm mass). In order to exhibit the behavior of the PDFs at both large and small x values, the x axis is scaled linearly in $\sqrt{x}$. To give some additional physical insight into these plots, the y axis is chosen to be $x^{1.5}f(x,Q)$, so that the area under each curve represents the momentum fraction carried by that distribution. This chosen scaling then focuses on the valence peak and suppresses the small x sea excitations. We see that the valence *u* quark and *d* quark distributions are well constrained by the precision DIS data; the scheme dependence is not pronounced.

Figures 4a and 4b show the same PDFs, except the axes are scaled differently in order to highlight the sea quark distributions. The differences between the CTEQ5 and CTEQ6 generations of PDFs are a result of the improved low x data. When we compare the HQ and $\overline{\text{MS}}$ type PDFs of the CTEQ6 generation, we observe the general pattern that the heavy quark (HQ) distributions overshoot the $\overline{\text{MS}}$ ones at low $x(\sim 10^{-3})$, and that the distributions meet again at ultrasmall $x(\lesssim 10^{-5})$. We can qualitatively understand the larger sea quark distributions for the HQ scheme since these distributions are compensating for the suppression of charm in this scheme; this suppression is absent in the zero-mass $\overline{\text{MS}}$ scheme. In the region of ultrasmall x, there are no data constraining these distributions; we show this effect graphically in the next section where we consider the PDF uncertainty band of the HQ and $\overline{\text{MS}}$ distributions.

Figure 5 shows the comparison of the gluon distributions at Q_0. Here the difference between the C5HQ and the CTEQ6 generations of gluon distributions is pronounced. As discussed in Ref. [6], the change in this least-well-determined parton distribution is due to the recent precision DIS data (most influential in the small x region) in conjunc-



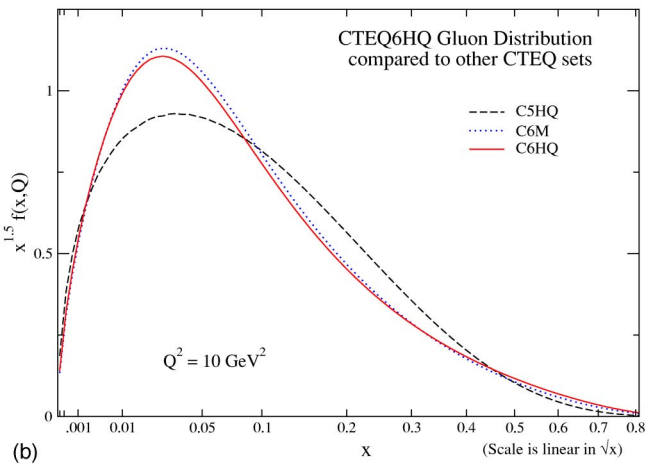


FIG. 7. Comparison of the charm and gluon distribution at $Q^2 = 10$ GeV^2.

tion with the greatly improved inclusive jet data from the Tevatron (critical for the medium to large x regions). The differences between C6HQ and C6M gluons at large x are due to a combination of scheme dependence and the inherent uncertainty range of the current analysis [6].

Figure 6 shows the comparison of the strange distributions at the same Q_0. The noticeable difference between the C5HQ curve and the others, in this case, is largely the result of different theoretical inputs: namely, the κ parameter [cf. Eq. (8)]. κ determines the ratio of strange to nonstrange sea quarks at the initial scale Q_0. This κ factor, known only approximately, was chosen to be 1/2 in both the CTEQ5 and CTEQ6 analyses, but for slightly different values of Q_0—1.0 GeV for C5HQ and 1.3 GeV for CTEQ6 sets.

B. Charm and gluon distributions at $Q^2=10$ GeV^2

To compare the differences of the charm distributions, we need to move some distance above the charm mass scale, since all current sets assume a zero charm distribution at $\mu = m_c$. Figure 7a makes this comparison at $Q^2=10$ GeV^2. We see a substantial difference between the C5HQ and the two CTEQ6 charm distributions. Unlike the case of the strange quark, this difference is physical; it is mainly a reflection of the difference in the gluon distributions (shown already for a lower Q value in Fig. 5) because charm is radiatively generated from the gluon. To confirm this hypothesis, Fig. 7b shows the gluon distributions at the same $Q^2 = 10$ GeV^2; the similarity is clear.

