Connected and disconnected sea partons from the CT18 parametrization of PDFs

Tie-Jiun Hou,[1,*] Mengshi Yan,[2] Jian Liang,[3] Keh-Fei Liu,[4] and C.-P. Yuan[5]

[1]*School of Nuclear Science and Technology, University of South China, Hengyang, Hunan 421001, China*
[2]*Department of Physics and State Key Laboratory of Nuclear Physics and Technology, Peking University, Beijing 100871, China*
[3]*Guangdong Provincial Key Laboratory of Nuclear Science, Institute of Quantum Matter, South China Normal University, Guangzhou 510006, China and Guangdong-Hong Kong Joint Laboratory of Quantum Matter, Southern Nuclear Science Computing Center, South China Normal University, Guangzhou 510006, China*
[4]*Department of Physics and Astronomy, University of Kentucky, Lexington, Kentucky 40506, USA*
[5]*Department of Physics and Astronomy, Michigan State University, East Lansing, Michigan 48824, USA*

(Received 19 June 2022; accepted 11 October 2022; published 10 November 2022)

The separation of the connected and disconnected sea partons, which were uncovered in the Euclidean path-integral formulation of the hadronic tensor, is accommodated with an extended parametrization of the nonperturbative parton distribution functions in the CT18 global analysis. This is achieved with the help of the distinct small x behaviors of these two sea partons and the constraint from the lattice calculation of the ratio of the strange momentum fraction to that of the $\bar{u}$ or $\bar{d}$ in the disconnected insertion. The whole dataset of CT18 is used in this CT18CS fit. The impact of the recent SeaQuest data on the $\bar{d}(x) - \bar{u}(x)$ distribution of CT18CS is also discussed. The separate momentum fractions for the valence, the connected sea and disconnected sea of u and d, the strange and the gluon partons are presented at the input scale $\mu = 1.3$ GeV for the first time. They can be compared term-by-term with systematic error controlled lattice calculations.

DOI: 10.1103/PhysRevD.106.096008

I. INTRODUCTION

In high energy experiments, such as those at hadron colliders, theoretical analyses depend on the parton structure of the hadronic beams in terms of their parton distribution functions (PDFs) in order to understand the $W^{\pm}, Z$ and Higgs productions in precision measurements of the Standard Model parameters and the search of new physics. The universal PDFs can be extracted from deep inelastic scattering (DIS) and Drell-Yan processes with the help of factorization theorem and global analyses which involve the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) evolution equations. Since the factorization formula involves an integral of the product of the parton distribution functions (PDFs) and the perturbative short distance kernel, extracting PDFs is intrinsically an inverse problem. The common approach is to model the PDFs in terms of the valence and sea partons with respective small and large x behaviors and perform a global fit of the available experimental data at different Q^2 values. As a result, the quality of the fit and its accuracy depend on the precision and availability of the experimental data in the relevant kinematic range. In particular, the flavor structure of the partons can be improved with experiments which directly address the flavor dependence. For example, the first experimental evidence that the sea patrons have nontrivial flavor dependence is shown in the experimental demonstration of the violation of Gottfried sum rule. The original Gottfried sum rule, $I_G \equiv \int_0^1 dx[F_2^p(x) - F_2^n(x)]/x = 1/3$, was obtained under the assumption that $\bar{u}$ and $\bar{d}$ sea partons are the same [1]. However, the New Muon Collaboration (NMC) measurement [2,3] of $\int_0^1 dx[F_2^p(x) - F_2^n(x)]/x$ turns out to be 0.235 ± 0.026, a 4σ difference from the Gottfried sum rule, which implies that the $\bar{u} = \bar{d}$ assumption was invalid. The recent SeaQuest experiment clearly shows that the $\bar{d}/\bar{u}$ ratio in the range $0.1 < x < 0.4$ is substantially larger than unity ($\sim$1.5) [4]. Other flavor-dependent issues under active experimental and theoretical pursuits include the intrinsic strange and charm partons [5–9], and the $s(x) - \bar{s}(x)$ [8–13] and $c(x) - \bar{c}(x)$ [14] differences.

The violation of the Gottfried sum rule prompted the Euclidean path-integral formulation of the hadronic tensor

*tjhou@msu.edu

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP[3].

of the nucleon for DIS which uncovered that there are two kinds of sea partons, one is the connected sea and the other disconnected sea [15,16]. They are so named to reflect the topology of the quark lines in the 4-point current-current correlator in the nucleon. The connected sea (CS) results from a connected insertion of the currents on the same "valence" quark line and the disconnected sea (DS) is from a disconnected insertion involving a vacuum polarization from the quark loop involving the external currents. These are "hand-bag" diagrams. On the other hand, the "cat's ears" diagrams, where two currents in the current-current correlator couple to different quark lines, are higher twists and are suppressed in the DIS region, but they are as important as the leading twists in low-energy lepton-nucleon scattering [17,18]. The suppression of the higher twist contributions at large Q^2 has been demonstrated in a recent lattice calculation which shows that the cat's ears diagrams drops out quickly as compared to those hand-bag diagrams when the three momentum transfer becomes large [18]. It is proved [15] that, in the isospin symmetric limit, the Gottfried sum rule violation originates only from the CS which is subject to Pauli blocking due to the unequal numbers of the valence u and d quarks in both the proton and the neutron. Attempts have been made [19–22] to separate out the CS and DS partons by combining strange parton distribution from a HERMES experiment [23], $\bar{u}+\bar{d}$ from the CT10 analysis [24], and the ratio $\langle x\rangle_s/\langle x\rangle_u$ (disconnected insertion) from lattice calculations [22,25].

In this work, we shall accommodate parton degrees of freedom delineated in the path-integral formulation of the hadronic tensor in the form of CT18 global analysis [26] of unpolarized PDFs. Adopting lattice results as constraints to perform the global fits has been applied to quark transversity distribution [27]. The present work goes one step further to explicitly separate the CS and DS degrees of freedom for the first time under the CT18 parametrization [28].

This manuscript is organized as follows. Section II gives a brief review of the path-integral formulation of the hadronic tensor which defines the parton degrees of freedom. Section III describes the parametrization of the PDF for each of the parton degrees of freedom and the details of the global analysis. The result of a global analysis with the inclusion of both CS and DS partons fitted to the original CT18 datasets, termed as CT18CS fit, is presented in Sec. IV. The second moments of the separate valence, the disconnected sea, and the gluon partons are presented for the first time at the input scale which can finally be compared directly with lattice calculations for each term and each flavor. We note that the E906 SeaQuest [4] data only became available after the completion of the CT18 analysis. Hence, we shall examine in Sec. V the impact of the E906 SeaQuest [4] data on a global fit similar to CT18CS, in which the E866 NuSea [29] data was already included. Section VI contains our summary.

II. PARTON DEGREES OF FREEDOM FROM EUCLIDEAN PATH-INTEGRAL FORMULATION OF THE HADRONIC TENSOR

The Euclidean hadronic tensor was formulated in the path-integral formalism to identify the origin of the Gottfried sum rule violation [15,16]. It is defined as the current-current correlator in the nucleon with Fourier transform in the spatial directions

$$\tilde{W}_{\mu\nu}(\vec{q},\vec{p},\tau)=\langle N(\vec{p})|\int d^3x\frac{e^{i\vec{q}\cdot\vec{x}}}{4\pi}J_\mu(\vec{x},\tau)J_\nu(0,0)|N(\vec{p})\rangle, \tag{1}$$

It is a function of τ, which is the Euclidean time separation between the currents. Formally, the inverse Laplace transform converts $\tilde{W}_{\mu\nu}(\vec{q},\vec{p},\tau)$ to the Minkowski hadronic tensor

$$W_{\mu\nu}(\vec{q},\vec{p},\nu)=\frac{1}{2m_N i}\int_{c-i\infty}^{c+i\infty}d\tau e^{\nu\tau}\tilde{W}_{\mu\nu}(\vec{q},\vec{p},\tau), \tag{2}$$

with $c>0$. However, this is not practical in lattice calculation, as there are no data on the imaginary τ. Instead, one can turn this into an inverse problem and find a solution from the Laplace transform [30]

$$\tilde{W}_{\mu\nu}(\vec{q},\vec{p},\tau)=\int d\nu e^{-\nu\tau}W_{\mu\nu}(\vec{q},\vec{p},\nu). \tag{3}$$

This has been studied [17,18,30,31] with the inverse algorithms such as Backus-Gilbert, maximum entropy and Bayesian reconstruction methods. The spectral density in lepton-nucleon scattering has several kinematic regions as the energy transfer ν increases—the elastic scattering, the inelastic reactions ($\pi N,\pi\pi M,\eta N$ etc.) and resonances (Δ, Roper, S_{11}, etc.), shallow inelastic scattering (SIS), and deep inelastic scattering (DIS) regions. To determine how large a ν is needed for DIS, we look at W, the total invariant mass of the hadronic state for the nucleon target at rest

$$W^2=(q+p)^2=m_p^2-Q^2+2m_p\nu. \tag{4}$$

The global analyses of the high energy lepton-nucleon and Drell-Yan experiments to extract the parton distribution functions usually make a cut with $W^2>12$ GeV^2 to avoid the elastic and inelastic regions. Thus, to be qualified in the DIS region, the energy transfer ν needs to be greater than 8 GeV for $Q^2=4$ GeV^2, a typical choice made in the CTEQ-TEA PDF global analysis.

It is shown [15,16,32,33] that, when the time ordering $t_f>t_2>t_1>t_0$ is fixed, the 4-point function for extracting the matrix element $\tilde{W}_{\mu\nu}(\vec{q},\vec{p},\tau)$ in Eq. (1) can be grouped in terms of 6 topologically distinct and gauge invariant path-integral insertions, according to different

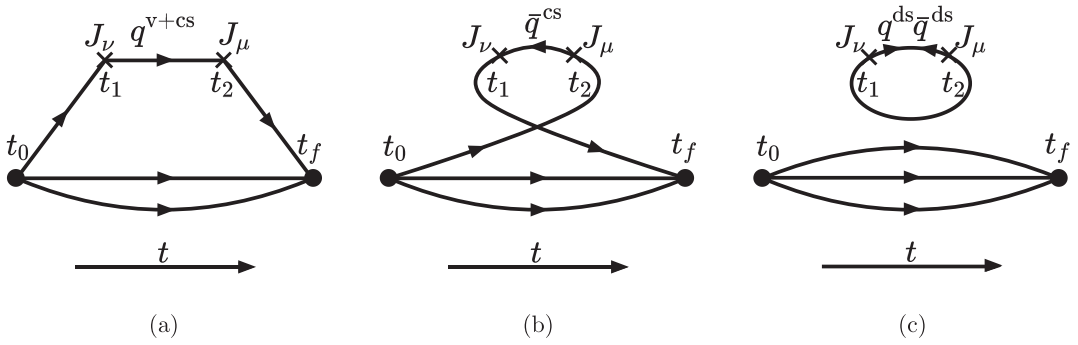

FIG. 1. Three gauge invariant and topologically distinct insertions in the Euclidean-path integral formulation of the nucleon hadronic tensor where the currents couple to the same quark propagator. In the DIS region, the parton degrees of freedom are (a) the valence and connected sea (CS) quarks q^{v+cs}, (b) the CS antiquarks $\bar{q}^{cs}$. Only u and d are present in (a) and (b) for the nucleon hadronic tensor. (c) the disconnected sea (DS) quarks q^{ds} and antiquarks $\bar{q}^{ds}$ with $q = u$, d, s, and c.

