

R reinterpretation of proton strangeness between ATLAS and CMS measurements

Alim Ruzi

*School of Physics and State Key Laboratory of Nuclear Physics and Technology,
Peking University, Beijing 100871, China*

Bo-Qiang Ma^{✉*}

*School of Physics and State Key Laboratory of Nuclear Physics and Technology, Peking University,
Beijing 100871, China,*

*Center for High Energy Physics, Peking University, Beijing 100871, China,
and Collaborative Innovation Center of Quantum Matter, Beijing, China*



(Received 2 June 2021; accepted 9 September 2021; published 6 October 2021)

We reexamine the shapes of the strange quark parton distribution functions (PDFs) of the proton by means of QCD analysis of HERA I + II $e^\pm p$ deep inelastic scattering cross section measurement at DESY and inclusive gauge boson production and W boson production associated with a charm quark from the LHC at CERN. We find that there is an overall agreement on the strange quark distributions obtained from CMS W + charm and ATLAS W/Z data at the parton momentum fraction range $x \lesssim 10^{-2}$. Meanwhile, there is also a strong tension between these data toward large x . We find that this tension fades away if the ATLAS measurement of W/Z production is analyzed together with the ATLAS W + charm data. The W/Z and W + charm data both from ATLAS and CMS experiments agree that the proton strangeness is enhanced toward small momentum fraction x and is smoothly suppressed at large x . Furthermore, a strong x dependence of the strange-to-nonstrange parton ratio $R_s(x, Q^2)$ is observed.

DOI: [10.1103/PhysRevD.104.076004](https://doi.org/10.1103/PhysRevD.104.076004)

I. INTRODUCTION

It is well known that nucleons consist of pointlike particles denominated as partons, revealed by lepton nucleon deep inelastic scattering (DIS) experiment [1,2] almost half a century ago. There have been tremendous efforts to precisely determine the parton distribution functions (PDFs) of the proton worldwide. Distributions of the valence-up quarks, valence-down quarks, and gluons are determined to a level that could give, by and large, a consistent result on physical quantities in the standard model (SM) obtained from high-energy collision experiments [3–7]. However, the strange content of the proton, which is poorly known because of limited sources of strangeness-sensitive data, is still ambiguous to our understandings of the partonic structure of the proton. Strange quark PDFs of nucleons play a vital role for a number of physics processes at the LHC, ranging from measurements

of W boson production in association with a charm jet [8] and determinations of electroweak (EW) interaction parameters to the formation of strange hadrons [9].

The size and shape of the strange quark PDFs have recently attracted a lot of interest and ignited a hot debate on specialized strange studies. The strange quark PDFs obtained from dimuon production in neutrino-nucleus DIS experiment by the NuTeV and NOMAD collaborations [10–12] are significantly suppressed, while the inclusive W/Z production data by the ATLAS Collaboration [13] give strongly enhanced distribution of strangeness. The ratio of strange to nonstrange sea quark PDFs, $R_s = (s + \bar{s})/(\bar{u} + \bar{d})$, from neutrino-nucleus DIS experiment is approximately 0.5 when evaluated at $x = 0.023$ and energy scale $Q = 1.6$ GeV, whereas the strange to antidown quark ratio, $r_s = (s + \bar{s})/2\bar{d}$, evaluated at $x = 0.023$ and $Q^2 = 1.9$ GeV² from the ATLAS analysis of the inclusive gauge boson production together with the combined HERA I + II $e^\pm p$ data [14], rises up to unity.

In their QCD analysis of accumulated W/Z data to extract ATLAS-epWZ16 [15], the ATLAS Collaboration obtains even larger strange-to-nonstrange ratio, $R_s = 1.131$ [15], showing agreement with their older analysis [13]. Another study of the ATLAS and CMS W/Z cross section measurements in Ref. [16] reports that there is no tension

*mabq@pku.edu.cn

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 SCOAP³.

between the LHC W/Z and the HERA data or between the ATLAS and CMS data sets, indicating that the LHC measurements of inclusive W/Z boson production support unsuppressed strangeness in the proton at low x at both low- and high-energy scales.