C. Q dependence of heavy quark mass effects

In this subsection, we compare the Q dependence of the C6M and C6HQ parton distributions at fixed values of x in order to see how the differences between PDFs in the zero-mass and general-mass schemes vary with increasing μ scale. We expect these differences (relative to the rising low x parton densities) will decrease with increasing Q. The purpose of this study is twofold. First, we check the self-consistency of the analysis, which is based on the expectation that power suppressed mass terms become unimportant asymptotically as $Q\to\infty$. Second, we find the specific scales where the differences due to mass effects become insignificant in practice. The latter question is relevant for high energy applications of PDFs derived from lower lower energy data.[16]

To examine the effects due to heavy flavor masses in the theoretical formalism, we cannot directly compare physical quantities that are being fitted—they are the same (within the errors of the global fit) by construction. On the other hand, parton distribution functions of individual flavors are not a good gauge either, because they are not physical—they are scheme dependent. Thus, for this comparison, we compromise and examine the quark singlet combination [$\Sigma\equiv\Sigma_q(q+\bar{q})$] and the gluon distribution.

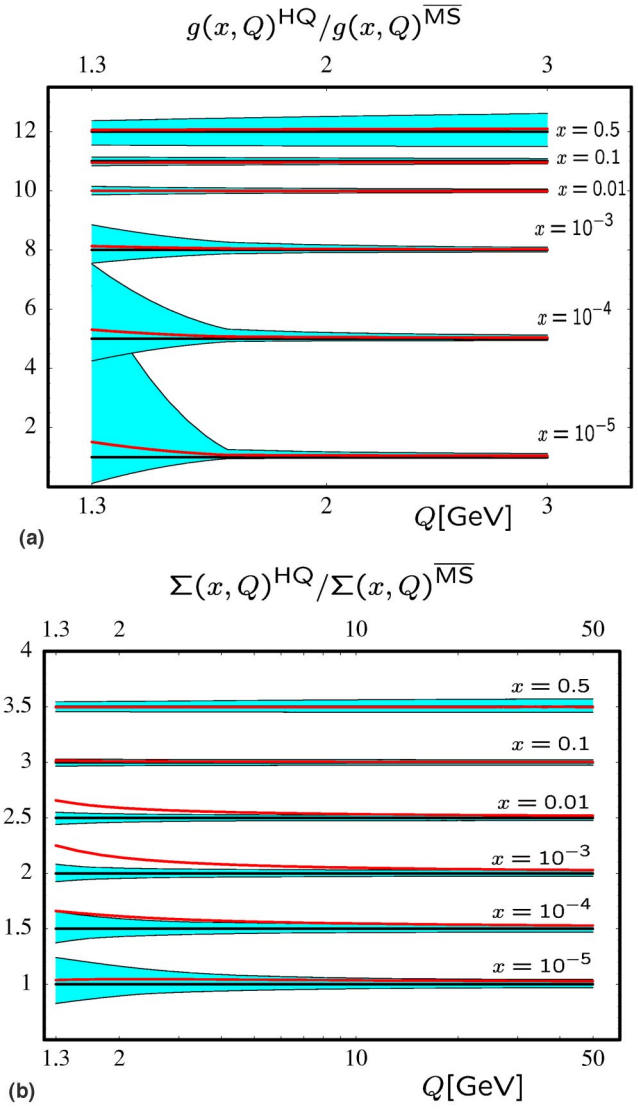


FIG. 8. Ratio of GM (C6HQ) to ZM (C6M) parton distributions as a function of Q (GeV) along with the uncertainty band of the latter. The offset on the y axis is arbitrary; i.e., the horizontal lines all correspond to $\Sigma(x,Q)^{\mathrm{HQ}}/\Sigma(x,Q)^{\overline{\mathrm{MS}}}=1$. Plotted are ratios for $x = \{10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 0.1, 0.5\}$ where x is increasing from the lowest to the uppermost curve. We display (a) the gluon and (b) the singlet quark $\Sigma\equiv\Sigma_q(q+\bar{q})$.

Figures 8a and 8b show, as solid lines, the C6HQ gluon and quark singlet distributions normalized to that of C6M. The shaded bands represent the ranges of uncertainty of the relevant quantities due to experimental sources, as estimated in the CTEQ6 global analysis [6]. We observe the following pattern: for most of the parameter space, the deviation of the C6HQ from C6M falls within the estimated uncertainty range. In particular, for the gluon distribution this is true for all values of x. For the quark distribution, the HQ fit overshoots the zero-mass fit significantly (i.e., beyond the uncertainty band) at moderately low $x\sim[10^{-2},10^{-3}]$. At very low $x\sim 10^{-5}$, the differences fall within the uncertainty bands again (primarily due to the increasing uncertainty). Although

[16]It also serves to validate the self-consistency of analyses where threshold effects have been neglected (such as for collider processes).

the relative difference between the C6HQ and C6M quark PDFs at $x \sim [10^{-2}, 10^{-3}]$ eventually becomes narrower as Q increases, the C6HQ fit does not evolve into the error band of the zero-mass fit. One may wonder why the two fits do not merge faster. The Wilson coefficients in GM-VFNS and ZM-VFNS differ by terms of order $\mathcal{O}(m^2/Q^2)$. Thus, the initial PDFs at Q_0, which are fitted to data, will differ by the same order of magnitude. Since the PDFs evolve logarithmically in the scaling variable, the differences in PDFs will persist over a much longer range of Q^2 than the differences in the power-law behaved Wilson coefficient terms. This effect is evident in the case of the singlet quark PDF at moderately low x values.