Wick contractions among the Grassmann numbers in the two currents and the source/sink interpolation fields. They can be further grouped into two classes. The first class includes those where the currents are coupled to the same quark propagator as illustrated in Fig. 1. The second class involves those where the two currents are coupled to different quark propagators as illustrated in Fig. 2. In low energy lepton-nucleon scattering, all 6 diagrams contribute and they are not separable [33]. However, in the DIS region, the first class are hand-bag diagrams which include the leading twist contributions, the second class are cat's ears diagrams, which are higher twists and are suppressed by $\mathcal{O}(1/Q^2)$.

The first class in Fig. 1 includes three path-integral diagrams that can be denoted as connected insertions (CI) [Figs. 1(a) and 1(b)], where the quark lines are all connected, and disconnected insertions (DI) [Fig. 1(c)], where there are vacuum polarizations associated with the currents in disconnected quark loops. We should note that Fig. 1(b) includes the exchange contribution to prevent the u or d quark in the loop in Fig. 1(c) from occupying the same Dirac eigenstate in the nucleon propagator, enforcing the Pauli principle. In fact, Figs. 1(c) and 1(b) are analogous to the direct and exchange diagrams in time-ordered Bethe-Goldstone diagrams in the many-body theory [34].

As far as the leading-twist DIS structure functions F_1, F_2, and F_3 are concerned, the three diagrams in Fig. 1 are additive with contributions classified as the valence and sea quarks q^{v+cs} in Fig. 1(a), the connected sea (CS) antiquarks $\bar{q}^{cs}$ in Fig. 1(b), and disconnected sea (DS) quarks q^{ds} and antiquarks $\bar{q}^{ds}$ in Fig. 1(c) [15,16,32,33]. Since the u and d partons in the quark loop in Fig. 1(c) appear in a different flavor trace than the one involving the nucleon propagator, they cancel in the Gottfried sum in the isospin symmetric limit. Thus, the Gottfried sum rule violation comes entirely from the connected sea (CS) difference $\bar{u}^{cs} - \bar{d}^{cs}$ in the F_2 structure functions in this case [15]. It is proven [16,33] from short distance expansion that the parton degrees of freedom defined in diagram in Fig. 1 are separable, unlike the case of low-energy lepton-nucleon scattering, where the higher twists are important [17,18]. Furthermore, these parton degrees of freedom are identical to those defined from the recent Feynman-x approaches [33], i.e., quasi-PDF [35], pseudo-PDF [36], and lattice cross section [37].

PDFs can be extracted from the factorization formula [38] where the experimental cross section or structure functions are expressed as a convolution integral of the coefficient functions and the PDFs. In practice, the global fitting programs adopt the parton degree of freedoms as $u, d, \bar{u}, \bar{d}, s, \bar{s}$ and g. We see that from the path-integral formalism of QCD, each of the u and d have two sources,

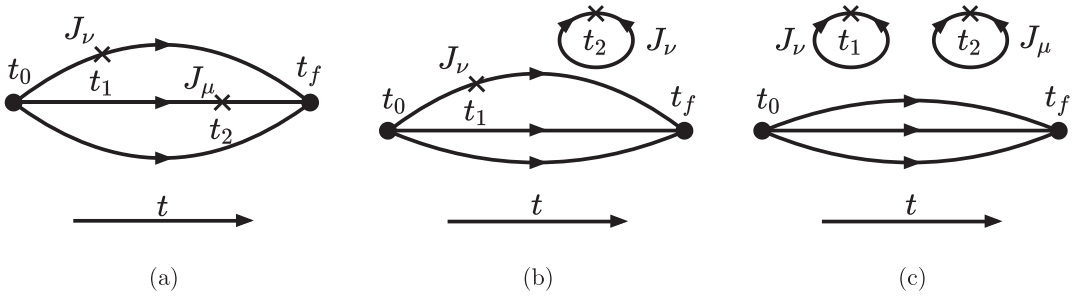

FIG. 2. Three other gauge invariant and topologically distinct insertions where the currents are inserted on different quark propagators. In the DIS region, they are higher-twist diagrams.

one from the connected insertion (CI) [Fig. 1(a)] and one from the disconnected insertion (DI) [Fig. 1(c)], so are $\bar{u}$ and $\bar{d}$ from Figs. 1(b) and 1(c). On the other hand, s, c and $\bar{s}, \bar{c}$ only come from the DI [Fig. 1(c)]. In other words,

$$\begin{aligned} u &= u^{v+cs} + u^{ds}, & d &= d^{v+cs} + d^{ds}, \\ \bar{u} &= \bar{u}^{cs} + \bar{u}^{ds}, & \bar{d} &= \bar{d}^{cs} + \bar{d}^{ds}, \\ s &= s^{ds}, & \bar{s} &= \bar{s}^{ds}, \\ c &= c^{ds}, & \bar{c} &= \bar{c}^{ds}. \end{aligned} \tag{5}$$

This classification of the parton degrees of freedom is richer than those in terms of q and $\bar{q}$ in the global analysis due to the fact that there are two sources for the quarks—q^{v+cs} and q^{ds}—and two sources for the antiquarks—$\bar{q}^{cs}$ and $\bar{q}^{ds}$. The distinguishing feature of CS and DS lies in their characteristic small-x behaviors, which we shall explore in this work to perform global analysis. In the Regge theory, the small-x behavior of q^{v+cs} and $\bar{q}^{cs}$, being in the flavor nonsinglet connected insertions, are dominated by the Reggeon exchange. Thus, we expect $q^{v+cs}(x), \bar{q}^{cs}(x) \to_{x\to 0} x^{-\alpha}$ for $q = u$, d, where $\alpha \sim 0.5$ is the slope of the Regge trajectory. Whereas, DS is flavor singlet and can have Pomeron exchanges. Hence, $q^{ds}(x), \bar{q}^{ds} \to_{x\to 0} x^{-1}$ for $q = u$, d, s, c. In an attempt to separate the CS and DS quarks [19] by combining strange quark distribution from a HERMES experiment and $\bar{u} + \bar{d}$ from CT10, it is found that $x(\bar{u}^{cs} - \bar{d}^{cs})$ spans the same x range as that of $x(\bar{u} - \bar{d})$, which suggests that they have similar small-x behaviors; whereas $x(\bar{u}^{ds} + \bar{d}^{ds})$ is much singular for $x < 0.05$ [19]. This is consistent with expectation.

Until the Feynman-x and/or the hadronic tensor approaches on the lattice have all the systematic errors, such as excited states contamination and large nucleon momentum, are under control so that all region of x can be compared with those from the global analyses, the most reliable comparison between global fittings and lattice calculations are via the parton moments. The latter are getting mature with all the systematic errors (e.g., continuum and infinite volume extrapolations, excited states contamination, physical pion mass, nonperturbative renormalization, and scale setting) have been taken into account [39,40]. However, as pointed out in [33,41], it is not possible to compare the moments from global analyses and those from the lattice calculations in detail, except for the limited isovector $(u - d)$ and stangeness moments. This is because the lattice calculation of moments in the three-point functions are organized in the connected insertions (CI) and disconnected insertions (DI). The CI includes both q^{v+cs} and $\bar{q}^{cs}$, while DI includes q^{ds} and $\bar{q}^{ds}$. On the other hand, in the present global analyses, CS and DS degrees of freedom are not separated. To make a comparison at the moment level, it is encumbered upon global analyses to disentangle the connected sea from the disconnected, so that the full lattice results of moments in CI and DI can be compared to them for each flavor.

III. GLOBAL FITTING

In this section, the general setting of the CT18CS global fit is presented. The CT18CS, as an extended parametrization of PDFs in accommodation with the Euclidean path-integral formalism of QCD, requires a different scheme of parton classification with more parton degrees of freedom. The specific parton degrees of freedom to be parametrized and a number of ansatzes imposed in this global analysis will be explained in the following Sec. III A. In the Sec. III B, we will introduce the settings of small-x and large-x behavior for CS and DS parton distributions.

A. Parton degrees of freedom

In the QCD global analysis of parton distributions in the proton, the PDFs of various partons are parametrized in some functional forms at the initial scale Q_0 (about 1 GeV), from where the PDFs are evolved to any arbitrary higher energy scale Q via DGLAP evolution equations. Typically, it is assumed that the charm and bottom quark PDFs are generated perturbatively from QCD evolution, though in some special studies, the possibility of having nonperturbaive charm PDF at the Q_0 scale was also considered, such as in Refs. [6–9,42]. Therefore, in general, the total number of parton degrees of freedom at the Q_0 scale is 7, which includes the following partons:

$$g, u^v, d^v, \bar{u}, \bar{d}, s, \bar{s} \tag{6}$$

In CT18 [26], it is also assumed that the strange PDFs $s = \bar{s}$ at the Q_0 scale, though $s \neq \bar{s}$ will be generated at large Q scale via NNLO QCD evolution. Given this ansatz, the number of parton degrees of freedom is 6 in CT18. We note that in MSHT20 PDFs [13] and NNPDF4.0 [9], a nonvanishing asymmetric strangeness $s(x) \neq \bar{s}(x)$ is imposed in the nonperturbative parametrization at their respective Q_0 scales.