A recent global analysis of vector gauge boson production in association with jets plus older W/Z and HERA $e^\pm p$ data by the ATLAS Collaboration [17] reports somewhat different R_s distribution from the one in ATLAS-epWZ16 PDF sets. The ratio R_s in Ref. [17] keeps consistency with ATLAS-epWZ16 at $x \lesssim 10^{-2}$ and declines faster at the other x value. But the strange quark in that analysis still stays comparable or even larger than the nonstrange sea quarks. The very recent global PDFs, MHST20 PDFs [18], extracted from almost all of the available measurements show an R_s value compatible with that of the ATLAS-epWZ16.

So, it comes as a puzzle that both the ATLAS and CMS measurements of W/Z boson production support an enhanced strangeness inside the proton, while the fixed target lepton-nuclear collision prefers a rather suppressed strangeness. Among all the standard model processes, the $Ws \rightarrow c$ interactions have the utmost sensitivity to the strange (antistrange) quark PDFs inside the proton. The QCD analysis of W production in association with a charm (anticharm) quark in the proton-proton (pp) collision at $\sqrt{s} = 7$ TeV by the ATLAS Collaboration shows the ratio of strange to down sea quark, $r_s = (s + \bar{s})/2\bar{d}$, to be 0.96 [19], supporting the results obtained from the analysis of W/Z measurements. However, the CMS Collaboration, in the analysis of the same process, $W +$ charm production plus W -lepton charge asymmetry [20–23], reports that they have not observed the enhanced strangeness as the ATLAS Collaboration did. Instead, they observed an x -dependent and suppressed ratio R_s . Yet, in another dedicated analysis on the strange quark distribution by the Hessian error updating method [24,25], it is reported that the CMS $W +$ charm quark production data [22] enhance strangeness of CT14 PDFs families. The recently extracted CT18 PDFs series from fixed target and up-to-date collider data by the Coordinated Theoretical-Experimental Project On QCD (CTEQ) group [26] demonstrate that the strange quark density is enhanced over the nominal fit CT18NNLO when the inclusive W/Z data are included. Another devoted study on the strangeness of the proton [27], which includes both the lepton-nucleus DIS cross section data from the NuTeV and NOMAD plus $W +$ charm data from the ATLAS and CMS experiments, reports a slightly enhanced strange quark PDF within uncertainties in terms of the ratio $R_s(x, Q^2)$ with huge uncertainties at large momentum fraction.

Given the important influence of the strange quark PDFs in SM physics inside the LHC, and the contradictory interpretations of the strange-sensitive data from the ATLAS and CMS measurements, we regard reinvestigation of the exact shapes of the strange quark PDFs of the

nucleon by a global analysis of strangeness-sensitive measurements, within an accurate theoretical and methodological framework, as necessary to clarify the controversial interpretations of the strangeness-related quantities, such as $xs(x, Q^2)$ and the ratio R_s . This paper is intended to fulfill this purpose. We present a fairly detailed PDFs extraction with the general-purpose PDFs fitting program XFITTER [28] from the following data: HERA I + II $e^\pm p$ [14], ATLAS combined inclusive W/Z production cross section [13,15], and W production in association with a charm quark [19]; CMS W -lepton charge asymmetry [20,21] and W production in association with charm quark [22,23], applying variants of different parametrization forms first to clarify whether the same physics processes in ATLAS and CMS experiments give compatible or incompatible distributions of the strange quark and, second, to pin down its exact shape.

II. INPUT DATA AND CORRESPONDING THEORETICAL PREDICTIONS

A. Data description

The process of $e^\pm p$ DIS has been the central dataset to many PDFs working groups in their global analyses and provides the most constraining power to quark PDFs when parametrized as a composite functional form. The gauge boson production data at the LHC also provide significant constraints on the flavor decomposition where the HERA data fail. The $W +$ charm production offers the best constraints on the strange quark PDFs and gluon distributions through the subprocesses: $g + s \rightarrow W^- + c$ and $\bar{s} + g \rightarrow W^+ + \bar{c}$. The W -lepton charge asymmetry data from CMS, available for 7 and 8 TeV, can constrain the valence quarks very well and have an indirect impact on the strange content.

