CT18 global PDF fit at leading order in QCD

Mengshi Yan,[1,*] Tie-Jiun Hou,[2,†] Pavel Nadolsky,[3,4] and C.-P. Yuan[5]

[1]*Department of Physics and State Key Laboratory of Nuclear Physics and Technology, Peking University, Beijing 100871, China*
[2]*School of Nuclear Science and Technology, University of South China, Hengyang, Hunan 421001, China*
[3]*Department of Physics, Southern Methodist University, Dallas, Texas 75275-0175, USA*
[4]*Fermilab, P.O. Box 500, Batavia, Illinois 60510-5011, USA*
[5]*Department of Physics and Astronomy, Michigan State University, East Lansing, Michigan 48824, USA*

(Received 18 December 2022; accepted 19 May 2023; published 1 June 2023)

In this paper, we present a CT18 PDF global QCD analysis at leading order in QCD perturbation theory. The CT18 LO PDFs is obtained within the general CT18 framework, along with two additional procedures imposed to improve the quality of the fit. We take the W-boson charge asymmetry and inclusive single-top production at LHC as examples to illustrate implications of the CT18 LO PDFs.

DOI: 10.1103/PhysRevD.107.116001

I. INTRODUCTION

Parton distribution functions (PDFs) describe the structure of hadrons as composed of (anti)quarks and gluons. PDFs are needed to make predictions for hard scattering processes in high-energy collisions. Commonly used event generators, such as PYTHIA, still rely on simulations of parton showers using LO splitting kernels to avoid negative weights at higher orders [1,2] Although that current parton shower algorithms provide consistent descriptions with next-to leading order (NLO) matrix elements, or even next-to-next-to leading order (NNLO) matrix elements, with matching and merging methods [3–5], the LO PDFs are still needed for consistent predictions by parton showering programs. An alternative approach is to redistribute event weights, so that negative weight would be avoided [6–10]. Currently, with measurements at the Large Hadron Collider (LHC) becoming unprecedentedly precise, PDFs must be known at a high level of accuracy and precision. Such precise PDF parametrizations are provided by several groups [11–13] by taking advantage of predictions at the NNLO in QCD coupling α_s for a large number of collider processes. Meanwhile, the predictive power of the leading order (LO) QCD theory is no longer sufficient for today's precise measurements, so that one is required to go beyond LO for a good description of data. The limitation of the LO QCD theory translate into the difficulty of LO PDFs' determination, as some of the missing higher-order terms in an hard-scattering cross section are absorbed into the extracted LO PDFs.

Recent LO PDFs are available from the MSHT20 [12] and NNPDF4.0 [13] analyses. To improve the description of Drell-Yan (DY) processes, a K factor of $1+\alpha_s C_F\pi/2$ has been adopted in LO PDF fits of the MSHT family [12,14,15], where $C_F=4/3$ in QCD. The fit quality of MSHT20 LO is $\chi^2/N_{pt}=2.58$, which is worse than their previous LO results and the MSHT20 (N)NLO PDFs. The NNPDF4.0 LO is not able to fit experimental data well either, with $\chi^2/N_{pt}=3.35$. Such a bad quality of these fits is due in part to the inclusion of high precise LHC measurements, for which NNLO corrections to theory predictions are already essential for a precise description.

The optimized CTEQ-TEA PDFs for LO event generators [16], CT09MCS, CT09MC1, and CT09MC2 were obtained in 2010 by generalizing the conventional QCD global analysis. The CT09MC candidate PDFs were constrained by not only the real experimental datasets, but also the NLO pseudodatasets for representative LHC processes as a joint input to the global fit. These PDFs described the underlying event at the Tevatron and LHC at a reasonably good level. Specifically, in CT09MCS, the factorization scales in the LO matrix elements and the normalization for each pseudodatasets were allowed to vary to reach the best agreement with NLO pseudodata. In CT09MC1 and CT09MC2 PDFs, the momentum sum rule, which reflects the conservation of total momentum carried by partons, was relaxed, and the factorization scales for pseudodatasets were constant. The normalization to each pseudodataset was still fitted as in the CT09MCS analysis. Two PDFs sets, CT09MC1 and CT09MC2 PDFs, were determined with

*msyan@pku.edu.cn
†tjhou@msu.edu

1-loop and 2-loop expressions of the running strong coupling α_s, respectively. For the most recent LO set of CTEQ-TEA family, CT14 LO PDFs [17], two PDF sets were provided. One is obtained with LO α_s evolution and $\alpha_s(M_Z) = 0.130$ to make up for the insufficient PDF evolution, as compared to (N)NLO analysis. The other PDF set is obtained with NLO α_s evolution and $\alpha_s(M_Z) = 0.118$. The quality of fits χ^2/N_{pt} was 1.99, 2.17 for CT14 LO with LO α_s evolution and with NLO α_s evolution, respectively.

Turning now to the CT18 global analysis [11], a wide variety of precise LHC measurements, on top of the combined HERA I + II DIS [18] datasets and CT14 datasets, are used to determine CT18 NLO and NNLO PDFs. We expect that a LO fit within the CT18 framework would suffer from the same difficulty of describing high-precise LHC data without higher-order corrections as in the case of MSHT20 LO and NNPDF4.0 LO.

In the following, experimental datasets and the special settings to improve the results of fits from the limitation of LO perturbation theory are introduced in Sec. II. Section III describes the results of fits and analyzes the impact of the new LHC precision data. In Sec. IV, we study implications of the CT18 LO PDFs for some of the LHC observables, such as the W-boson charge asymmetry and single-top production. Section V contains our conclusions.

II. DESCRIPTION OF THE CT18 LO GLOBAL FIT

A. CT18 datasets

The CT18 QCD global analysis [11] is obtained by fitting a selection of high-statistics LHC datasets, the combined HERA I + II DIS dataset, and datasets already included in CT14 global QCD analysis [17], for a total of 40 datasets. The global fits are performed with NLO and NNLO perturbative QCD predictions. Both NLO and NNLO fits can describe this large dataset well, as quantified by χ^2/N_{pt} shown in Tables I and II. In Sec. II B, we will show that a subset of the CT18 dataset, particularly consisting of high-precision LHC datasets from run-2, is difficult to fit with LO QCD theory.

Out of 11 high-precision LHC datasets included in the CT18 analyses, five datasets correspond to W and Z vector boson production. From ATLAS, there is the 8 TeV transverse momentum p_T of lepton pairs distribution in Z/γ^* production with 20.3 fb^{-1} (ID = 253) [26]. From CMS, the 8 TeV muon charge asymmetry A_{ch} for inclusive $W^{\pm}$ production with 18.8 fb^{-1} (ID = 249) [24] is included. From LHCb, the datasets included are 7 TeV W and Z forward rapidity cross section distribution measurement with 1.0 fb^{-1} (ID = 245) [22], 8 TeV $Z \to e^+e^-$ forward rapidity cross section with 2.0 fb^{-1} (ID = 246) [23], and 8 TeV W and Z production cross section with 2.0 fb^{-1} (ID = 250) [25]. The ATLAS 7 TeV W and Z combined cross section (ID = 248) dataset [42] is not included in the nominal CT18, since this dataset is observed to have tension with other datasets (Sec. II C of [11]). Alternative PDF sets, CT18A and CT18Z, have been generated with the inclusion of ATLAS $\sqrt{s} = 7$ TeV W and Z dataset.

In the CT18 LO analysis, we start from the CT18 datasets, without the inclusion of the above mentioned ATLAS $\sqrt{s} = 7$ TeV W and Z dataset. Then, as will be shown in the next section, we will exclude five datasets that cannot be correctly described at LO.

B. Special procedures adopted in the CT18 LO fit

Theory predictions at LO are less complicated than at higher orders, but they miss some vital quantum corrections for describing the most precise experimental data. Many NLO predictions of LHC cross sections tend to be larger than predictions at LO. Predictions of differential cross sections with LO PDFs and LO matrix elements are unreliable also in terms of shape. See, for example, Fig. 1 of Ref. [16] for a comparison between LO and NLO predictions for SM boson rapidity distributions at the

TABLE I. Quality of fit to the datasets for which $\chi^2/N_{\rm pt}$ values of CT18 LOpert are larger than 6. The values with an asterisk are $\chi^2/N_{\rm pt}$ of the datasets that are not included in the corresponding fit. Totally, the number of data points in CT18 LO is 3547, which is reduced by the first choice from the number of data points in CT18 LOpert and (N)NLO, 3861.