As mentioned in the last section, when the separation of CS partons and DS partons are considered, we would have more partonic degrees of freedom. The classification of partons becomes:

$$g, u^{v+cs}, u^{ds}, \bar{u}^{cs}, \bar{u}^{ds}, d^{v+cs}, d^{ds}, \bar{d}^{cs}, \bar{d}^{ds}, s^{ds}, \bar{s}^{ds}, \tag{7}$$

totally 11 of them. To implement all the degrees of freedom in Eq. (7) and obtain their Q^2 dependence would require generalized DGLAP evolution equations as developed in Ref. [41]. In the present study, we shall pararmetrize the extended set of partons in Eq. (7) at the input scale Q_0. with some specific assumptions to be listed below, and then combine the CS and DS into the conventional partons in

Eqs. (5) and (6) so that we can use the same NNLO evolution equations as for CT18. In this way, we can compare with the results of CT18 to discern the different roles played by CS and DS and their respective impacts on physics at this stage. When the generalized evolution code is ready, we can fully explore the CS and DS effects at all scales.

For the present work, we have adopted the following assumptions to reduce the number of parton degrees of freedom from 11 to 6, similar to that in CT18.

(i) Similar to CT18, we assume the symmetric disconnected sea parton distributions:

$$u^{ds} = \bar{u}^{ds}, d^{ds} = \bar{d}^{ds}, \quad \text{and} \quad s^{ds} = \bar{s}^{ds}. \tag{8}$$

(ii) The isospin symmetry is imposed for the u and d quarks:

$$u^{ds} = \bar{u}^{ds} = d^{ds} = \bar{d}^{ds}. \tag{9}$$

(iii) The DS components of u and d quark PDFs are proportional to the s quark PDF, i.e.,

$$u^{ds} = \bar{u}^{ds} = d^{ds} = \bar{d}^{ds} = R \times s. \tag{10}$$

Since the DS in the lattice calculation involves a correlation between the quark loop and the nucleon propagator via the gluons, it is not as sensitive to the nucleon wave function as are the valence and CS partons in the connected insertion. The only difference between u^{ds}, d^{ds}, and s is their quark masses. Thus, it is reasonable to postulate that their distribution are the same modulo a proportional constant R. In this work, we determine the value of R from the ratio of the second moment between the strange and the sum of u and d in the disconnected insertion, predicted by a lattice QCD calculation which has taken all the systematic errors into account [22]. It yields $1/R = \langle x \rangle_{s+\bar{s}}/\langle x \rangle_{\bar{u}+\bar{d}}(\mathrm{DI}) = 0.822(69)(78)$ at 1.3 GeV, where $\langle x \rangle_{\bar{u}+\bar{d}}(\mathrm{DI})$ is the momentum fraction carried by the light quark (either u or d) in the disconnect insertions. This result was obtained by properly evolving the matching coefficients from 2 GeV to 1.3 GeV [43], using the known result of $1/R$ at 2 GeV, which was found to be 0.795(79)(77) [22].

In the CTEQ-TEA PDF global analysis, the normalizations for individual sea quark PDFs are computed using the valence quark and momentum sum rules, and the first moments $\langle x \rangle_g$ and the ratio $\langle x \rangle_{\bar{s}+\bar{s}}/\langle x \rangle_{\bar{u}+\bar{d}}$ fitted as free parameters. Since the parametrizations do not determine the ratio of the strange-to-nonstrange PDFs, we restrict this ratio in the present work by the above-mentioned prediction from lattice-QCD. Specifically, we require that the ratio $(s^{ds}(x) + \bar{s}^{ds}(x))/(\bar{u}^{ds}(x) + \bar{d}^{ds}(x))$, at $Q_0 = 1.3$ GeV, is constrained at the 68% confident level to be in the interval [0.718, 0.926] with a central value of 0.822, by imposing the appropriate Lagrange multiplier constraint in the CT18CS fit.

Finally, we note that the assumption in Eq. (10) can be checked by the similar $\langle x^3 \rangle$ ratio in lattice calculations in the future.

(iv) We further define

$$u^{cs} \equiv \bar{u}^{cs} \quad \text{and} \quad d^{cs} \equiv \bar{d}^{cs}, \tag{11}$$

so that

$$u^{v} = u^{v+cs} - u^{cs} = u^{v+cs} - \bar{u}^{cs}, \tag{12}$$

which agrees with the usual definition of the *valence* quark: $q^v \equiv q - \bar{q} = [u^{v+cs} + u^{ds}] - [\bar{u}^{cs} + \bar{u}^{ds}]$, when $u^{ds} = \bar{u}^{ds}$. It was pointed out in [41] that when u^{ds} and $\bar{u}^{ds}$ are not equal, the $q^v \equiv q - \bar{q}$ definition leads to conceptual puzzles, such as the valence u can evolve into valence d in NNLO evolution and that the strangeness can have valence distribution when $s \neq \bar{s}$. These puzzles are resolved with the definition in Eq. (12) [41].

With all the above conditions taken into account, the remaining parton degrees of freedom are $g, u^v, u^{cs}, d^v, d^{cs}, s^{ds}$. As discussed earlier below Eq. (7), we shall combine CS and DS into the usual $\bar{u}/\bar{d}$ d.o.f., i.e., $\bar{u} = \bar{u}^{cs} + \bar{u}^{ds} = \bar{u}^{cs} + Rs$ and $\bar{d} = \bar{d}^{cs} + \bar{d}^{ds} = \bar{d}^{cs} + Rs$ at the input scale and evolve them in the same NNLO equations as CT18 in the global fitting.

B. Small-x and large-x behavior

At the starting Q_0 scale, the nonperturbative PDFs are parametrized as

$$f(x) = x^{a_1 - 1}(1 - x)^{a_2 - 1}\mathrm{Poly(x)}, \tag{13}$$

where the parameters a_1 and a_2 dominate the behavior of PDFs as x approaches 0 or 1, respectively, and the $\mathrm{Poly}(x)$, constructed with a set of Bernstein polynomials in CTEQ-TEA PDF family, is responsible for the shape of PDFs in a wide range of x. In practice, we implemented the following ansatzes to parametrize various parton distributions at the Q_0 scale.

(i) $\bar{d}/\bar{u} \xrightarrow{x \to 0} 1$ Based on the isospin symmetry in strong interaction, we require $\bar{u}$ and $\bar{d}$ to have the same small-x behavior, where the disconnected sea dominates. Specifically, this ansatz is implemented by setting $a_1^{\bar{u}} = a_1^{\bar{d}}$ to preserve the isospin symmetry in the small-x region. This ansatz was also applied in the CT18 fit. See, Appendix C of Ref. [26].

(ii) $u^{ds}, \bar{u}^{ds}, d^{ds}, \bar{d}^{ds}, s^{ds}, \bar{s}^{ds} \xrightarrow{x \to 0} x^{-1}$. Since the DS partons are flavor singlet and can have Pomeron

TABLE I. The fitted values of a_1 and a_2 parameters of various partons in the CT18 and CT18CS NNLO fits. The numbers marked with "⋆" indicate that they are not fitted, but input values in the CT18CS fit. Furthermore, the ansatz made in the CT18CS fit includes $u^{cs} = \bar{u}^{cs}$ and $d^{cs} = \bar{d}^{cs}$, cf. Eq. (11), and $u^{ds} = \bar{u}^{ds} = d^{ds} = \bar{d}^{ds} = R \times s$, with $1/R = \langle x \rangle_{s+\bar{s}} / \langle x \rangle_{\bar{u}+\bar{d}}$ (DI) $= 0.822(69)(78)$ at 1.3 GeV, cf. Eq. (10).

CT18	u^v	d^v	g	$\bar{u}$	$\bar{d}$	s
a_1	0.763	0.763	0.531	−0.022	−0.022	−0.022
a_2	3.036	3.036	3.148	7.737	7.737	10.31
CT18CS	u^v	d^v	g	$\bar{u}^{cs}$	$\bar{d}^{cs}$	s^{ds}
a_1	0.739	0.739	0.553	0.739	0.739	0.000⋆
a_2	3.036⋆	3.036⋆	3.371	7.737⋆	7.737⋆	11.57

exchanges, their small-x behavior goes like x^{-1}. Based on Eq. (10), this ansatz is implemented by setting $a_1^s = 0$, which is the value of the shape parameter a_1 of the strangeness PDF. We note that $a_1^s = 0$ is consistent with the CT18 error PDF sets, with the value of a_1^s of the CT18 central set shown in the first row of Table I.

(iii) $u^{cs}, d^{cs} \xrightarrow{x\to 0} u^v, d^v$. Like valence partons, the CS partons are in the connected insertion, which is flavor nonsinglet. Thus, we set the small-x behavior of CS partons to be the same as those of valence-quark partons: $a_1^{u^{cs}} = a_1^{d^{cs}} = a_1^{u^v} = a_1^{d^v}$.

(iv) $d/u \xrightarrow{x\to 1} d/u$ of CT18. In the CT18 fit, the ratio d/u was required to approach a finite number as $x \to 1$. This assumption is also kept in the CT18CS fit, which is done by setting $a_2^{u^v} = a_2^{d^v} = a_{2,\mathrm{CT18}}^{u^v} = a_{2,\mathrm{CT18}}^{d^v} = 3.036$. Since the ratio d/u at $x \to 1$ is dominated by valence partons and the parameter a_2 controls the PDF behavior as $x \to 1$, we fix the a_2 values of valence partons as those in CT18 fit, for simplicity.

(v) $\bar{d}/\bar{u} \xrightarrow{x\to 1} \bar{d}/\bar{u}$ of CT18. As shown in Refs. [4,26], CT18 PDFs can describe reasonably well both the E866 NuSea [29] and E906 SeaQuest [4] data, though the SeaQuest data were not included in the CT18 fit, as they only became available after the completion of the CT18 fit. Since both datasets provide important constraints on the ratio $\bar{d}/\bar{u}$ as $x \to 1$, and the CS component of anti-quarks dominates the sea parton behavior in the large x region, we set $a_2^{\bar{u}^{cs}} = a_2^{\bar{d}^{cs}} = a_{2,\mathrm{CT18}}^{\bar{u}} = a_{2,\mathrm{CT18}}^{\bar{d}} = 7.737$ in the CT18CS fit, for simplicity.

For a quick comparison, we list in Table I the fitted values of a_1 and a_2 parameters of various partons in the CT18 and CT18CS NNLO fits. The numbers marked with "⋆" indicate that they are not fitted, but input values in the CT18CS fit. Note that we did not list the values of the other shape parameters used in these fits. In total, there are 28 such shape parameters to be fitted in both the CT18 and CT18CS fits, with the same number (6) of parton degrees of freedom. We note that the published CT18 PDF error sets include an additional pair of eigenvector sets to account for the larger error of gluon PDF in the small-x region.