We do not include neutrino-induced dimuon data in our analysis. The dimuon data were obtained from $\nu_\mu(\bar{\nu}_\mu)A$ fixed target DIS process where the target is iron, whose internal structure is more complex than the free nucleon. In this process, however, the nuclear European Muon Collaboration (EMC) effect should be taken into account in addition to some nuclear corrections; see Ref. [29] and references therein. Considering the non-negligible nuclear corrections in the lepton-nuclear collision experiment, we believe that pure-proton-related data, which are free from the EMC effect, are significantly useful for our analysis. Besides that, the strange quark distribution extracted from dimuon production in neutrino scattering differs significantly from the one extracted from LHC data, especially the ATLAS W/Z production process. The latter prefers a rather higher R_s value, and the former prefers a lower value. One more disadvantage of dimuon data is that their constraining power vanishes at $x \lesssim 0.01$ due to shadowing and antishadowing effects in the neutrino-nucleus DIS process [30].

The description of data is given below:

- (i) The HERA I + II $e^\pm p$ combined data [14] include measurements acquired from run 1 during 1992 to 2000 and run 2 during 2002 to 2007 in $e^\pm p$ collision at different electron/positron beam energies with approximately the total integrated luminosities of 500 pb^{-1} . The datasets are divided between $e^\pm p$ charged current (CC) and neutral current (NC) scattering according to the intermediate particle being W^\pm or Z/γ^* . The $e^\pm p$ CC and $e^- p$ NC interaction scattering cross sections are measured at $\sqrt{s} = 318 \text{ GeV}$. The rest of the datasets are made up of $e^+ p$ NC scattering cross sections with positron beam energy of $E_e = 27.5 \text{ GeV}$ and proton beam energies of $E_p = 920, 820, 575$ and 460 GeV , corresponding to total center-mass energies $\sqrt{s} = 318, 300, 251, \text{ and } 225 \text{ GeV}$. These datasets when combined give in total 1307 data points on differential cross sections being a function of the Bjorken variable x , approximately the light-cone momentum fraction of the quark, and the transferred-energy squared Q^2 . For the published datasets, these two kinematic variables can cover $6 \times 10^{-7} \leq x \leq 0.65$, $0.045 \leq Q^2 \leq 50000 \text{ GeV}^2$ in the NC interaction scattering and $1.3 \times 10^{-2} \leq x \leq 0.40$, $200 \leq Q^2 \leq 50000 \text{ GeV}^2$ in the CC interaction scattering.
- (ii) The ATLAS W/Z data [15] include the cross section measurements of W^\pm production in the leptonic decay channels: $W^+ \rightarrow l^+ \nu$, $W^- \rightarrow l^- \bar{\nu}$, and $Z \rightarrow ll$ ($l = e, \mu$) production processes at $\sqrt{s} = 7 \text{ TeV}$ with an integrated luminosity of 4.6 fb^{-1} . The differential cross sections for W^\pm are given as the function of lepton pseudorapidity in the range $|\eta_l| < 2.5$ with 22 data points in total. The cross section of Z/γ^* production is measured as a function of the absolute dilepton rapidity $|y_{ll}|$ in the central and forward region, for three intervals of dilepton invariant mass or the intermediate Z boson mass, $46 < m_{ll} < 66 \text{ GeV}$, $66 < m_{ll} < 116 \text{ GeV}$ (forward and central rapidity), and $116 < m_{ll} < 150 \text{ GeV}$ (forward and central rapidity). The Z/γ^* data include 39 data points in total.
- (iii) The data of $W + \text{charm}$ measurement come from ATLAS and CMS experiments. The ATLAS measurements of W^\pm production in association with a(n) (anti)charm quark are available for $\sqrt{s} = 7 \text{ TeV}$ with total 22 data points [19], while the CMS data of this process are available for 7 [22] and 13 TeV [23] averaged for $W^+ + \bar{c}$ and $W^- + c$ with a total of five points. Both ATLAS and CMS measurements of $W + \text{charm}$ production are given as function of lepton pseudorapidity, η^μ .
- (iv) The W -lepton charge asymmetry, defined as $\mathcal{A} = \frac{\sigma(W^+) - \sigma(W^-)}{\sigma(W^+) + \sigma(W^-)}$, from CMS is available for 7 [20] and 8 TeV [21].