ID	Dataset	CT18 LO	CT18 LOpert	CT18 NLO	CT18 NNLO
145	H1 σ_r^b [19]	6.14*	6.26	1.49	0.68
147	Combined HERA charm production [20]	21.14*	11.54	0.80	1.24
169	H1 F_L [21]	17.15*	17.15	0.77	1.89
245	LHCb 7 TeV W/Z rap. [22]	5.85	8.36	2.29	1.63
246	LHCb 8 TeV Z → ee rap. [23]	5.84	11.06	2.09	1.00
249	CMS 8 TeV W A_{ch} [24]	2.17	9.14	0.60	1.03
250	LHCb 8 TeV W/Z rap. [25]	10.59	13.61	3.36	2.17
253	ATLAS 8 TeV Z p_T^{ll} [26]	19.21*	19.20	2.06	1.12
	Total	1.60	2.15	1.17	1.17

TABLE II. Similar to Table I, quality of fits $\chi^2/N_{\rm pt}$ for DIS and jet datasets in the CT18 dataset are compared.

ID	Dataset	CT18 LO	CT18 LOpert	CT18 NLO	CT18 NNLO
160	HERAI + II 1 fb^{-1} H1 and ZEUS NC and CC e$^{\pm}$p reduced cross sec. comb. [18]	1.72	1.64	1.22	1.26
101	BCDMS F_2^p [27]	1.16	1.26	1.08	1.11
102	BCDMS F_2^d [28]	1.29	1.64	1.13	1.12
104	NMC F_2^d/F_2^p [29]	1.17	1.43	0.97	1.02
108	CDHSW F_2^p [30]	1.09	1.52	0.90	1.01
109	CDHSW F_3^p [30]	1.22	1.42	0.82	0.90
110	CCFR F_2^p [31]	1.78	2.55	1.15	1.14
111	CCFR xF_3^p [32]	0.53	0.80	0.41	0.39
124	NuTeV $\nu\mu\mu$ SIDIS [33]	1.02	2.11	0.52	0.49
125	NuTeV $\bar{\nu}\mu\mu$ SIDIS [33]	1.73	2.46	1.03	1.17
126	CCFR $\nu\mu\mu$ SIDIS [34]	0.60	1.43	0.79	0.75
127	CCFR $\bar{\nu}\mu\mu$ SIDIS [34]	0.58	1.15	0.54	0.52
504	CDF run-2 inclusive jet production [35]	1.54	1.46	1.49	1.70
514	DØ run-2 inclusive jet production [36]	1.22	1.38	1.07	1.03
542	CMS 7 TeV 5 fb^{-1} single incl. jet cross sec., $R = 0.7$ (extended in y) [37]	1.41	1.56	1.23	1.23
544	ATLAS 7 TeV 4.5 fb^{-1} single incl. jet cross sec., $R = 0.6$ [38]	1.50	1.58	1.40	1.45
545	CMS 8 TeV 19.7 fb^{-1} single incl. jet cross sec., $R = 0.7$, (extended in y) [39]	1.65	1.78	1.10	1.14
573	CMS 8 TeV 19.7 fb^{-1} $t\bar{t}$ norm. double-diff. top p_T and y cross sec. single [40]	2.05	2.38	1.56	1.18
580	ATLAS 8 TeV 20.3 fb^{-1} $t\bar{t}$ p_T^t and $m_{t\bar{t}}$ absolute spectrum [41]	2.41	3.38	1.29	0.63
	Total	1.60	2.15	1.17	1.17

LHC. To resolve this difficulty, the conventional approach for PDF determination needs to be extended.

The experimental uncertainty in the LHC run-2 datasets from vector boson production is at a percent level; a LO fit with the inclusion of these datasets is difficult. To illustrate this point, a PDF set named CT18 LOpert has been generated, where all theoretical predictions are computed at LO, and no other adjustments have been applied. In Tables I and II, the CT18 LOpert presents undesirable qualities of fit to both total and specific datasets. In order to improve the fit at the LO, in our final PDF set, CT18 LO, we have made the following two special choices:

(1) From the CT18 dataset, we exclude ID 169 H1 F_L [18], ID 145 H1 bottom reduced cross section [19], ID 147 combined HERA charm production [20], ID 253 ATLAS 8 TeV Z boson p_T^{ll} distribution [26], and ID 268 ATLAS 7 TeV W and Z bosons rapidity distribution plus W charge asymmetry distribution [43], since these datasets cannot be well described by the QCD theory at the leading order. Furthermore, the ID248 ATLAS 7 TeV precision W and Z data [42] were not included in the nominal CT18 NLO and NNLO analysis, but in the alternative CT18A and CT18Z NLO and NNLO PDFs sets [11]. After excluding all the above-mentioned datasets, the total number of remaining data points is 3547, which is 134 points less than in the CT18 dataset.

(2) In the rest of Drell-Yan datasets, for inclusive production of either $W^{\pm}$ or Z bosons, we multiply the LO hard cross sections by a K factor $K(Q)$ given in Eq. (1) to partially make up for the deficit of production rate,

$$K(Q) = 1 + \frac{\alpha_s(Q)}{\pi}\frac{C_F\pi^2}{2}. \tag{1}$$

.

To explain the first choice, we first note that, at the leading order $\mathcal{O}(\alpha_s^0)$, the longitudinal structure function $F_L = F_2 - 2xF_1$ respects the Callan-Gross relation [44],

$$F_2 - 2xF_1 = 0. \tag{2}$$

Beyond the LO, the gluon emission gives rise to non-vanishing F_L, and the Callan-Gross relation is violated. In the CT18 dataset, the H1 F_L data measure the longitudinal structure function: a LO PDF fit cannot predict this data. The large $\chi^2/N_{\rm pt}$ values of the H1 F_L data for CT18 LO and CT18 LOpert in Table I reflect the inconsistency between the H1 F_L data and the Callan-Gross relation $F_L = 0$ predicted by LO theory. Hence, we exclude this dataset from the LO fit. The H1 bottom reduced cross section and the combined HERA charm reduced DIS cross sections are sensitive to higher-order QCD corrections and the mass of heavy partons [45–49]. Therefore, we also exclude these two datasets. Moreover, a LO QCD calculation, at $\mathcal{O}(\alpha_s)$, cannot describe well the ATLAS 8 TeV Z boson p_T^{ll} distribution [26] because of the presence of a large logarithm $\ln(M_Z/p_T^{ll})$. Likewise, a QCD calculation

at $\mathcal{O}(\alpha_s)$ cannot describe well the inclusive W or Z production cross sections, if asymmetric kinematic cuts are applied to the two decay leptons, such as in the ATLAS charge asymmetry measurement. Needless to say that, at $\mathcal{O}(\alpha_s^0)$, the p_T distribution of the Drell-Yan pair produced at the LHC is a delta function with a peak at zero, and the two decay leptons must have the same transverse momenta, as predicted by the parton model with longitudinal PDFs. Hence, these datasets are also excluded from our LO fits.

The importance of the second choice will be discussed in Sec. III B. In the LO PDF studies of the MSHT family [12,14,15], this K factor, Eq. (1), was adopted to help describe vector boson production data. Apart from these two special choices, no other modifications in the fitting procedure were made. As for the strong coupling α_s, it is expected that a LO PDF analysis would prefer a larger value of α_s. We decided to fix it to be $\alpha_s(m_Z) = 0.135$ at the Z boson mass scale, which yields the best quality of fit at the LO. The α_s dependence of CT18 LO fit will be discussed in Sec. III C.