IV. RESULTS

In this section, we present the results of the CT18CS global fit on aspects of quality of the fit, the configuration of PDFs, and various PDF Mellin moments. The comparison between CT18CS and the standard CT18 NNLO fits shows that CT18CS, with an extended parametrization, is consistent with the CT18 global analysis. Note that this global analysis uses the same datasets as the ones used in the CT18 analysis. There are in total 39 datasets, with 3681 data points included [26].

A. Quality of the fit

In Table II, we compare the quality of the CT18CS fit to that of the CT18 fit. It turns out, both have the total

TABLE II. The χ^2 of selected datasets included in the CT18 and CT18CS fits, with non-negligible $\Delta\chi^2 = |\chi^2_{\mathrm{CT18}} - \chi^2_{\mathrm{CT18CS}}|$. $N_{pt,E}$ is the number of data points of individual dataset, and χ^2_{CT18} and χ^2_{CT18CS} are the χ^2 values obtained by using the central set of CT18 and CT18CS PDFs, respectively.

ID	Experimental dataset	$N_{pt,E}$	χ^2_{CT18}	χ^2_{CT18CS}
104	NMC F_2^d/F_2^p [44]	123	125.7	120.6
124	NuTeV $\nu\mu\mu$ SIDIS [45]	38	18.49	19.75
125	NuTeV $\bar{\nu}\mu\mu$ SIDIS [45]	33	38.45	40.05
201	E605 Drell-Yan process [46]	119	103.4	107.1
203	E866 Drell-Yan process $\sigma_{pd}/(2\sigma_{pp})$ [29]	15	16.09	13.50
204	E866 Drell-Yan process $Q^3 d^2\sigma_{pp}/(dQdx_F)$ [47]	184	244.4	240.3
246	LHCb 8 TeV 2.0 fb^{-1} $Z \to e^-e^+$ forward rapidity cross section [48]	17	25.82	23.63
249	CMS 8 TeV W cross section and A_{ch}. [49]	11	11.37	8.089
	All other datasets	3141	3708	3719
	Total	3681	4292	4292

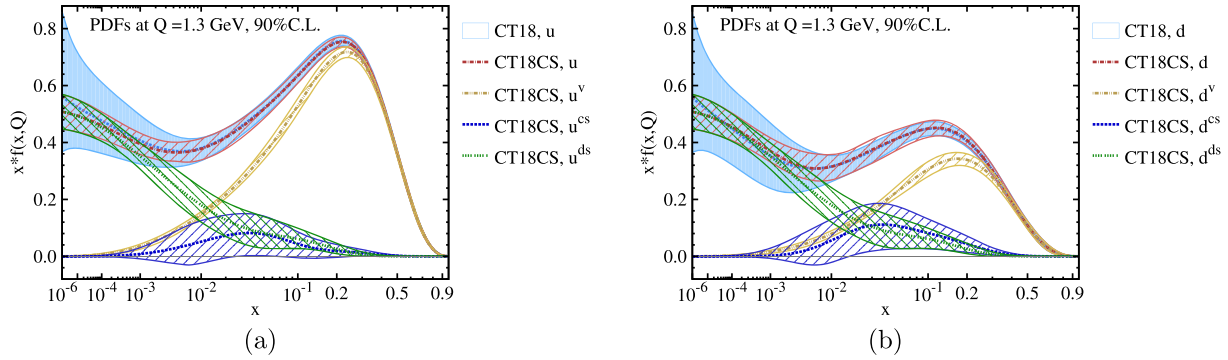

FIG. 3. The decomposition of CT18CS u and d quark distributions in the CS and DS classification at $Q = 1.3$ GeV. The CT18CS PDFs are compared to the nominal CT18 NNLO, which is presented in blue dot line (for central set prediction) and blue error band (for PDF uncertainty).

$\chi^2_{\mathrm{CT18}} = 4292$ for a total of 3681 data points. The experimental datasets which made non-negligible contributions to the change in χ^2 of these two fits are also listed in Table II for comparison. As expected, they are the data most sensitive to valence and sea quark PDFs.

B. Comparison of PDFs

In this section, we compare the fitted CT18CS PDFs obtained in this analysis to the published CT18 PDFs [26].

In CT18CS, the u and d quark distributions are represented by the combination of valence, connected sea, and disconnected sea quark distributions. For $\bar{u}$ and $\bar{d}$ distributions, they are also made of connected sea, and disconnected sea distributions. In Figs. 3 and 4, the decomposition of u, d, $\bar{u}$, $\bar{d}$ in terms of CS and DS parton distributions, at $Q_0 = 1.3$ GeV, are shown, respectively. The PDF error bands, obtained at the 90% confidence level (C.L.), are also shown for comparison.

As shown in Figs. 3(a) and 3(b), the summation of valence, connected sea, and disconnected sea quark distributions of u and d agrees well with the CT18 central PDF values. The CS components of u and d provide sizable contributions in the intermediate-x region only, i.e., $10^{-3} < x < 0.4$. The u and d PDFs in the large-x region are dominated by u^v and d^v. At small-x, the DS components are, as expected, dominating u and d PDFs, where both the valence and CS components are suppressed. Similar comparisons made for $\bar{u}$ and $\bar{d}$ are displayed in Fig. 4. As shown, CT18CS is in good agreement with CT18 NNLO for these parton distributions. Furthermore, in CT18CS, the novel CS parton distribution is found to be responsible for u and d sea quark distributions in the intermediate-x region. On the contrary, the DS patron distribution plays a more important role in the small-x region. We should note that the errors of $u, d, \bar{u}$ and $\bar{d}$ at small x (*i.e.*, $x < 10^{-3}$) from CT18CS are substantially smaller than those in CT18. This is mainly due to the ansatz that we imposed on their small x behavior to be x^{-1} in Sec. III B.

A useful format to compare $\bar{d}$ and $\bar{u}$ PDFs resulted from the CT18CS and CT18 fits is to examine their difference, as shown in Fig. 5. Since we have assumed in this analysis that the DS component of $\bar{d}$ and $\bar{u}$ are the same, hence $(\bar{d} - \bar{u}) = (\bar{d}^{cs} - \bar{u}^{cs})$ in CT18CS PDFs. It shows that the CT18CS central value is close to that of CT18 NNLO for $x > 0.03$. In the small-x region, the difference $(\bar{d}^{cs} - \bar{u}^{cs})$ of CT18CS vanishes. This result is consistent with the prediction presented in Fig. 4 of Ref. [19] in which the E866

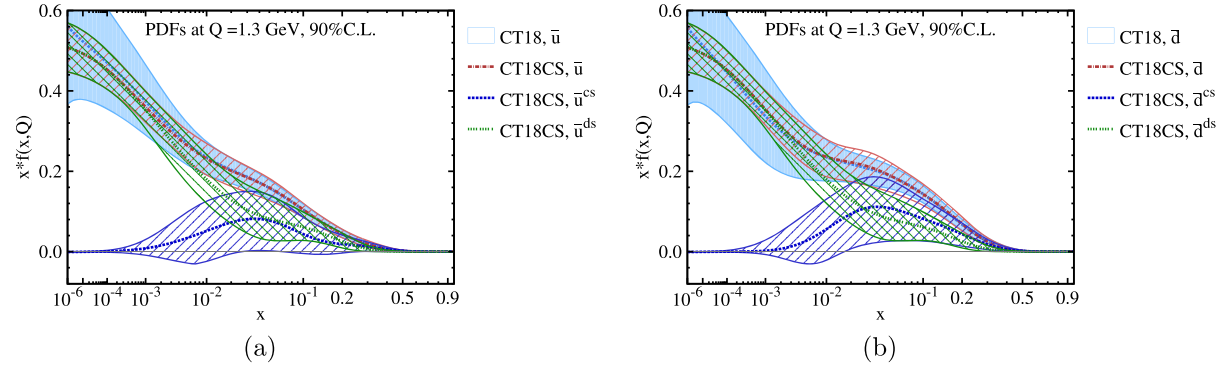

FIG. 4. Similar to Fig. 3, but for the decomposition of CT18CS $\bar{u}$ and $\bar{d}$.

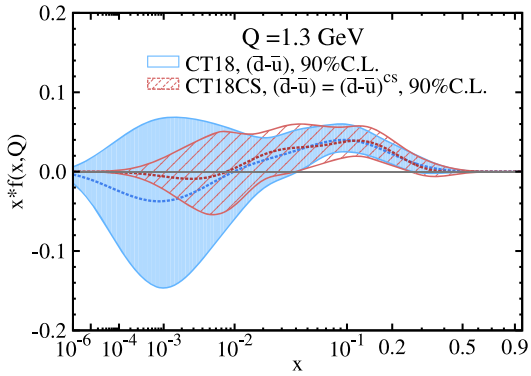

FIG. 5. Comparison of CT18CS and CT18, for $\bar{d}-\bar{u}$. In CT18CS, $(\bar{d}-\bar{u})=(\bar{d}-\bar{u})^{cs}$ due to the ansatz $\bar{d}^{ds}=\bar{u}^{ds}$, cf. Sec. III A.

NuSea [29] and HERMES [23] data were compared to CT10 PDFs [24] in the framework of leading order analysis.

We also show in Fig. 6 similar comparison for g-PDF, s-PDF, $d-u$, and PDF ratio $(s+\bar{s})/(\bar{u}+\bar{d})$ at $Q=1.3$ GeV. In Fig. 6(a), gluon distributions in CT18CS and CT18 fits are in very good agreement across the whole x range. The quark CS and DS separation has no effect on the gluon distribution. As shown in Fig. 6(b), the uncertainty in the strangeness distribution in CT18CS is reduced by a large margin, as compared to CT18 for $x<0.03$. This is of the same pattern as for u, d and $\bar{u}$, $\bar{d}$ since we have adopted the same ansatz, i.e., $a_1^s=0$, in the CT18CS fit, cf. Sec. III B. The central value of the s-PDF distribution of CT18CS for small x is correlated with those of $\bar{u}$ and $\bar{d}$ via the R ratio introduced from the lattice result.

The comparison of $d-u$ distribution between CT18CS and CT18 is shown in Fig. 6(c). The $d-u$ distribution corresponds to the $d^{v+cs}-u^{v+cs}$ distribution in CT18CS. This is because u^{ds} and d^{ds} are assumed to be the same under isospin symmetry, cf. Eq. (9). For $x>0.005$, both central values and sizes of the uncertainty bands of the two PDFs are in good agreement. In the low-x region, the ansatz that CS and valence partons have the same behavior for $x\to 0$ in CT18CS leads to a significant reduction in the uncertainty size.