B. Theoretical calculation

The corresponding predictions of $e^\pm p$ cross sections for HERA I + II $e^\pm p$ data [14] are obtained by solving the DGLAP evolution equations [31–34] at next-to-next-to-leading-order (NNLO) in the $\overline{\text{MS}}$ scheme [35] through the QCDNUM program [36] interfaced with XFITTER. The renormalization and factorization scales are chosen to be $\mu_r^2 = \mu_f^2 = Q^2$. The initial factorization scale is set to be $Q_0^2 = 1.9 \text{ GeV}^2$ just below the charm quark mass $m_c = 1.43 \text{ GeV}$ in this analysis. The evolved PDFs at each scale point are convoluted with the coefficient functions of the structure functions of the proton to obtain the corresponding differential cross sections for $e^\pm p$ scattering. The contributions to the structure functions from heavy quarks are calculated in the general-mass-variable-flavor-number scheme [37–39] in the NC interactions. For the structure functions in CC interactions, the zero-mass approximation is used because the HERA CC data have $Q^2 \gg M_H^2$, where M_H is the heavy quark mass, i.e., the masses of charm and bottom quarks. The calculation of the differential cross section for the gauge boson production is done using the well-known Monte Carlo program MCFM [40,41] with the CT18NNLO PDF set [26]. The W^\pm and Z boson cross sections are available up to NNLO. Because of the intensiveness of this calculation, a fast convolution technique is applied. First of all, the partonic cross sections in pp collision are obtained in the form of grid files using MCFM with the help of another fast parton convolution program APPLGRID [42] interfaced both with MCFM and XFITTER; then, the corresponding cross sections are obtained convoluting the grid files with the evolved PDFs during the fit. So far, APPLGRID can only calculate cross sections up to next-to-leading-order (NLO) accuracy. To get the NNLO total and differential cross sections, a K -factor formalism is applied. The K -factor is defined as

$$K = \frac{\sigma_{\text{NNLO QCD}}^{\text{NLO EW}}}{\sigma_{\text{NLO QCD}}^{\text{LO EW}}}. \quad (1)$$

For both the CMS and ATLAS $W + \text{charm}$ data, only the NLO calculations are used, and this is obtained via MCFM. The theoretical prediction for $W + \text{charm}$ process is only available at NLO calculation, while that for the Drell-Yan and W^\pm production can be obtained up to NNLO. The measurements for the W/Z data [15] are given in high precision, so it is important to get as accurate theoretical predictions as possible. During the fit, the order of calculation is specified according to the physics process under consideration. All the predictions are calculated in the respective fiducial phase space of the experimental data. The K -factors are evaluated bin by bin with the same PDFs in both the numerator and denominator.