III. RESULTS

A. Quality of the fit

Goodness-of-fit figures, χ^2/N_{pt}, for individual datasets are summarized in Tables I and II. The overall χ^2/N_{pt} of CT18 LO is 1.60, significantly improved from 2.15, the total χ^2/N_{pt} of CT18 LOpert fit. The fit to the CT18 LO dataset is clearly improved by the steps described in Sec. II B, though it is still worse than in the CT18 NLO and NNLO fits.

After multiplication by the Drell-Yan K factor, Eq. (1), the high-precision LHC W and Z bosons production datasets in Table I acquire a better, still too high, χ^2/N_{pt} in CT18 LO. Fits to DIS and jet datasets shown in Table II are mostly improved from CT18 LOpert; again, it is difficult to obtain a good fit with $\chi^2/N_{pt} \sim \mathcal{O}(1)$ at this order. In CT18 LO, the ID160 HERA I + II combined reduced cross section [18] has a slightly larger χ^2/N_{pt} than in CT18 LOpert. This increase only comes from several data points with very low energy scale $Q = 2.121$ GeV in the neutral-current channel of e^+p scattering, where the correlated systematic errors pull the central values of data points away from theory prediction. In the CTEQ-TEA program, the best-fit value of χ^2 is the combination of the best fit χ^2 to the shifted data and the contribution from the optimal nuisance parameters [11,50]. The optimal nuisance parameters for these low energy neutral-current HERA data points thus have very large values, suggesting that there is a noticeable systematical bias in the CT18 LO fit to these data points. The agreement with the rest of HERA data points is about the same in the CT18 LO and LOpert fits. A similar phenomenon, that one data point χ^2/N_{pt} increases from CT18 LOpert to CT18 LO, also happens to ID504 CDF run-2 inclusive jet production [35].

TABLE III. The second moments of CT18 LO and CT18 NLO PDFs at the initial scale $Q_0 = 1.3$ GeV. The numbers showing in brackets are the CT18 NLO uncertainties on the last quoted decimal at the 68% confidence level. The CT18 NLO uncertainties for $\langle x\rangle_{\bar{u}}$, $\langle x\rangle_{\bar{d}}$, and $\langle x\rangle_s$ are negligible.

PDFs	$\langle x\rangle_u$	$\langle x\rangle_d$	$\langle x\rangle_{u_v}$	$\langle x\rangle_{d_v}$
CT18 LO	0.34	0.16	0.31	0.12
CT18 NLO	0.35(4)	0.17(4)	0.32(4)	0.13(4)
	$\langle x\rangle_{\bar{u}}$	$\langle x\rangle_{\bar{d}}$	$\langle x\rangle_s$	$\langle x\rangle_g$
CT18 LO	0.03	0.04	0.02	0.40
CT18 NLO	0.03(0)	0.04(0)	0.01(0)	0.39(1)

B. Functional dependence and moments of PDFs

In Table III, we summarize the second moments $\langle x\rangle$ that quantify the momentum carried by individual parton flavors. The moments are computed at the initial scale Q_0. In comparison to the CT18 NLO central values, the strange and gluon second moments are increased to compensate for the missing higher-order contributions to the Wilson coefficients of a number of scattering processes. Under the total momentum sum rule, all the other flavors are allocated with less momenta than in CT18 NLO, as more momentum is carried by the gluon and strange distributions. The CT18 LO second moments at Q_0 scale in general are consistent with CT18 NLO within 1 standard deviation, except for the strange PDF. Without higher-order corrections, a better determination of strange PDF is difficult, as to be seen below. In terms of the PDF configuration, the gluon and strange distributions of CT18 LO are enhanced, while the bumps in valence quarks are reduced. The figures of PDF comparison between CT18 LO and NLO are available at the website [51].

The Fig. 1 shows comparisons of CT18 LO PDFs to MSHT20 LO [12] and NNPDF4.0 LO [13] PDFs at 2.0 GeV. We note that we do not generate the CT18 PDF error sets at the LO, as done in the CT18 (N)NLO fits, because the LO theoretical uncertainty is large. Three LO PDFs are found to be inconsistent with each other across a wide range of x at 2.0 GeV. As shown in Fig. 2, the differences among three LO PDFs at the electroweak scale are reduced as a result of the DGLAP evolution. But their differences are still significant in typical LHC measurements. More figures for the comparison of LO PDFs are available at the website [51].

Today, vector boson production at the LHC can be measured so precisely that total experimental uncertainty is at a percent level. The Born-level cross section is not capable of describing precise experimental data as accurately as those beyond the leading order. In Fig. 3, we

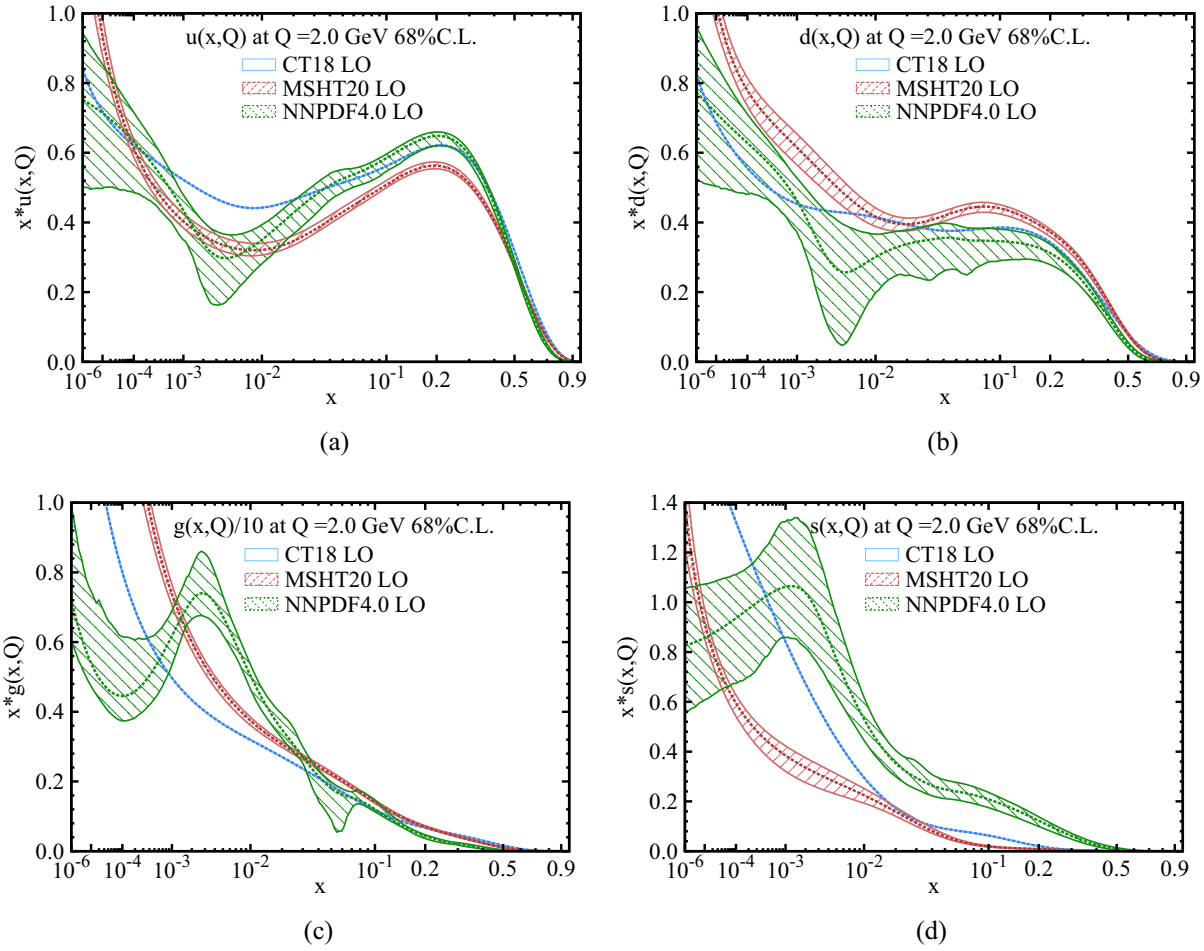


FIG. 1. A comparison of $u(x)$, $d(x)$, $s(x)$, and $g(x)$ PDFs at 2.0 GeV for CT18 LO, MSHT20 LO [12], and NNPDF4.0 LO [13].