In Fig. 6(d), the ratio of $(s+\bar{s})/(\bar{u}+\bar{d})$ in CT18CS is compared to that in CT18. In the small-x region, where the DS parton dominates the sea quark distribution, this ratio for CT18CS is constrained by the lattice input, in addition to the ansatz $a_1^s=0$, which reflects its central value and

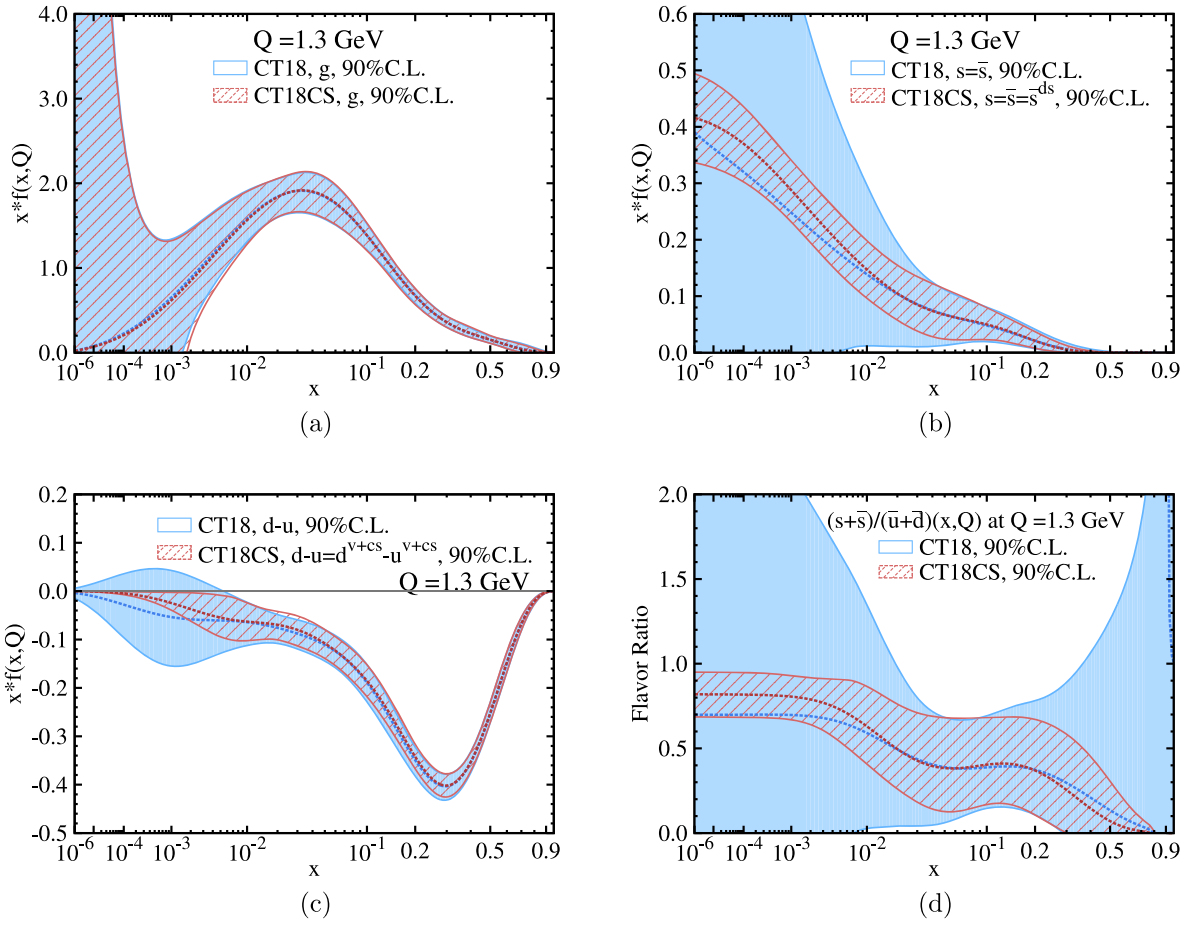

FIG. 6. Similar to Fig. 5, but for comparing g-PDF, s-PDF, $d-u$ and PDF ratio $(s+\bar{s})/(\bar{u}+\bar{d})$ at $Q=1.3$ GeV.

TABLE III. The second moment $\langle x\rangle$ of CT18CS and CT18 NNLO at 1.3 GeV. The superscript "$*$" indicates that due to the fourth ansatz imposed in Eq. (11), the second moments for CS components between quarks and antiquarks are identical, namely, $\langle x\rangle_{\bar{u}^{cs}} = \langle x\rangle_{u^{cs}}$ and $\langle x\rangle_{\bar{d}^{cs}} = \langle x\rangle_{d^{cs}}$. The superscript "†" indicates that due to the second ansatz imposed in Eq. (9), the second moments of DS components of u, $\bar{u}$, d, and $\bar{d}$ are identical.

PDF	$\langle x\rangle_{u^v}$	$\langle x\rangle_{d^v}$	$\langle x\rangle_g$	$\langle x\rangle_{\bar{u}}$	$\langle x\rangle_{\bar{d}}$	$\langle x\rangle_s$
CT18	0.325(5)	0.134(4)	0.385(10)	0.0284(22)	0.0361(27)	0.0134(52)
CT18CS	0.323(4)	0.136(3)	0.384(12)	0.0287(25)	0.0364(34)	0.0137(39)
PDF	$\langle x\rangle_{u^{v+cs}}$	$\langle x\rangle_{d^{v+cs}}$	$\langle x\rangle^*_{\bar{u}^{cs}}$	$\langle x\rangle^*_{\bar{d}^{cs}}$	$\langle x\rangle^\dagger_{u^{ds}}$	
CT18CS	0.335(7)	0.155(8)	0.0120(64)	0.0197(70)	0.0167(49)	

small uncertainty. As a consequence, the error in CT18CS is greatly reduced as compared to that of CT18 for $x < 0.03$. In the larger-x region, where the CS parton becomes important, this ratio is constrained by imposing the same large-x behaviors for $\bar{u}$ and $\bar{d}$ as in CT18. It is noted [22] that the PDF ratio $(s+\bar{s})/(\bar{u}+\bar{d})$ starts to dip for $x > 0.01$. This is due to the fact that $\bar{u}$ or $\bar{d}$ has two components—CS and DS, in contrast to $\bar{s}$ which only has DS. As shown in Fig. 4(b), when $x > 0.01$ the CS components start to show up and contribute to the denominator of the ratio, making it smaller.

C. PDF Mellin moments

The momentum carried by a certain flavor parton can be calculated in terms of the second moment $\langle x\rangle$ of its PDF. In Table III, we compare the predictions of CT18CS to CT18 PDFs on the second moments of various partons at the input scale. The $\bar{u}$ and $\bar{d}$ are split into CS and DS in CT18CS and $\langle x\rangle_{u^{v+cs}}$ is from the direct insertion calculation of the u quark on the lattice, which corresponds to the sum of the valence and CS, cf. Eq. (12). Other similar comparisons can be found in Table VII of Ref. [26]. Without the CS and DS separation, one is not able to compare separate flavor-dependent PDF moments to those from the lattice calculation [41], since the disconnected insertion lattice calculation corresponds to the DS, while the CS is lumped with the valence in the connected insertion. The only exceptions are the strange moments which only have DS and those of $u-d$ which only involve the connected insertion. They are quite limited. One cannot compare the moments for $u, d, \bar{u}$ and $\bar{d}$.

Now that the CS and DS are separated (although at the input scale) in CT18CS, the lower half of Table III shows that, at $Q_0 = 1.3$ GeV, $\bar{u}^{cs}$ and $\bar{d}^{cs}$ carry about 1.20% and 1.97% of the total momentum of the proton, respectively. Namely, $\bar{d}^{cs}$ carries more momentum than $\bar{u}^{cs}$. For comparison, $\bar{u}^{ds}$ and $\bar{d}^{ds}$ each carries about 1.67% of the total momentum of the proton. Totally, the CS and DS components of up- and down-quarks carry about 6.34% and 6.68% of the total momentum of the proton, respectively. In addition, the strange PDF only has DS component which accounts for 2.74% of proton's total momentum, with both s and $\bar{s}$ contributions included. This is driven by the input value of R taken from the lattice prediction of $1/R = \langle x\rangle_{s+\bar{s}}/\langle x\rangle_{\bar{u}+\bar{d}}(\mathrm{DI}) = 0.822(69)(78)$ at $Q = 1.3\,\mathrm{GeV}$, where $\langle x\rangle_{\bar{u}+\bar{d}}(\mathrm{DI})$ is the momentum fraction carried by the DS component of $\bar{u}$ and $\bar{d}$ partons. By separating the CS and DS components of partons in the global analysis, the predictions in Table III can be directly compared to lattice calculations of separate flavors in both the connected and disconnected insertions, term by term.

In Table IV, we collect the second moments of $u^+ - d^+ = (u+\bar{u}) - (d+\bar{d})$ and $s^+ = s + \bar{s}$ predicted by CT18CS and CT18 calculations, at 1.3 GeV and 2.0 GeV, respectively. Lattice results of $\langle x\rangle_{u^+-d^+}$ and $\langle x\rangle_{s+}$ at $Q = 2.0$ GeV are also given and they are found

TABLE IV. The second moments of $(u^+ - d^+)$ and s^+ predicted by CT18 [26] and CT18CS at 2.0 GeV and 1.3 GeV, respectively. We also show lattice results at 2.0 GeV. For $\langle x\rangle_{u^+-d^+}$, we follow Ref. [40] in supplying ranges obtained from various calculations, grouped according to the number of active flavors, N_f, in the lattice action used.