III. ANALYSIS SETUP

We are curious about the ratio R_s given by CMS analysis of $W + \text{charm}$ production together with W -lepton charge asymmetry data [23] being much different from the one reported by the ATLAS group [15]. The CMS analysis [23] made a comparison of the ratio R_s to the one obtained with ATLAS-epWZ16 PDFs [15] and ABMP16 PDFs [43]. R_s from ATLAS is almost constant throughout the entire x except at high x , and the property of being greater than 1 is interpreted as enhancement of strange quark PDFs relative to nonstrange quark PDFs by the ATLAS group. On the contrary, R_s of CMS is fully x dependent and declining faster as x is getting larger, interpreted as suppressed strangeness by CMS. This is the main difference between ATLAS W/Z and CMS $W + \text{charm}$ analyses in terms of the ratio R_s . Considered as dominant data sources in constraining the light flavor quark and antiquark PDFs, the ATLAS W/Z and CMS $W + \text{charm}$ data should not be controversial as far as the strange quark PDFs are concerned. Below, we are going to explore both the analysis setup of the ATLAS [15] and CMS [23] analyses. The ATLAS-epWZ16 is extracted at NNLO accuracy in perturbative QCD using inclusive HERA I + II $e^\pm p$ [14] and ATLAS W/Z data [15] through the xFITTER framework. The PDFs in ATLAS-epWZ16 are parametrized at a starting scale $Q_0^2 = 1.9 \text{ GeV}^2$ as

$$xg(x) = A_g x^{B_g} (1-x)^{C_g} - A'_g x^{B'_g} (1-x)^{C'_g}, \quad (2a)$$

$$xu_v(x) = A_{u_v} x^{B_{u_v}} (1-x)^{C_{u_v}} (1 + E_{u_v} x^2), \quad (2b)$$

$$xd_v(x) = A_{d_v} x^{B_{d_v}} (1-x)^{C_{d_v}}, \quad (2c)$$

$$x\bar{u}(x) = A_{\bar{u}} x^{B_{\bar{u}}} (1-x)^{C_{\bar{u}}}, \quad (2d)$$

$$x\bar{d}(x) = A_{\bar{d}} x^{B_{\bar{d}}} (1-x)^{C_{\bar{d}}}, \quad (2e)$$

$$x\bar{s}(x) = xs(x) = A_{\bar{s}} x^{B_{\bar{s}}} (1-x)^{C_{\bar{s}}}. \quad (2f)$$

The relationship $A_{\bar{u}} = A_{\bar{d}}$ and $B_{\bar{u}} = B_{\bar{d}} = B_{\bar{s}}$ are set for the sea quark parameters based on the assumption $\bar{u} = \bar{d}$ as $x \rightarrow 0$, complying with the usual way quoted by the HERA Collaboration [14].

On the contrary, the CMS Collaboration applies similar shape functions at $Q_0^2 = 1.9 \text{ GeV}^2$ for their analysis of $W + \text{charm}$ quark production together with the W -lepton charge asymmetry data to extract PDFs at NLO using xFITTER. The applied parametrization form in the CMS analysis is

$$xg(x) = A_g x^{B_g} (1-x)^{C_g}, \quad (3a)$$

$$xu_v(x) = A_{u_v} x^{B_{u_v}} (1-x)^{C_{u_v}} (1 + E_{u_v} x^2), \quad (3b)$$

$$xd_v(x) = A_{d_v} x^{B_{d_v}} (1-x)^{C_{d_v}}, \quad (3c)$$

$$x\bar{u}(x) = A_{\bar{u}} x^{B_{\bar{u}}} (1-x)^{C_{\bar{u}}} (1 + E_{\bar{u}} x^2), \quad (3d)$$

$$x\bar{d}(x) = A_{\bar{d}} x^{B_{\bar{d}}} (1-x)^{C_{\bar{d}}}, \quad (3e)$$

$$x\bar{s}(x) = xs(x) = A_{\bar{s}} x^{B_{\bar{s}}} (1-x)^{C_{\bar{s}}}. \quad (3f)$$

In their parametrization form, the CMS Collaboration applies no extra relationship between sea quark parameters except for the renormalization parameters of \bar{u} and \bar{d} , $A_{\bar{u}} = A_{\bar{d}}$. They put an extra parameter that directly affects the strange quark PDFs: strangeness fraction $f_s = \bar{s}/(\bar{d} + \bar{s})$, often used by the HERA group in their PDFs analysis [14] as well. This strangeness fraction number is set as a free parameter in the fit. Power parameters B of all the light flavor sea quarks, \bar{u} , \bar{d} , \bar{s} are also set free from each other.