illustrate the impact of these gauge boson production datasets, by comparing a series of LO fits, in which LHC run-II vector boson production datasets, namely, ID 245, 246, 249, 250, and 253, are treated with different prescriptions. In CT18 LOpert, the strange PDF vanishes in the range $2 \times 10^{-4} < x < 0.02$ at $Q_0 = 1.3$ GeV under the influence of these datasets. When evolved to 100 GeV, the strange distribution is still quite small in this range of x. We note that PDFs at the initial scale Q_0 in CT analyses are parametrized to be non-negative, so that in CT18 LOpert $s(2 \times 10^{-4} < x < 0.02, Q_0) = 0$. This is to ensure that the predicted LO cross sections are always positive. Such a strong suppression of the strange PDF is resolved in CT18 LOpert-lw0.01, where these Drell-Yan datasets are weighted lightly (by a factor of 0.01, instead of the nominal value of 1), and in CT18 LO by applying the Drell-Yan K factor, Eq. (1), to those Drell-Yan datasets. As an example, in Fig. 4, we compare CT18 LO, CT18 LOpert, and CT18 NLO predictions to experimental data for ID 245 LHCb W and Z boson production at 7 TeV [22]. The theory predictions are calculated by using the APPLgird package [52]. Without applying the K factor, Eq. (1), CT18 LOpert cannot provide enough cross sections for either W or Z production. In CT18 LO, the prediction of overall magnitude of this process is improved by the Drell-Yan K factor and is consistent with predictions by CT18 NLO. Consequently, the s-quark PDF is not as suppressed as in CT18 LOpert. But improving the shape of rapidity distribution still requires NLO QCD corrections.

C. α_s dependence of CT18 LO

The strong coupling constant is one of key elements in computing theory predictions, and hence, fed as an input into a PDFs global fit. In the PDF fits beyond LO, it is widely accepted [11–15,17,53] that the value of the strong coupling at Z-boson mass scale is fixed at its PDG value $\alpha_s(m_Z) = 0.118$ [54]. Due to the missing important quantum corrections and to generate more sea quarks via parton evolution, the value of $\alpha_s(m_Z)$ in LO fits is often fixed at a higher value than $\alpha_s(m_Z) = 0.118$. A number of LO PDFs [12,17,53] take $\alpha_s(m_Z)$ to be at 0.130. For NNPDF4.0 [13], NNPDF3.1 [53], and CT14 [17], LO PDFs with $\alpha_s(m_Z) = 0.118$ are also provided. In MSTW08 LO [14] and MMHT14 LO [15], $\alpha_s(m_Z)$ is fixed at 0.140 and 0.135, respectively.

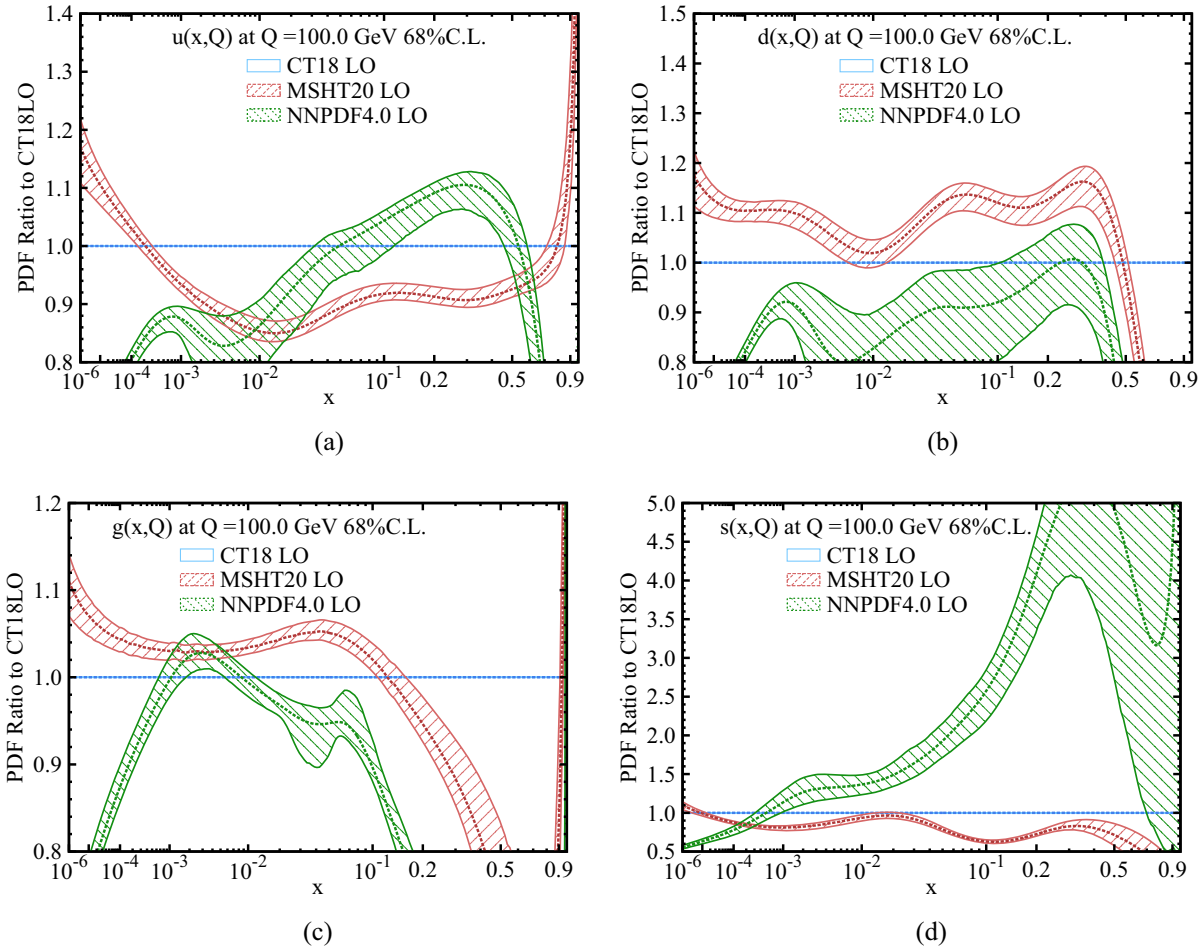


FIG. 2. The comparison of $u(x)$, $d(x)$, $s(x)$, and $g(x)$ PDF ratios to CT18 LO at 100 GeV for CT18 LO, MSHT20 LO [12], and NNPDF4.0 LO [13].