VI. CONCLUDING REMARKS

The CTEQ6HQ PDFs presented here complement the previously published CTEQ6 sets by providing distributions which can be used in the generalized $\overline{\text{MS}}$ scheme with nonzero-mass partons. This analysis includes the complete set of NLO processes including the real and virtual quark-initiated terms. Additionally, the ACOT(χ) scheme is used to introduce a generalized scaling variable which provides numerically stable results for the entire energy range—from the heavy quark thresholds to the high energy limit.

While the zero-mass-parton scheme is sufficient for many purposes, the fully massive scheme can be important when physical quantities are sufficiently sensitive to heavy quark contributions. This is evident when comparing the CTEQ6HQ and CTEQ6M fits to the mismatched sets (Table I) where the precise DIS data from HERA highlight the discrepancies.

The CTEQ6HQ fits also provide the basis for a series of further studies involving more quantitative analysis of strange, charm, and bottom quark distributions inside the nucleon. For example, the CTEQ6HQ PDFs are necessary for a consistent analysis of resummed differential distributions for heavy quark production such as in Ref. [37]. Using the full range of data from both the charged and neutral current processes, these distributions can reduce the uncertainties in the calculations; hence, they have significant implications for charm and bottom production and can help resolve questions about intrinsic heavy quark constituents inside the proton, the $\Delta x F_3$ structure function, and the extraction of $\sin\theta_W$.

ACKNOWLEDGMENTS

We thank our colleagues J. Huston, P. Nadolsky, J. Pumplin, and D. Stump, for fruitful collaboration on the CTEQ6 project that forms the foundation of this study. F.O. acknowledges the hospitality of MSU and BNL where a portion of this work was performed. S.K. is grateful to RIKEN, Brookhaven National Laboratory, and the U.S. Department of Energy (Contract No. DE-AC02-98CH10886) for providing the facilities essential for the completion of this work. This research was supported by the National Science Foundation (Grant No. 0100677) and by the Lightner-Sams Foundation.

[1] M.A. Aivazis, J.C. Collins, F.I. Olness, and W.K. Tung, Phys. Rev. D **50**, 3102 (1994).
[2] M. Kramer, F.I. Olness, and D.E. Soper, Phys. Rev. D **62**, 096007 (2000).
[3] J.C. Collins, Phys. Rev. D **58**, 094002 (1998).
[4] R.S. Thorne and R.G. Roberts, Phys. Rev. D **57**, 6871 (1998); Phys. Lett. B **421**, 303 (1998).
[5] A. Chuvakin *et al.*, Phys. Rev. D **61**, 096004 (2000).
[6] J. Pumplin, D.R. Stump, J. Huston, H.L. Lai, P. Nadolsky, and W.K. Tung, J. High Energy Phys. **07**, 012 (2002).
[7] J.C. Collins and W. Tung, Nucl. Phys. **B278**, 934 (1986).
[8] F.I. Olness and W. Tung, Nucl. Phys. **B308**, 813 (1988); M.A. Aivazis, F.I. Olness, and W. Tung, Phys. Rev. Lett. **65**, 2339 (1990).
[9] J. Amundson, C. Schmidt, W.K. Tung, and X. Wang, J. High Energy Phys. **10**, 031 (2000).
[10] E. Laenen, S. Riemersma, J. Smith, and W.L. van Neerven, Nucl. Phys. **B392**, 162 (1993).
[11] M. Glück, E. Reya, and M. Stratmann, Nucl. Phys. **B422**, 37 (1994).
[12] E. Witten, Nucl. Phys. **B104**, 445 (1976).
[13] J.C. Collins, F. Wilczek, and A. Zee, Phys. Rev. D **18**, 242 (1978).
[14] M. Buza *et al.*, Eur. Phys. J. C **1**, 301 (1998); Phys. Lett. B **411**, 211 (1997).
[15] M. Cacciari *et al.*, J. High Energy Phys. **05**, 007 (1998); M. Cacciari, S. Frixione, and P. Nason, *ibid.* **03**, 006 (2001).
[16] W.K. Tung, S. Kretzer, and C. Schmidt, J. Phys. G **28**, 983 (2002).
[17] S. Kretzer *et al.* (in preparation).
[18] S. Kretzer and I. Schienbein, Phys. Rev. D **58**, 094035 (1998).
[19] H1 Collaboration, C. Adloff *et al.*, Eur. Phys. J. C **13**, 609 (2000).
[20] H1 Collaboration, C. Adloff *et al.*, Phys. Lett. B **528**, 199 (2002).
[21] ZEUS Collaboration, J. Breitweg *et al.*, Eur. Phys. J. C **12**, 35 (2000).
[22] S.J. Brodsky, P. Hoyer, C. Peterson, and N. Sakai, Phys. Lett. **93B**, 451 (1980); S.J. Brodsky, C. Peterson, and N. Sakai, Phys. Rev. D **23**, 2745 (1981); B.W. Harris, J. Smith, and R. Vogt, Nucl. Phys. **B461**, 181 (1996); J.F. Gunion and R. Vogt, Report No. UCD-97-14, LBNL-40399, hep-ph/9706252.
[23] J. Pumplin, D. Stump, R. Brock, D. Casey, J. Huston, J. Kalk, H.L. Lai, and W.K. Tung, Phys. Rev. D **65**, 014013 (2002).
[24] D. Stump, J. Pumplin, R. Brock, D. Casey, J. Huston, J. Kalk, H.L. Lai, and W.K. Tung, Phys. Rev. D **65**, 014012 (2002).
[25] A.D. Martin, R.G. Roberts, W.J. Stirling, and R.S. Thorne, Eur. Phys. J. C **23**, 73 (2002); **28**, 455 (2003).
[26] ZEUS Collaboration, J. Breitweg *et al.*, Eur. Phys. J. C **12**, 35 (2000).