To know whether there is a tension between ATLAS W/Z and CMS $W + \text{charm}$ data on the strange-related distributions, we conduct series of fits to the HERA I + II $e^\pm p$, ATLAS, and CMS $W + \text{charm}$ data together with the available W -lepton charge asymmetry measurement, applying both Eqs. (2) and (3) using the xFITTER program. After clarifying the reason behind the contradictory results between ATLAS and CMS data, it is necessary to pin down what the distribution of strange quark PDFs looks like. To achieve this, we consider that the combined usage of all the data available here in a fit with a flexible parametrization form to accommodate the experimental uncertainties is much preferable to the analysis of merely ATLAS or CMS data using Eqs. (2) and (3). In this way, the strange quark PDFs could be even more constrained. Thus, we do one more round of fit with the following parametrization form at $Q_0^2 = 1.9 \text{ GeV}^2$ including all the data at once:

$$xg(x) = A_g x^{B_g} (1-x)^{C_g} - A'_g x^{B'_g} (1-x)^{C'_g}, \quad (4a)$$

$$xu_v(x) = A_{u_v} x^{B_{u_v}} (1-x)^{C_{u_v}} (1 + E_{u_v} x^2), \quad (4b)$$

$$xd_v(x) = A_{d_v} x^{B_{d_v}} (1-x)^{C_{d_v}}, \quad (4c)$$

$$x\bar{u}(x) = A_{\bar{u}} x^{B_{\bar{u}}} (1-x)^{C_{\bar{u}}} (1 + D_{\bar{u}} x), \quad (4d)$$

$$x\bar{d}(x) = A_{\bar{d}} x^{B_{\bar{d}}} (1-x)^{C_{\bar{d}}} (1 + D_{\bar{d}} x), \quad (4e)$$

$$x\bar{s}(x) = xs(x) = A_{\bar{s}} x^{B_{\bar{s}}} (1-x)^{C_{\bar{s}}} (1 + D_{\bar{s}} x). \quad (4f)$$

In this parametrization form, we add an extra parameter D to the sea quarks and put no restrictions between the individual PDF parameters, so that it might adapt the tensions between individual partons and measurements well. Setting $s = \bar{s}$ is the usual way used by many PDFs working groups because of the limited data constraints on the strange quark distribution. By taking difference of

$\sigma(W^+\bar{c})$ and $\sigma(W^-c)$, it is possible to extract information on $x(s-\bar{s})$. Recently, the newly extracted PDFs MSHT20 [18] and the reference [27] obtain very small, order of 10^{-3} , $x(s-\bar{s})$ distribution by analyzing both the dimuon and LHC data. Given some limitations in data on the strange quark distribution, we use symmetric strangeness throughout in our analysis. Actually, it was predicted [44] that the momentum and helicity distributions of the strange and anti-strange quarks are different from each other. Hopefully, this problem will be solved by analyzing relevant data from the near-future high-luminosity and high-energy experiments and with more precise theoretical calculations.

The denomination for each round of fit goes as follows:

- (i) *Fit 1*.—Applying the parametrization form of Eq. (2) with HERA I + II $e^\pm p$ [14] and ATLAS W/Z data [15].
- (ii) *Fit 2*.—Applying the parametrization form of Eq. (3) with HERA I + II $e^\pm p$ [14] and CMS W + charm and W -lepton charge asymmetry data [20–23].
- (iii) *Fit 1p*.—Adding ATLAS W + charm data [19] to Fit 1.
- (iv) *PKU-fit*.—Applying Eq. (4) with all the data.