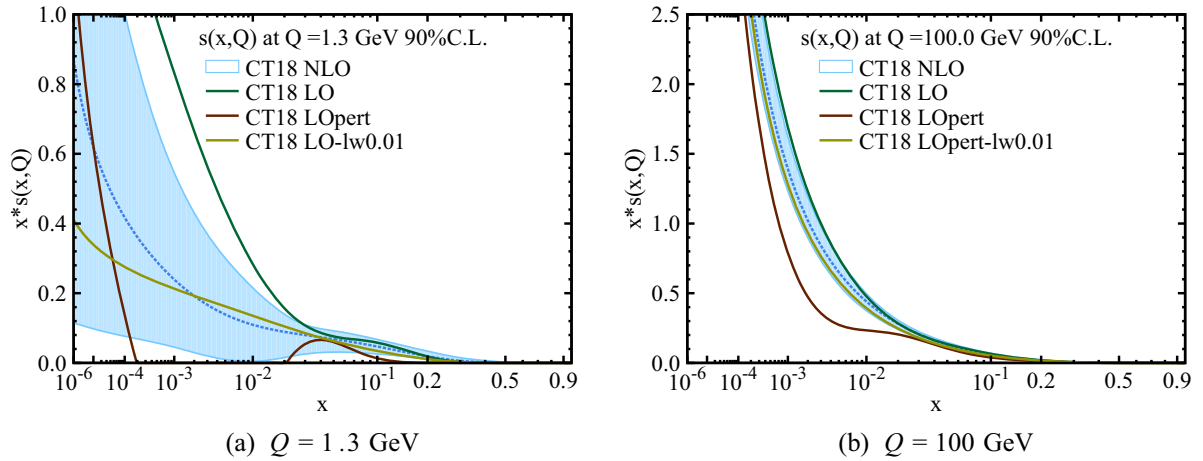
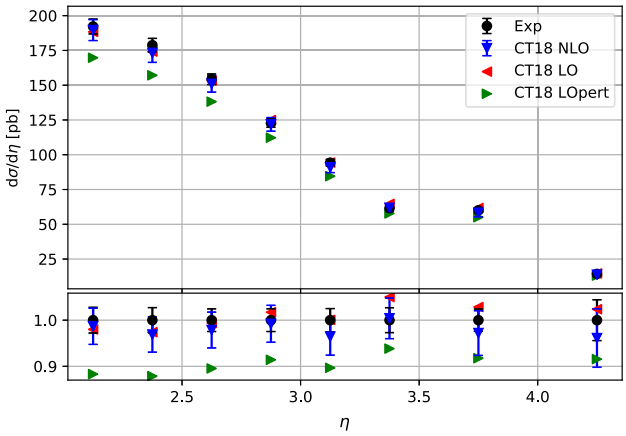
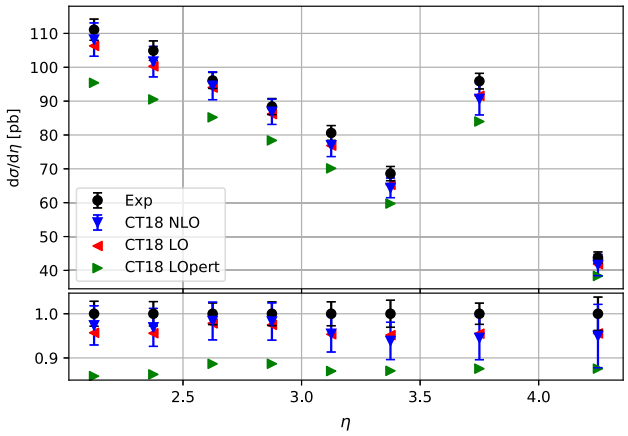


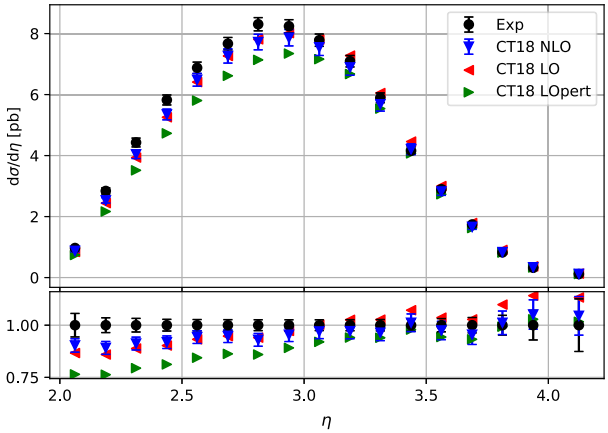
FIG. 3. A comparison of $s(x)$ at $Q_0 = 1.3$ GeV and 100 GeV among CT18 LO, CT18 LOpert, CT18 LOpert-lw0.01, and CT18 NLO. In CT18 LOpert-lw0.01, ID 245, 246, 249, 250, 253 datasets are assigned a weight of 0.01. In CT18 LOpert, i.e., for the brown curves in both panels, these datasets have a unit weight, as in CT18 LO and CT18 NLO.



(a) W^+ production



(b) W^- production



(c) Z production

FIG. 4. Predictions for the rapidity distribution of $W^{\pm}$ and Z production at LHCb 7 TeV, ID 245, by CT18 LO, CT18 LOpert, and CT18 NLO. Theory predictions are also compared to the unshifted experimental values.

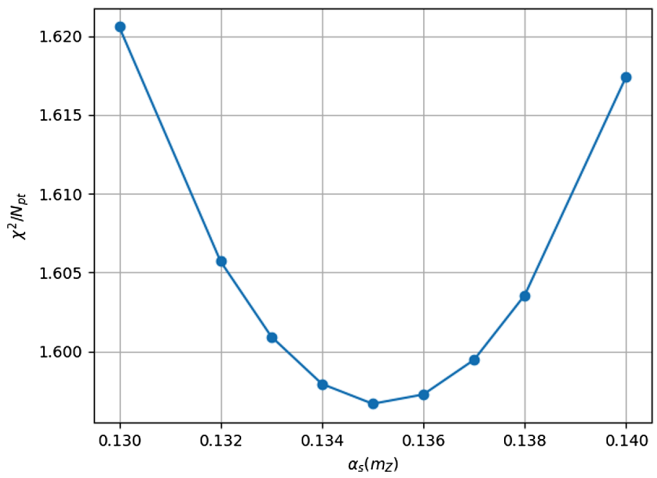


FIG. 5. The $\chi^2/N_{\rm pt}$ scan over various values of $\alpha_s(m_Z)$ for CT18 LO.

The input value of $\alpha_s(m_Z)$ of CT18 LO PDFs is set to be equal to its value providing the optimal description of the data. We scan over various values of $\alpha_s(m_Z)$ by performing a series of fits with all the other theoretical and experimental setups identical to those of CT18 LO. Figure 5 shows the value of $\chi^2/N_{\rm pt}$ of global fits while varying the value of $\alpha_s(m_Z)$. As shown in the figure, the best-fit value of $\alpha_s(m_Z)$ is found to be around 0.135. Compared to the usual choice $\alpha_s(m_Z) = 0.130$ for most of the LO PDFs, a CT18 LO PDFs fit, in which $\alpha_s(m_Z) = 0.135$, could reduce the $\chi^2/N_{\rm pt}$ by 0.024, equivalent to about 84 units for total χ^2. A larger value of $\alpha_s(m_Z)$ than 0.135 does not improve the fit to the datasets.

IV. PHENOMENOLOGY

In this section, we examine implications of the CT18 LO PDFs for some LHC predictions. In these computations, the input physical parameters are set as follows:

$$m_W = 80.385\ \text{GeV}, \qquad G_F = 1.16639 \times 10^{-5}\ \text{GeV}^{-2},$$
$$\Gamma_W = 2.06\ \text{GeV}. \tag{3}$$

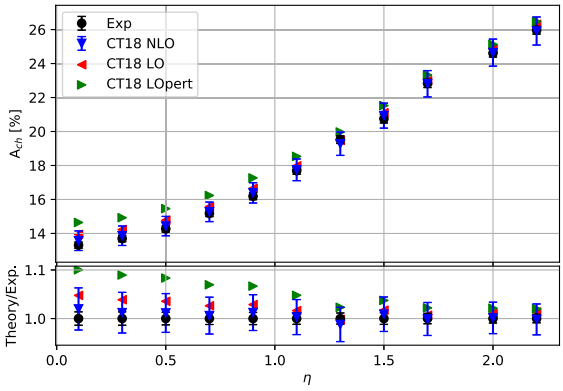
A. Charge asymmetry A_{ch} in W boson production

The charge asymmetry in W boson production provides important constraint on the valence and sea quark distribution functions. Figure 6 compares experimental data points and predictions for ID 249 CMS muon charge asymmetry A_{ch} at 8 TeV [24], as a function of the pseudorapidity of muon from W boson decay. The $pp \to W^{\pm} + X \to \mu^{\pm}\nu + X$ differential cross sections rapidity distributions were calculated with APPLgrid [52].