	$Q = 2.0$ GeV		$Q = 1.3$ GeV	
	CT18	Lattice	CT18CS	CT18
		$0.111 - 0.209^{N_f=2+1}$		
$\langle x\rangle_{u^+-d^+}$	0.156(7)	$0.153 - 0.194^{N_f=2+1+1}$ [40]	0.173(7)	0.175(8)
		$0.166 - 0.212^{N_f=2}$		
$\langle x\rangle_{s^+}$	0.033(9)	0.051(26)(5) [43]	0.027(8)	0.027(10)

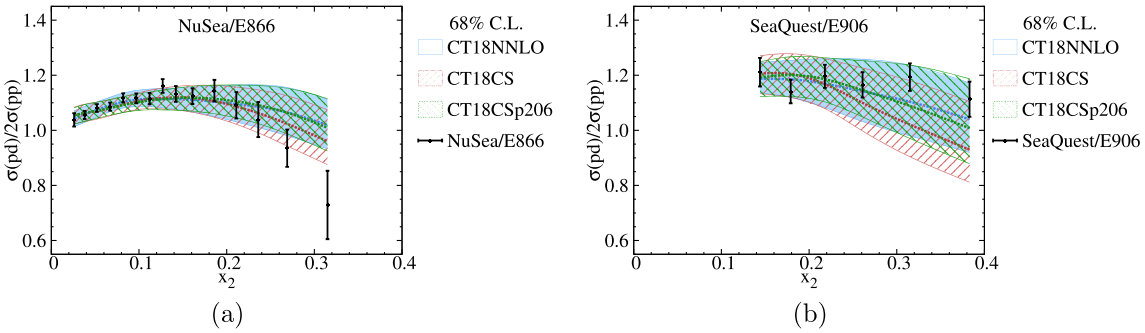

FIG. 7. Comparison of CT18 NNLO, CT18CS and CT18CSp206 predictions to the (a) E866 SeaQuest and (b) E906 SeaQuest data. Note that SeaQuest data were not included in the CT18NNLO and CT18CS fits, but are in the CT18CSp206 fit.

to be consistent with the CT18 predictions. However, we note that the deviation of the lattice calculations from different groups are large and not all systematic errors have been taken into account.

V. THE IMPACT OF SEAQUEST DATA

Fixed-target Drell-Yan measurements provide an important probe of the x dependence of the nucleon PDFs. This fact motivated the Fermilab E866 NuSea experiment [29], which determined the deuteron-to-proton cross section ratio $\sigma_{pd}/2\sigma_{pp}$ out to relatively large x_2, the momentum fraction of the target. Intriguingly, E866 found evidence that the cross section ratio dropped below unity, $\sigma_{pd}/2\sigma_{pp} < 1$, as x_2 approached and exceeded $x \gtrsim 0.25$. The E866 results stimulated an interest in performing a similar measurement out to larger x_2 with higher precision—the main objective of the subsequent E906 SeaQuest experiment at Fermilab [4]. Comparing to the NuSea data, the recent SeaQuest data include an extra bin which records data around $x \sim 0.4$ with high precision. In Fig. 7, we compare the predictions by CT18CS to the NuSea and SeaQuest data. For $x_2 > 0.2$, the NuSea and the SeaQuest data exhibit different shapes of $\sigma(pd)/2\sigma(pp)$. The ratio $\sigma(pd)/2\sigma(pp)$ for the NuSea data clearly decreases as x_2 becomes higher than 0.2, while for the SeaQuest data, this ratio seems to remain the same up to $x_2 = 0.4$. The difference in the shape of $\sigma(pd)/2\sigma(pp)$ distribution implies that NuSea and SeaQuest data have different preference for the PDF-ratio $\bar{d}/\bar{u}$ or the PDF-difference $\bar{d} - \bar{u}$ in the large-x region. In view of the fact that, in the CT18CS analysis, $\bar{q} = \bar{q}^{cs} + \bar{q}^{ds}$ for $q = u$ or d and $\bar{u}^{ds} = \bar{d}^{ds}$, the deviation of $\bar{d}/\bar{u}$ from unity is thus due to the different $\bar{u}^{cs}$ and $\bar{d}^{cs}$ contributions in the proton. Hence, it is interesting to know how the inclusion of the SeaQuest data in a global fit such as CT18CS could modify the PDF-difference $(\bar{d} - \bar{u})$, cf. Fig. 5, which is equal to $(\bar{d} - \bar{u})^{cs}$.

Below, we discuss the result of a new fit, referred to as "CT18CSp206" below, which follows the same approach as CT18CS, but with the inclusion of the E906 SeaQuest data to the original CT18 dataset. In Table V, we compare the quality of the CT18CSp206 fit to that of CT18CS. The only datasets with non-negligible $\Delta\chi^2 = |\chi^2_{\text{CT18CS}} - \chi^2_{\text{CT18CSp206}}|$ are just the E866 NuSea data and E906 SeaQuest data. From CT18CS to CT18CSp206 fit, the χ^2 for E866 NuSea data is increased by about 5 units, while the fit to the E906 SeaQuest data is improved (with a reduction of 12 units in its χ^2). This tension in the change of χ^2 reflects the different preferences of PDF-ratio $\bar{d}/\bar{u}$ or the PDF-difference $\bar{d} - \bar{u}$ in the large-x region. In Figs. 7 and 8, we compare the predictions of CT18CS and CT18CSp206 to the NuSea and SeaQuest data. In Fig. 7(a), the prediction of CT18CS is closer to the E866 NuSea data points for $x_2 > 0.2$, comparing to those of CT18 and CT18CSp206. For Fig. 7(b), the CT18CS prediction presents a different shape from E906 SeaQuest data points particularly for $x_2 > 0.2$, while CT18 and CT18CSp206 PDFs show better consistencies with these data points. Figure 8 shows the comparison of uncertainty sizes between the total experimental uncertainty and the PDF-induced uncertainty in

TABLE V. The χ^2 of selected datasets included in the CT18CS and CT18CSp206 fits. Only those with non-negligible $\Delta\chi^2 = |\chi^2_{\text{CT18CS}} - \chi^2_{\text{CT18CSp206}}|$ are listed. $N_{pt,E}$ is the number of data points of individual dataset, and χ^2_{CT18CS} and $\chi^2_{\text{CT18CSp206}}$ are the χ^2 values predicted by using the CT18CS and the CT18CSp206 fit. Note that the E906 SeaQuest data [4] (ID = 206) are not included in the CT18CS fit, but are in the CT18CSp206 fit.

ID	Experimental dataset	$N_{pt,E}$	χ^2_{CT18CS}	$\chi^2_{\text{CT18CSp206}}$
203	E866 Drell-Yan process $\sigma_{pd}/(2\sigma_{pp})$ [29]	15	13.5	18.8
206	E906 SeaQuest Drell-Yan process $\sigma_{pd}/(2\sigma_{pp})$ [4]	6	20.6	8.24

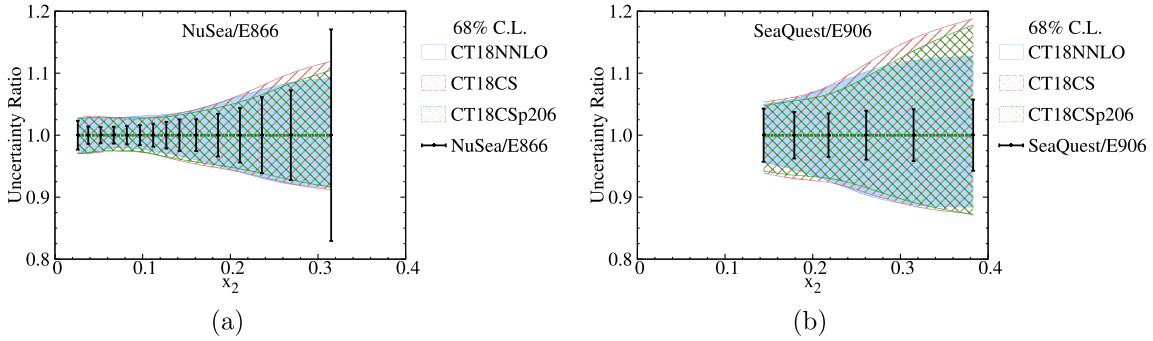

FIG. 8. Similar to Fig. 7, but for the comparison of the sizes of (total) experimental uncertainty of (a) E866 NuSea and (b) E906 SeaQuest experiments to the PDF-induced uncertainty predicted by CT18NNLO, CT18CS, and CT18CSp206.

predictions for both E866 NuSea and E906 SeaQuest data. All of three above-mentioned PDFs sets exhibit conservative uncertainties, so that the PDF-induced uncertainties in predictions are larger than the experimental uncertainty for both datasets, except for the data point with the highest x_2 value in E866 NuSea measurement. For $x_2 > 0.2$, the CT18CS predictions for both datasets possess slightly larger error bands than predictions of CT18 and CT18CSp206. For most of the range of x_2, the error band of CT18CSp206 is comparable to the CT18 error band, while in the prediction of E906 SeaQuest data with $x_2 > 0.3$, CT18CSp206 has a larger uncertainty.

Finally, we remark that the impact of SeaQuest data to modifying the CT18CS PDFs can also be studied by using the ePump-updating method, detailed in Refs. [50,51]. The idea is to add the SeaQuest data, with a given weight, to the original CT18 dataset and perform a new global fit using the ePump-updating method. This will update the original CT18CS PDFs and produce a new set of PDFs. Given this new set of PDFs, one can calculate the change in the total χ^2 of each dataset included in the global fit, as compared to that given by the original CT18CS PDFs. Instead of examining $\chi^2_E(N_{pt,E})/N_{pt,E}$ for the individual experiment E, which has different probability distribution and is dependent on the total number of data point $N_{pt,E}$, we provide an equivalent information in the form of the effective Gaussian variables $S_E = \sqrt{2\chi^2_E} - \sqrt{2N_{pt,E} - 1}$ [24]. A well-fitted dataset should have S_E between -1 and 1. An S_E smaller than -1 means the dataset is fitted too well (maybe due to large experimental errors) and an S_E larger than 1 indicates poor fitting. To examine the potential tensions between the E906 SeaQuest data and the datasets included in the CT18CS fit, we plot in Fig. 9(a) the change of the effective Gaussian variable S_E for some datasets included in CT18CS as the weight of SeaQuest data is increased from 0 to 10. Only the datasets with non-negligible change in S_E are shown. Note that a weight of zero corresponds to the CT18CS fit, in which the SeaQuest data were not included, and a weight of one leads to the above-mentioned CT18CSp206 fit. As the weight of SeaQuest data increases, the S_E of SeaQuest data becomes smaller, as expected, while the E866 NuSea data becomes much larger, indicating tension with the SeaQuest data. Both the NMC F_2^d/F_2^p (ID = 104) and CMS 8 TeV W and A_{ch} (ID = 249) data show very slight increase in their

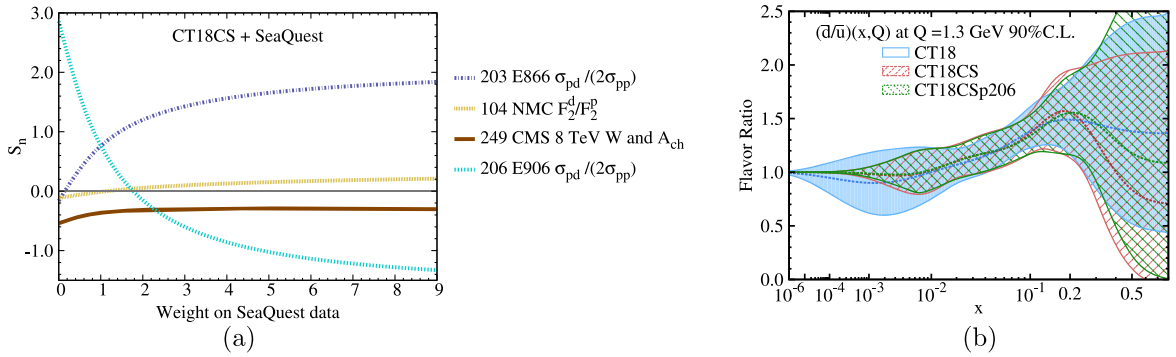

FIG. 9. (a) The change of the effective Gaussian variable S_E for some datasets included in CT18CS, as the weight of E906 SeaQuest data increases from 0 to 10. Only the datasets with notable changes in S_E are shown. Note that a weight of zero corresponds to the CT18CS fit, in which the SeaQuest data is not included. (b) Comparison of the PDF ratio $\bar{d}/\bar{u}$, as a function of x at $Q = 1.3$ GeV, among CT18, CT18CS, and CT18CSp206.