[27] NuTeV Collaboration, G.P. Zeller *et al.*, Phys. Rev. Lett. **88**, 091802 (2002).

[28] See, e.g., S. Davidson, S. Forte, P. Gambino, N. Rius, and A. Strumia, J. High Energy Phys. **02**, 037 (2002). A full list of references with possible interpretations of [27] is beyond the scope of the present publication.

[29] S. Kretzer, F.I. Olness, R.J. Scalise, R.S. Thorne, and U.-K. Yang, Phys. Rev. D **64**, 033003 (2001), and references therein.

[30] CHARM II Collaboration, P. Vilain *et al.*, Eur. Phys. J. C **11**, 19 (1999); CCFR Collaboration, A.O. Bazarko *et al.*, Z. Phys. C **65**, 189 (1995); A.O. Bazarko, Ph.D. thesis, Report No. NEVIS-1504; CCFR Collaboration, S.A. Rabinowitz *et al.*, Phys. Rev. Lett. **70**, 134 (1993); CDHSW Collaboration, H. Abramowicz *et al.*, Z. Phys. C **15**, 19 (1982); NOMAD Collaboration, P. Astier *et al.*, Phys. Lett. B **486**, 35 (1900); NuTeV Collaboration, M. Goncharov *et al.*, Phys. Rev. D **64**, 112006 (2001).

[31] CCFR/NuTeV Collaboration, U.K. Yang *et al.*, Phys. Rev. Lett. **86**, 2742 (2001).

[32] CCFR Collaboration, W.G. Seligman *et al.*, Phys. Rev. Lett. **79**, 1213 (1997); W. G. Seligman, Ph.D. thesis, Columbia University, Report No. NEVIS-292, 1997.

[33] M.A.G. Aivazis, F.I. Olness, and W.-K. Tung, Phys. Rev. Lett. **65**, 2339 (1990); M. Glück, S. Kretzer, and E. Reya, Phys. Lett. B **380**, 171 (1996); **405**, 391(E) (1996); V. Barone, M. Genovese, N. Nikolaev, E. Predazzi, and B. Zakharov, *ibid.* **328**, 143 (1994); V. Barone, U. D'Alesio, and M. Genovese, *ibid.* **357**, 435 (1995).

[34] As a point of entry, see, e.g., A.L. Kataev, in ''La Thuile 2001, Results and Perspectives in Particle Physics,'' pp. 205–221, hep-ph/0107247; J.T. Londergan and A.W. Thomas, Phys. Lett. B **558**, 132 (2003); S. Kumano, Phys. Rev. D **66**, 111301 (2002); S. Kovalenko, I. Schmidt, and J.-J. Yang, Phys. Lett. B **546**, 68 (2002).

[35] NuTeV Collaboration, R. Bernstein, presented at DIS2002, Cracow, 2002.

[36] NuTeV Collaboration, M. Goncharov *et al.*, Phys. Rev. D **64**, 112006 (2001); S. Kretzer, D. Mason, and F. Olness, *ibid.* **65**, 074010 (2002).

[37] P.M. Nadolsky, N. Kidonakis, F.I. Olness, and C.P. Yuan, Phys. Rev. D **67**, 074015 (2003).