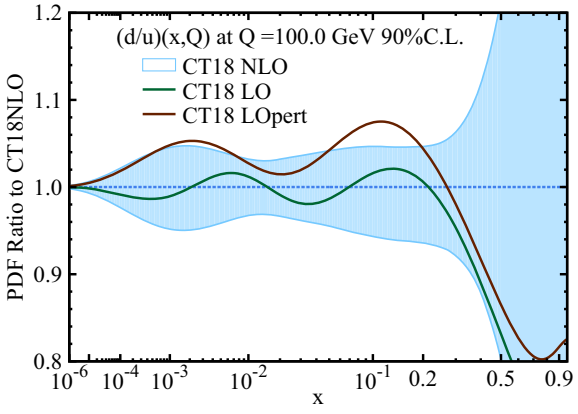
As shown in the panel (a) of Fig. 6, predictions by CT18 LO exhibit a different shape compared to the experimental data. In the lower rapidity region, the CT18 LO predictions are about ~5% higher than the measurements. For larger rapidity, the CT18 LO predictions for A_{ch} tend to be more consistent with measurements. On the other hand, the CT18 LO predictions for muon charge asymmetry generally lie within the CT18 NLO uncertainty band over whole rapidity range. CT18 LOpert prediction for A_{ch}, particularly for $\eta < 1.0$, is inconsistent with both the CMS measurement and CT18 NLO prediction. In the bottom panels of Fig. 6, the $(d/u)(x)$ and $(\bar{d}/\bar{u})(x)$ PDF ratios at 100 GeV are compared among CT18 LO, CT18 LOpert, and CT18 NLO, to which the charge asymmetry A_{ch} is sensitive. The CT18 LO PDF ratios are mostly consistent with CT18 NLO PDF ratios, except for the $(\bar{d}/\bar{u})(x)$ in range $0.01 < x < 0.03$. The $(\bar{d}/\bar{u})(x)$ ratio of CT18 LOpert is consistent with that of CT18 NLO, while the $(d/u)(x)$ PDF ratio for CT18 LOpert is close to the upper bound of CT18 NLO uncertainty.

B. Single-top production

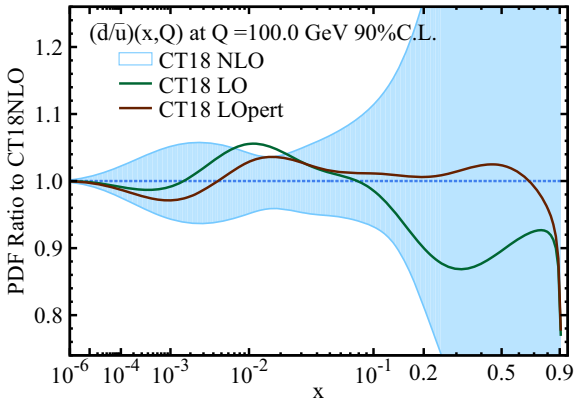
The total cross section for t-channel inclusive single-top production is a representative observable to study the implication of CT18 LO PDFs. It has been measured at the LHC [55–63] at various center-of-mass energies. The theoretical calculation of this process could serve as a test of the consistency in PDFs at different perturbation orders [64,65], since the total inclusive cross sections at different



(a) predictions to ID 249



(b) ratios of $\bar{d}/\bar{u}$ at 100 GeV



(c) ratios of $\bar{d}/\bar{u}$ at 100 GeV

FIG. 6. Panel (a): similar to Fig. 4, but for CMS W-boson charge asymmetry A_{ch}, ID 249. The PDF uncertainty in predictions corresponds to 68% confidence level. Panel(b): a comparison of the ratio $(d/u)(x)$ in between CT18 LO and CT18 NLO at 100 GeV. Panel(c): similar to panel (c) but for $(\bar{d}/\bar{u})(x)$.

TABLE IV. Total cross sections of inclusive single-top quark and antiquark production in pp collision at 14 TeV at LO and NLO for various PDF sets. The PDF uncertainty corresponds to the 68% confidence level.

PDFs	$\alpha_s(m_Z)$	$\sigma_{\rm LO}(t)$ [pb]	$\sigma_{\rm LO}(\bar{t})$ [pb]	$\sigma_{\rm NLO}(t)$ [pb]	$\sigma_{\rm NLO}(\bar{t})$ [pb]
CT18 LO	0.135	153.5	93.5	⋯	⋯
CT18 NLO [11]	0.118	⋯	⋯	156.8 ± 2.5	93.4 ± 1.6
CT14 LO [17]	0.118	137.6	80.6	⋯	⋯
CT14 NLO [17]	0.118	⋯	⋯	157.0 ± 3.3	93.2 ± 1.9
MSHT20 LO [12]	0.130	142.3 ± 1.0	97.3 ± 0.6	⋯	⋯
MSHT20 NLO [12]	0.118	⋯	⋯	156.5 ± 1.1	94.7 ± 0.7
NNPDF4.0 LO [13]	0.118	133.6 ± 1.1	83.2 ± 0.8	⋯	⋯
NNPDF4.0 NLO [13]	0.118	⋯	⋯	155.1 ± 0.9	93.7 ± 0.5
NNPDF3.1 LO [53]	0.118	141.7 ± 1.8	85.4 ± 2.0	⋯	⋯
NNPDF3.1 LO [53]	0.130	153.5 ± 1.8	92.6 ± 1.6	⋯	⋯
NNPDF3.1 NLO [53]	0.118	⋯	⋯	155.2 ± 1.2	93.4 ± 0.7
NNPDF3.0 LO [69]	0.118	149.1 ± 13.1	90.5 ± 8.5	⋯	⋯
NNPDF3.0 LO [69]	0.130	159.4 ± 7.7	96.6 ± 5.1	⋯	⋯
NNPDF3.0 NLO [69]	0.118	⋯	⋯	162.6 ± 1.9	99.2 ± 1.5
HERAPDF20 LO [18]	0.130	154.6 ± 0.8	92.1 ± 0.7	⋯	⋯
HERAPDF20 NLO [18]	0.118	⋯	⋯	162.5 ± 0.9	96.5 ± 0.7

orders are all expected to reproduce the data that is used in PDF fits. We make use of this property to illustrate the consistency of CT18 LO with CT18 NLO. In our calculation, the t-channel inclusive single-top production cross section is computed by MCFM [66–68]. For this calculation, the input parameters take the values shown in Eq. (3), along with the renormalization and factorization scales chosen as $\mu_R = \mu_F = m_t$.

Table IV presents predictions for the t-channel inclusive single-top production with a variety of PDFs. In general, due to the lack of higher-order corrections, the LO predictions for this process tend to be smaller than the corresponding NLO ones. The CT18 LO prediction of single top quark production is slightly outside of the CT18 NLO uncertainty band, while they give consistent predictions for single top antiquark production. Comparing to CT14 LO predictions, the CT18 LO ones have increased substantially and better agree with the NLO.

V. CONCLUSION

In this paper, we presented CT18 LO PDFs, which were obtained within the general framework of CT18 global analysis with two modifications described in Sec. II B. One is to discard some datasets, that cannot be properly described at LO (such as Drell-Yan data with different cuts on the transverse momemta of the two final-state leptons). The other is to apply a K factor to predictions for Drell-Yan processes to make up for the insufficiency of LO predictions. As a result, the quality of the LO fit with these two choices, called CT18 LO, is substantially improved compared to a naive LO fit, called CT18 LOpert. In CT18 LOpert, the strange quark distribution behave unphysicallyas shown in Fig. 3. This is because the LO Wilson coefficients cannot provide enough normalization to describe the high-precision W and Z production data from LHC run-2, as the NLO K factors being typically positive. When fitting the vector boson production, the up and down sea quark distributions is driven by data to increase their magnitudes to compensate for the deficiency in the LO Wilson coefficients. Since all flavors are correlated under the total momentum sum rule, the magnitude of the strange PDF has to be reduced when that of the others is increased. The strong suppression of $s(x)$ at small x is relieved in CT18 LO via the implementation of the Drell-Yan K factor. We have checked that the CT18 LO predictions are close to the CT18 NLO ones for the rapidity distribution of charge asymmetry in W-boson production and the total cross section of t-channel inclusive single-top production. Still, we stress that the CT18 LO PDFs are inferior to CT18 (N)NLO PDFs in many aspects and should not be used except for order-of-magnitude estimation. We do not provide error sets for CT18 LO because of its very large theoretical uncertainties. In cases when a leading-order uncertainty estimation for a given experimental observable is needed, we recommend to quote the uncertainty predicted by the CT18 NLO PDF error sets. Variations in QCD scales and α_s value contribute as additional sources of uncertainties.