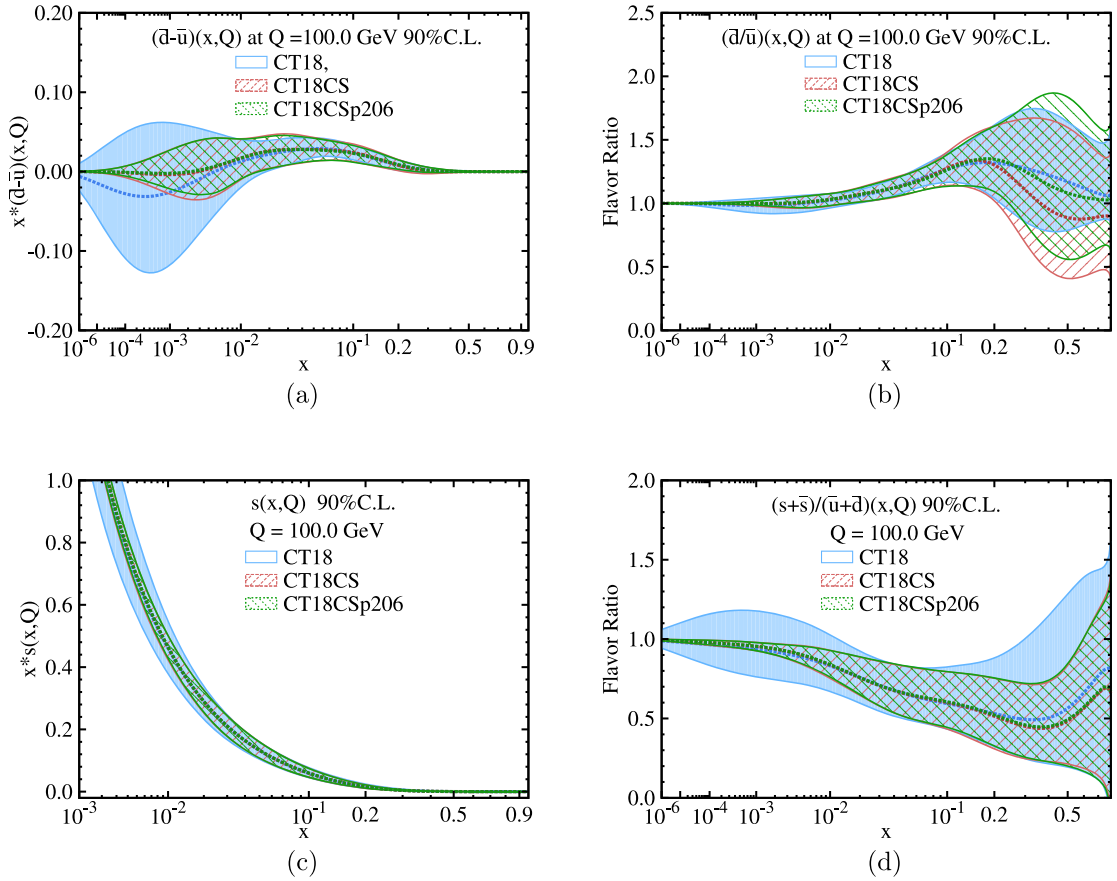

FIG. 10. The comparison of $\bar{d}-\bar{u}$, $\bar{d}/\bar{u}$, $s(x)$, and $(s+\bar{s})/(\bar{u}+\bar{d})$ PDFs at 100 GeV for CT18, CT18CS, and CT18CSp206.

S_E values as the weight of SeaQuest data increases from zero. In Fig. 9(b), we compare the PDF-ratio $\bar{d}/\bar{u}$, as a function of x at $Q = 1.3$ GeV, among CT18 NNLO, CT18CS, and CT18CSp206, where the E906 SeaQuest data (labeled as ID = 206 in Table V) is included, via the ePump-updating method [50,51]. It shows that CT18CSp206 has a larger PDF ratio $\bar{d}/\bar{u}$ at $x > 0.2$, as compared to CT18CS. On the other hand, the uncertainty of the PDF ratio $\bar{d}/\bar{u}$ of CT18CSp206 in large-x region is enlarged from that of CT18CS to tolerate the tension between the two datasets. For completeness, we also show in Fig. 10 the comparison of $\bar{d}/\bar{u}$, $(\bar{d}-\bar{u})$, s, and $(s+\bar{s})/(\bar{u}+\bar{d})$, respectively, as predicted by these three different global fits at $Q = 100$ GeV. In Figs. 10(a) and 10(b), at 100 GeV the comparison of PDF ratio $\bar{d}/\bar{u}$, or of the PDF difference $(\bar{d}-\bar{u})$, is similar to those at 1.3 GeV, cf. Fig. 10. The impact of the SeaQuest data on $s(x)$ and PDF ratio $(s+\bar{s})/(\bar{u}+\bar{d})$ at 100 GeV is negligible, as shown in Figs. 10(c) and 10(d).

VI. SUMMARY

In this work, we present a NNLO QCD global analysis named CT18CS where the connected sea partons and disconnected sea partons, as revealed in the path-integral formulation of the hadronic tensor in QCD, are separately parametrized at the input scale of $Q_0 = 1.3$ GeV. The CS and DS are mainly distinguished by their respective small-x behaviors. Furthermore, we assumed that the DS of u and d are proportional to the s with the proportional constant constrained by a recent complete lattice calculation of the second moment ratio [22] $\langle x\rangle_{s+\bar{s}}/\langle x\rangle_{\bar{u}+\bar{d}}(\mathrm{DI}) = 0.822(69)(78)$ at $Q = 1.3$ GeV, where $\langle x\rangle_{\bar{u}+\bar{d}}(\mathrm{DI})$ is the momentum fraction carried by the DS component of $\bar{u}$ and $\bar{d}$ partons. This lattice QCD constraint was included in the CT18CS fit via the Lagrange multiplier method. Together with the ansatz $a_1^s = 0$, this lattice input has helped reduce the error of the ratio $\langle x\rangle_{s+\bar{s}}/\langle x\rangle_{\bar{u}+\bar{d}}$ greatly for $x < 0.03$ as compared to that of the CT18 fit.

Short of applying the evolution equations where CS and DS partons are evolved separately, we impose a number of ansatzes regarding small-x behaviors and isospin symmetry, as described in Sec. III A. in the input scale and evolve the combined CS and DS partons during evolution. In this way, the PDFs are still evolved from Q_0 with the usual parton classification, namely $g, u, \bar{u}, d, \bar{d}, s$ and the results can be compared with CT18 at Q_0.

It is found that the fit quality of CT18CS is comparable to that of CT18 NNLO. The CT18CS PDFs, obtained with an extended parametrization, are consistent with CT18 NNLO in a wide range of x, but the errors of the quark partons in CT18CS are greatly reduced at small x as compared to those of CT18, mainly due to the small-x behaviors imposed and the lattice QCD input. As expected, the DS components primarily contribute to $\bar{u}$ and $\bar{d}$ in small-x region, and the CS components provide sizable contribution in the intermediate-x region. We give the second moments of CS and DS in different flavors at scale Q_0. They can be compared with systematic error controlled lattice calculations term by term for the first time. At $Q = 1.3$ GeV, we find that up and down quarks in the CS sector takes about 6.34% of total momentum, while the momentum in DS sector is about 6.68% of total amount. They are comparable in size at this low scale. The implication of CT18CS PDFs are studied in the comparison of predictions for the NuSea data and SeaQuest data between CT18CS and CT18 NNLO PDFs. A new global fit (referred to as CT18CSp206) on the basis of CT18CS is obtained with the SeaQuest data included. Through a scan of the effective Gaussian variable S_E over various weights to the E906 SeaQuest data, using the ePump-updating method [50,51], it is found that the SeaQuest data and the NuSea data are in tension.

In the future, global analyses should incorporate the extended evolution equations [41] where the connected sea and the disconnected sea are evolved separately so that they will remain separated at all Q^2 for better and more detailed delineation of the PDF degrees of freedom and compared to lattice results term by term.