We will make available the central CT18 LO PDFs described above as a part of the LHAPDF library [70,71] and at the CTEQ-TEA website [51,72].

ACKNOWLEDGMENTS

We thank our CTEQ-TEA colleagues for useful discussions. P. M. N. is partially supported by the

U.S. Department of Energy under Grant No. DE-SC0010129 and by the Fermilab URA award. The work of C. P. Y. is partially supported by the U.S. National Science Foundation under Grant No. PHY-2013791. C. P. Y is also grateful for the support from the Wu-Ki Tung endowed chair in particle physics.

[1] A. Buckley *et al.*, General-purpose event generators for LHC physics, Phys. Rep. **504**, 145 (2011).
[2] S. Amoroso *et al.* (HSF Physics Event Generator WG Collaboration), Challenges in Monte Carlo event generator software for high-luminosity LHC, Comput. Software Big Sci. **5**, 12 (2021).
[3] J. Bellm *et al.*, HERWIG 7.2 release note, Eur. Phys. J. C **80**, 452 (2020).
[4] E. Bothmann *et al.* (Sherpa Collaboration), Event generation with SHERPA 2.2, SciPost Phys. **7**, 034 (2019).
[5] C. Bierlich *et al.*, A comprehensive guide to the physics and usage of PYTHIA 8.3, SciPost Phys. Codebases 8 (2022).
[6] J. R. Andersen, C. Gütschow, A. Maier, and S. Prestel, A positive resampler for Monte Carlo events with negative weights, Eur. Phys. J. C **80**, 1007 (2020).
[7] B. Nachman and J. Thaler, Neural resampler for Monte Carlo reweighting with preserved uncertainties, Phys. Rev. D **102**, 076004 (2020).
[8] J. R. Andersen and A. Maier, Unbiased elimination of negative weights in Monte Carlo samples, Eur. Phys. J. C **82**, 433 (2022).
[9] R. Frederix, S. Frixione, S. Prestel, and P. Torrielli, On the reduction of negative weights in MC@NLO-type matching procedures, J. High Energy Phys. 07 (2020) 238.
[10] K. Danziger, S. Höche, and F. Siegert, Reducing negative weights in Monte Carlo event generation with SHERPA, arXiv:2110.15211.
[11] T.-J. Hou *et al.*, New CTEQ global analysis of quantum chromodynamics with high-precision data from the LHC, Phys. Rev. D **103**, 014013 (2021).
[12] S. Bailey, T. Cridge, L. A. Harland-Lang, A. D. Martin, and R. S. Thorne, Parton distributions from LHC, HERA, Tevatron and fixed target data: MSHT20 PDFs, Eur. Phys. J. C **81**, 341 (2021).
[13] R. D. Ball *et al.*, The path to proton structure at one-percent accuracy, Eur. Phys. J. C **82**, 428 (2022).
[14] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, Parton distributions for the LHC, Eur. Phys. J. C **63**, 189 (2009).
[15] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne, Parton distributions in the LHC era: MMHT 2014 PDFs, Eur. Phys. J. C **75**, 204 (2015).
[16] H.-L. Lai, J. Huston, S. Mrenna, P. Nadolsky, D. Stump, W.-K. Tung, and C. P. Yuan, Parton distributions for event generators, J. High Energy Phys. 04 (2010) 035.
[17] S. Dulat, T.-J. Hou, J. Gao, M. Guzzi, J. Huston, P. Nadolsky, J. Pumplin, C. Schmidt, D. Stump, and C.-P. Yuan, New parton distribution functions from a global analysis of quantum chromodynamics, Phys. Rev. D **93**, 033006 (2016).
[18] H. Abramowicz *et al.* (H1 and ZEUS Collaborations), Combination of measurements of inclusive deep inelastic $e^{\pm}p$ scattering cross sections and QCD analysis of HERA data, Eur. Phys. J. C **75**, 580 (2015).
[19] A. Aktas *et al.* (H1 Collaboration), Measurement of $F_2^{c\bar{c}}$ and $F_2^{b\bar{b}}$ at high Q^2 using the H1 vertex detector at HERA, Eur. Phys. J. C **40**, 349 (2005).
[20] H. Abramowicz *et al.* (H1 and ZEUS Collaborations), Combination and QCD analysis of charm production cross section measurements in deep-inelastic ep scattering at HERA, Eur. Phys. J. C **73**, 2311 (2013).
[21] F. D. Aaron *et al.* (H1 Collaboration), Measurement of the inclusive $e^{\pm}p$ scattering cross section at high inelasticity y and of the structure function F_L, Eur. Phys. J. C **71**, 1579 (2011).
[22] R. Aaij *et al.* (LHCb Collaboration), Measurement of the forward Z boson production cross-section in pp collisions at $\sqrt{s} = 7$ TeV, J. High Energy Phys. 08 (2015) 039.
[23] R. Aaij *et al.* (LHCb Collaboration), Measurement of forward $\mathrm{Z} \to \mathrm{e}^{+}\mathrm{e}^{-}$ production at $\sqrt{s} = 8$ TeV, J. High Energy Phys. 05 (2015) 109.
[24] V. Khachatryan *et al.* (CMS Collaboration), Measurement of the differential cross section and charge asymmetry for inclusive $\mathrm{pp} \to \mathrm{W}^{\pm} + X$ production at $\sqrt{s} = 8$ TeV, Eur. Phys. J. C **76**, 469 (2016).
[25] R. Aaij *et al.* (LHCb Collaboration), Measurement of forward W and Z boson production in pp collisions at $\sqrt{s} = 8$ TeV, J. High Energy Phys. 01 (2016) 155.
[26] G. Aad *et al.* (ATLAS Collaboration), Measurement of the transverse momentum and ϕ_{η}^{*} distributions of Drell–Yan lepton pairs in proton–proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector, Eur. Phys. J. C **76**, 291 (2016).
[27] A. C. Benvenuti *et al.* (BCDMS Collaboration), A high statistics measurement of the deuteron structure functions $F_2(X, Q^2)$ and R from deep inelastic muon scattering at high Q^2, Phys. Lett. B **237**, 592 (1990).
[28] A. C. Benvenuti *et al.* (BCDMS Collaboration), A high statistics measurement of the proton structure functions $F_2(x, Q^2)$ and R from deep inelastic muon scattering at high Q^2, Phys. Lett. B **223**, 485 (1989).
[29] M. Arneodo *et al.* (New Muon Collaboration), Measurement of the proton and deuteron structure functions, $F_2{}^p$ and $F_2{}^d$, and of the ratio σ_{L} σ_{T}, Nucl. Phys. **B483**, 3 (1997).
[30] J. P. Berge *et al.*, A measurement of differential cross-sections and nucleon structure functions in charged current neutrino interactions on iron, Z. Phys. C **49**, 187 (1991).
[31] U.-K. Yang *et al.* (CCFR/NuTeV Collaboration), Measurements of F_2 and $xF_3^{\nu} - xF_3^{\bar{\nu}}$ from CCFR ν_{μ} − Fe and $\bar{\nu}_{\mu}$ − Fe Data in a Physics Model Independent Way, Phys. Rev. Lett. **86**, 2742 (2001).

[32] W. G. Seligman *et al.*, Improved Determination of α_s from Neutrino Nucleon Scattering, Phys. Rev. Lett. **79**, 1213 (1997).
[33] D. A. Mason, Measurement of the strange-antistrange asymmetry at NLO in QCD from NuTeV dimuon data, Ph.D. thesis, Oregon University, 2006.
[34] M. Goncharov *et al.* (NuTeV Collaboration), Precise measurement of dimuon production cross-sections in ν_μ Fe and $\bar{\nu}_\mu$ Fe deep inelastic scattering at the tevatron, Phys. Rev. D **64**, 112006 (2001).
[35] T. Aaltonen *et al.* (CDF Collaboration), Measurement of the inclusive jet cross section at the Fermilab tevatron $p\bar{p}$ collider using a cone-based jet algorithm, Phys. Rev. D **78**, 052006 (2008).
[36] V. M. Abazov *et al.* (D0 Collaboration), Measurement of the Inclusive Jet Cross-Section in $p\bar{p}$ Collisions at $s^{(1/2)} = 1.96$-TeV, Phys. Rev. Lett. **101**, 062001 (2008).
[37] S. Chatrchyan *et al.* (CMS Collaboration), Measurement of the ratio of inclusive jet cross sections using the anti-k_T algorithm with radius parameters $R = 0.5$ and 0.7 in pp collisions at $\sqrt{s} = 7$ TeV, Phys. Rev. D **90**, 072006 (2014).
[38] G. Aad *et al.* (ATLAS Collaboration), Measurement of the inclusive jet cross-section in proton-proton collisions at $\sqrt{s} = 7$ TeV using 4.5 1/fb of data with the ATLAS detector, J. High Energy Phys. 02 (2015) 153.
[39] V. Khachatryan *et al.* (CMS Collaboration), Measurement and QCD analysis of double-differential inclusive jet cross sections in pp collisions at $\sqrt{s} = 8$ TeV and cross section ratios to 2.76 and 7 TeV, J. High Energy Phys. 03 (2017) 156.
[40] A. M. Sirunyan *et al.* (CMS Collaboration), Measurement of double-differential cross sections for top quark pair production in pp collisions at $\sqrt{s} = 8$ TeV and impact on parton distribution functions, Eur. Phys. J. C **77**, 459 (2017).
[41] G. Aad *et al.* (ATLAS Collaboration), Measurements of top-quark pair differential cross-sections in the lepton + jets channel in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector, Eur. Phys. J. C **76**, 538 (2016).
[42] M. Aaboud *et al.* (ATLAS Collaboration), Precision measurement and interpretation of inclusive W^+, W^- and Z/γ^* production cross sections with the ATLAS detector, Eur. Phys. J. C **77**, 367 (2017).
[43] G. Aad *et al.* (ATLAS Collaboration), Measurement of the inclusive $W^\pm$ and Z/gamma cross sections in the electron and muon decay channels in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector, Phys. Rev. D **85**, 072004 (2012).
[44] C. G. Callan, Jr. and D. J. Gross, High-Energy Electroproduction and the Constitution of the Electric Current, Phys. Rev. Lett. **22**, 156 (1969).
[45] M. A. G. Aivazis, J. C. Collins, F. I. Olness, and W.-K. Tung, Leptoproduction of heavy quarks. 2. A Unified QCD formulation of charged and neutral current processes from fixed target to collider energies, Phys. Rev. D **50**, 3102 (1994).
[46] M. A. G. Aivazis, F. I. Olness, and W.-K. Tung, Leptoproduction of heavy quarks. 1. General formalism and kinematics of charged current and neutral current production processes, Phys. Rev. D **50**, 3085 (1994).
[47] M. Kramer, F. I. Olness, and D. E. Soper, Treatment of heavy quarks in deeply inelastic scattering, Phys. Rev. D **62**, 096007 (2000).
[48] W.-K. Tung, S. Kretzer, and C. Schmidt, Open heavy flavor production in QCD: Conceptual framework and implementation issues, J. Phys. G **28**, 983 (2002).
[49] M. Guzzi, P. M. Nadolsky, H.-L. Lai, and C. P. Yuan, General-mass treatment for deep inelastic scattering at two-loop accuracy, Phys. Rev. D **86**, 053005 (2012).
[50] J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, P. M. Nadolsky, and W. K. Tung, New generation of parton distributions with uncertainties from global QCD analysis, J. High Energy Phys. 07 (2002) 012.
[51] T.-J. Hou, P. Nadolsky, K. Xie, and C.-P. Yuan, CTEQ-TEA Hepforge, https://ct.hepforge.org.
[52] T. Carli, D. Clements, A. Cooper-Sarkar, C. Gwenlan, G. P. Salam, F. Siegert, P. Starovoitov, and M. Sutton, *A posteriori* inclusion of parton density functions in NLO QCD final-state calculations at hadron colliders: The APPLGRID project, Eur. Phys. J. C **66**, 503 (2010).
[53] R. D. Ball *et al.* (NNPDF Collaboration), Parton distributions from high-precision collider data, Eur. Phys. J. C **77**, 663 (2017).
[54] P. A. Zyla *et al.* (Particle Data Group), Review of particle physics, Prog. Theor. Exp. Phys. **2020**, 083C01 (2020).
[55] G. Aad *et al.* (ATLAS Collaboration), Measurement of the t-channel single top-quark production cross section in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector, Phys. Lett. B **717**, 330 (2012).
[56] G. Aad *et al.* (ATLAS Collaboration), Comprehensive measurements of t-channel single top-quark production cross sections at $\sqrt{s} = 7$ TeV with the ATLAS detector, Phys. Rev. D **90**, 112006 (2014).
[57] S. Chatrchyan *et al.* (CMS Collaboration), Measurement of the t-Channel Single Top Quark Production Cross Section in pp Collisions at $\sqrt{s} = 7$ TeV, Phys. Rev. Lett. **107**, 091802 (2011).
[58] S. Chatrchyan *et al.* (CMS Collaboration), Measurement of the single-top-quark t-channel cross section in pp collisions at $\sqrt{s} = 7$ TeV, J. High Energy Phys. 12 (2012) 035.
[59] M. Aaboud *et al.* (ATLAS Collaboration), Fiducial, total and differential cross-section measurements of t-channel single top-quark production in pp collisions at 8 TeV using data collected by the ATLAS detector, Eur. Phys. J. C **77**, 531 (2017).
[60] V. Khachatryan *et al.* (CMS Collaboration), Measurement of the t-channel single-top-quark production cross section and of the $|V_{tb}|$ CKM matrix element in pp collisions at $\sqrt{s} = 8$ TeV, J. High Energy Phys. 06 (2014) 090.
[61] M. Aaboud *et al.* (ATLAS Collaboration), Measurement of the inclusive cross-sections of single top-quark and top-antiquark t-channel production in pp collisions at $\sqrt{s} =$ 13 TeV with the ATLAS detector, J. High Energy Phys. 04 (2017) 086.
[62] A. M. Sirunyan *et al.* (CMS Collaboration), Cross section measurement of t-channel single top quark production in pp collisions at $\sqrt{s} = 13$ TeV, Phys. Lett. B **772**, 752 (2017).
[63] A. M. Sirunyan *et al.* (CMS Collaboration), Measurement of the single top quark and antiquark production cross sections

in the t channel and their ratio in proton-proton collisions at $\sqrt{s} = 13$ TeV, Phys. Lett. B **800**, 135042 (2020).
[64] Z. Sullivan, Are PDFs still consistent with Tevatron data?, EPJ Web Conf. **172**, 03008 (2018).
[65] J. Campbell, T. Neumann, and Z. Sullivan, Testing parton distribution functions with t-channel single-top-quark production, Phys. Rev. D **104**, 094042 (2021).
[66] J. M. Campbell and R. K. Ellis, MCFM for the Tevatron and the LHC, Nucl. Phys. B, Proc. Suppl. **205–206**, 10 (2010).
[67] R. Boughezal, J. M. Campbell, R. K. Ellis, C. Focke, W. Giele, X. Liu, F. Petriello, and C. Williams, Color singlet production at NNLO in MCFM, Eur. Phys. J. C **77**, 7 (2017).
[68] J. M. Campbell, R. K. Ellis, and W. T. Giele, A multi-threaded version of MCFM, Eur. Phys. J. C **75**, 246 (2015).
[69] R. D. Ball *et al.* (NNPDF Collaboration), Parton distributions for the LHC Run II, J. High Energy Phys. 04 (2015) 040.
[70] A. Buckley, J. Ferrando, S. Lloyd, K. Nordström, B. Page, M. Rüfenacht, M. Schönherr, and G. Watt, LHAPDF6: Parton density access in the LHC precision era, Eur. Phys. J. C **75**, 132 (2015).
[71] https://lhapdf.hepforge.org/.
[72] https://ct.hepforge.org/.