ACKNOWLEDGMENTS

The authors are indebted to J. C. Peng, J. W. Qiu, and Y. B. Yang for insightful discussions. The work of K.-F. L. is partially support by the U.S. DOE Grant No. DE-SC0013065 and DOE Grant No. DE-AC05-06OR23177 which is within the framework of the TMD Topical Collaboration. This research used resources of the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725. This work used Stampede time under the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation Grant No. ACI-1053575. We also thank the National Energy Research Scientific Computing Center (NERSC) for providing HPC resources that have contributed to the research results reported within this paper. We acknowledge the facilities of the USQCD Collaboration used for this research in part, which are funded by the Office of Science of the U.S. Department of Energy. The work of J. L. is partially supported by the National Science Foundation of China (NSFC) under Grant No. 12175073. 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] Kurt Gottfried, Sum Rule for High-Energy Electron—Proton Scattering, Phys. Rev. Lett. **18**, 1174 (1967).
[2] P. Amaudruz *et al.*, The Gottfried Sum from the Ratio F2(n)/F2(p), Phys. Rev. Lett. **66**, 2712 (1991).
[3] M. Arneodo *et al.*, A reevaluation of the Gottfried sum, Phys. Rev. D **50**, R1 (1994).
[4] J. Dove *et al.*, The asymmetry of antimatter in the proton, Nature (London) **590**, 561 (2021).
[5] S. J. Brodsky, P. Hoyer, C. Peterson, and N. Sakai, The intrinsic charm of the proton, Phys. Lett. **93B**, 451 (1980).
[6] Tie-Jiun Hou, Sayipjamal Dulat, Jun Gao, Marco Guzzi, Joey Huston, Pavel Nadolsky, Carl Schmidt, Jan Winter, Keping Xie, and C. P. Yuan, CT14 intrinsic charm parton distribution functions from CTEQ-TEA global analysis, J. High Energy Phys. 02 (**2018**) 059.
[7] Richard D. Ball, Valerio Bertone, Marco Bonvini, Stefano Carrazza, Stefano Forte, Alberto Guffanti, Nathan P. Hartland, Juan Rojo, and Luca Rottoli, A determination of the charm content of the proton, Eur. Phys. J. C **76**, 647 (2016).
[8] Richard D. Ball *et al.*, Parton distributions from high-precision collider data, Eur. Phys. J. C **77**, 663 (2017).
[9] Richard D. Ball *et al.* (NNPDF Collaboration), The path to proton structure at 1% accuracy, Eur. Phys. J. C **82**, 428 (2022).
[10] S. Davidson, S. Forte, P. Gambino, N. Rius, and A. Strumia, Old and new physics interpretations of the NuTeV anomaly, J. High Energy Phys. 02 (2002) 037.
[11] Stefan Kretzer, Fredrick Olness, Jon Pumplin, Daniel Stump, Wu-Ki Tung, and Mary Hall Reno, The Parton Structure of the Nucleon and Precision Determination of the Weinberg Angle in Neutrino Scattering, Phys. Rev. Lett. **93**, 041802 (2004).
[12] H. L. Lai, Pavel M. Nadolsky, J. Pumplin, D. Stump, W. K. Tung, and C. P. Yuan, The strange parton distribution of the nucleon: Global analysis and applications, J. High Energy Phys. 04 (2007) 089.
[13] 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).

[14] Raza Sabbir Sufian, Tianbo Liu, Andrei Alexandru, Stanley J. Brodsky, Guy F. de Téramond, Hans Günter Dosch, Terrence Draper, Keh-Fei Liu, and Yi-Bo Yang, Constraints on charm-anticharm asymmetry in the nucleon from lattice QCD, Phys. Lett. B **808**, 135633 (2020).
[15] Keh-Fei Liu and Shao-Jing Dong, Origin of Difference between Anti-d and Anti-u Partons in the Nucleon, Phys. Rev. Lett. **72**, 1790 (1994).
[16] Keh-Fei Liu, Parton degrees of freedom from the path integral formalism, Phys. Rev. D **62**, 074501 (2000).
[17] Jian Liang, Terrence Draper, Keh-Fei Liu, Alexander Rothkopf, and Yi-Bo Yang, Towards the nucleon hadronic tensor from lattice QCD, Phys. Rev. D **101**, 114503 (2020).
[18] Jian Liang and Keh-Fei Liu, PDFs and neutrino-nucleon scattering from hadronic tensor, *Proc. Sci.*, LATTICE2019 (2020) 046.
[19] Keh-Fei Liu, Wen-Chen Chang, Hai-Yang Cheng, and Jen-Chieh Peng, Connected-Sea Partons, Phys. Rev. Lett. **109**, 252002 (2012).
[20] Jen-Chieh Peng, Wen-Chen Chang, Hai-Yang Cheng, Tie-Jiun Hou, Keh-Fei Liu, and Jian-Wei Qiu, On the momentum dependence of the flavor structure of the nucleon sea, Phys. Lett. B **736**, 411 (2014).
[21] J. C. Peng, W. C. Chang, H. Y. Cheng, and K. F. Liu, Flavor structure of intrinsic nucleon sea, Few-Body Syst. **56**, 349 (2015).
[22] Jian Liang, Mingyang Sun, Yi-Bo Yang, Terrence Draper, and Keh-Fei Liu, Ratio of strange to u/d momentum fraction in disconnected insertions, Phys. Rev. D **102**, 034514 (2020).
[23] A. Airapetian *et al.*, Measurement of parton distributions of strange quarks in the nucleon from charged-kaon production in deep-inelastic scattering on the deuteron, Phys. Lett. B **666**, 446 (2008).
[24] Hung-Liang Lai, Marco Guzzi, Joey Huston, Zhao Li, Pavel M. Nadolsky, Jon Pumplin, and C. P. Yuan, New parton distributions for collider physics, Phys. Rev. D **82**, 074024 (2010).
[25] Takumi Doi, Mridupawan Deka, Shao-Jing Dong, Terrence Draper, Keh-Fei Liu, Devdatta Mankame, Nilmani Mathur, and Thomas Streuer, Strangeness and glue in the nucleon from lattice QCD, *Proc. Sci.*, LATTICE2008 (2008) 163.
[26] Tie-Jiun Hou *et al.*, New CTEQ global analysis of quantum chromodynamics with high-precision data from the LHC, Phys. Rev. D **103**, 014013 (2021).
[27] Huey-Wen Lin, W. Melnitchouk, Alexei Prokudin, N. Sato, and H. Shows, First Monte Carlo Global Analysis of Nucleon Transversity with Lattice QCD Constraints, Phys. Rev. Lett. **120**, 152502 (2018).
[28] Tie-Jiun Hou, Jian Liang, Keh-Fei Liu, Mengshi Yan, and C. P. Yuan, Connected and Disconnected Sea Partons from CT18 Parametrization of PDFs, SciPost Phys. Proc. **8**, 067 (2022).
[29] R. S. Towell *et al.*, Improved measurement of the anti-d/anti-u asymmetry in the nucleon sea, Phys. Rev. D **64**, 052002 (2001).
[30] Keh-Fei Liu, Parton distribution function from the hadronic tensor on the lattice, Proc. Sci., LATTICE2015 (**2016**) 115.
[31] Jian Liang, Keh-Fei Liu, and Yi-Bo Yang, Lattice calculation of hadronic tensor of the nucleon, EPJ Web Conf. **175**, 14014 (2018).
[32] K. F. Liu, S. J. Dong, Terrence Draper, D. Leinweber, J. H. Sloan, W. Wilcox, and R. M. Woloshyn, Valence QCD: Connecting QCD to the quark model, Phys. Rev. D **59**, 112001 (1999).
[33] Keh-Fei Liu, PDF in PDFs from hadronic tensor and LaMET, Phys. Rev. D **102**, 074502 (2020).
[34] Hans A. Bethe and Jeffrey Goldstone, Effect of a repulsive core in the theory of complex nuclei, Proc. R. Soc. A **238**, 551 (1957).
[35] Xiangdong Ji, Parton Physics on a Euclidean Lattice, Phys. Rev. Lett. **110**, 262002 (2013).
[36] A. V. Radyushkin, Quasi-parton distribution functions, momentum distributions, and pseudo-parton distribution functions, Phys. Rev. D **96**, 034025 (2017).
[37] Yan-Qing Ma and Jian-Wei Qiu, Exploring Partonic Structure of Hadrons Using *ab initio* Lattice QCD Calculations, Phys. Rev. Lett. **120**, 022003 (2018).
[38] John C. Collins, Davison E. Soper, and George F. Sterman, Factorization of hard processes in QCD, Adv. Ser. Dir. High Energy Phys. **5**, 1 (1989).
[39] Huey-Wen Lin *et al.*, Parton distributions and lattice QCD calculations: A community white paper, Prog. Part. Nucl. Phys. **100**, 107 (2018).
[40] Martha Constantinou *et al.*, Parton distributions and lattice-QCD calculations: Toward 3D structure, Prog. Part. Nucl. Phys. **121**, 103908 (2021).
[41] Keh-Fei Liu, Evolution equations for connected and disconnected sea parton distributions, Phys. Rev. D **96**, 033001 (2017).
[42] Richard D. Ball, Alessandro Candido, Juan Cruz-Martinez, Stefano Forte, Tommaso Giani, Felix Hekhorn, Kirill Kudashkin, Giacomo Magni, and Juan Rojo, Evidence for intrinsic charm quarks in the proton, Nature (London) **608**, 483 (2022).
[43] Yi-Bo Yang, Jian Liang, Yu-Jiang Bi, Ying Chen, Terrence Draper, Keh-Fei Liu, and Zhaofeng Liu, Proton Mass Decomposition from the QCD Energy Momentum Tensor, Phys. Rev. Lett. **121**, 212001 (2018).
[44] M. Arneodo *et al.*, Measurement of the proton and deuteron structure functions, F2(p) and F2(d), and of the ratio sigma-L/sigma-T, Nucl. Phys. **B483**, 3 (1997).
[45] David Alexander Mason, Measurement of the strange-antistrange asymmetry at NLO in QCD from NuTeV dimuon data, Ph.D. thesis, Oregon University, 2006.
[46] G. Moreno *et al.*, Dimuon production in proton-copper collisions at $\sqrt{s} = 38.8$-GeV, Phys. Rev. D **43**, 2815 (1991).
[47] J. C. Webb *et al.*, Absolute Drell-Yan dimuon cross-sections in 800 GeV/c pp and pd collisions, arXiv:hep-ex/0302019.
[48] Roel Aaij *et al.*, Measurement of forward $Z \to e^+e^-$ production at $\sqrt{s} = 8$ TeV, J. High Energy Phys. 05 (2015) 109.
[49] Vardan Khachatryan *et al.*, Measurement of the differential cross section and charge asymmetry for inclusive $pp \to W^{\pm} + X$ production at $\sqrt{s} = 8$ TeV, Eur. Phys. J. C **76**, 469 (2016).

[50] Carl Schmidt, Jon Pumplin, and C.-P. Yuan, Updating and optimizing error parton distribution function sets in the Hessian approach, Phys. Rev. D **98**, 094005 (2018).

[51] Tie-Jiun Hou, Zhite Yu, Sayipjamal Dulat, Carl Schmidt, and C. P. Yuan, Updating and optimizing error parton distribution function sets in the Hessian approach. II, Phys. Rev. D **100**, 114024 (2019).