To get a data-driven characteristics of the PDFs, we get rid of all the restrictions set on the fit parameters, $A_{\bar{u}} = A_{\bar{d}} = A_{\bar{s}}$ and $B_{\bar{u}} = B_{\bar{d}} = B_{\bar{s}}$ in the above fits. The initial factorization and renormalization scale is still 1.9 GeV^2 . The heavy quark masses are set as $m_c = 1.43 \text{ GeV}$ and $m_b = 4.5 \text{ GeV}$. The strong coupling constant α_s is set to be 0.111 suggested in

the HERAPDF2.0 [14] through a χ^2 scan on α_s . We did conduct a few rounds of fits with α_s values equal to 0.110, 0.111, 0.113, 0.115, 0.117, and 0.118 using Eq. (2), imposing specific cuts on $Q^2 \gtrsim 3.5, 7.5, 10 \text{ GeV}^2$, and found that fit results are almost independent on the strong coupling constant. To use as much as data, we set a minimum $Q_{\min}^2 = 3.5 \text{ GeV}^2$ as a universal cut on the squared 4-momentum transferred. The quality of the fit is judged by $\chi^2/(\text{degrees of freedom (DOF)})$, defined as the difference of the number of data points used between the number of free fit parameters), where χ^2 is constructed by including the measurement, corresponding theoretical calculation, and the measurement uncertainties (statistical, correlated and uncorrelated systematic). The detailed form of the χ^2 function is constructed as

$$\chi^2(M, \lambda) = \sum_{i=1}^{N_{\text{data}}} \frac{1}{\Delta_i^2} \left(M_i + \sum_{\alpha=1}^{N_{\text{sys.corr}}} \Gamma_{\alpha i} T_i \lambda_\alpha - T_i \right)^2 + \sum_{\alpha=1}^{N_{\text{sys.corr}}} \lambda_\alpha + \sum_{i=1}^{N_{\text{data}}} \ln \frac{\Delta_i^2}{M_i^2 (\delta_{i,\text{stat}}^2 + \delta_{i,\text{uncor}}^2)}, \quad (5)$$

where M_i and T_i stand for the measured and theoretical values of the cross sections. $\Gamma_{\alpha i}$, together with the nuisance parameter λ_α , quantifies the contribution of each correlated systematic error source α . $\delta_{i,\text{stat}}$ and $\delta_{i,\text{uncor}}$ are the corresponding relative statistical and uncorrelated systematical errors which are proportional to the measured value.

TABLE I. The corresponding partial, total χ^2/DOF function for four fits.

Dataset	Fit 1	Fit 2	Fit 1p	PKU-fit
HERA I + II CC e^+p	49/39	49/39	47/39	44/39
HERA I + II CC e^-p	64/42	55/42	66/42	68/42
HERA I + II NC e^+p	226/159	222/159	226/159	226/159
HERA I + II NC e^+p , $E_p = 820 \text{ GeV}$	66/70	70/70	66/70	66/70
HERA I + II NC e^+p , $E_p = 920 \text{ GeV}$	451/377	449/377	450/377	441/377
HERA I + II NC e^+p , $E_p = 460 \text{ GeV}$	220/204	219/204	221/204	220/204
HERA I + II NC e^+p , $E_p = 575 \text{ GeV}$	222/254	224/254	222/254	220/254
CMS W - μ charge asymmetry $\sqrt{s} = 7 \text{ TeV}$		14/11		13/11
CMS W - μ charge asymmetry $\sqrt{s} = 8 \text{ TeV}$		3.2/11		5.5/11
CMS W + charm $\sqrt{s} = 7 \text{ TeV}$		2.6/5		1.5/5
CMS W + charm $\sqrt{s} = 13 \text{ TeV}$		2.5/5		1.2/5
ATLAS Z/γ^* , $46 < m_z < 66 \text{ GeV}$	23/6		23/6	24/6
ATLAS central Z/γ^* , $66 < m_z < 116 \text{ GeV}$	14/12		15/12	14/12
ATLAS forward Z/γ^* , $66 < m_z < 116 \text{ GeV}$	5.6/9		6/9	4.7/9
ATLAS central Z/γ^* , $116 < m_z < 150 \text{ GeV}$	6.9/6		6.7/6	6.8/6
ATLAS forward Z/γ^* , $116 < m_z < 150 \text{ GeV}$	4.5/6		4.4/6	4.1/6
ATLAS $W^- \rightarrow l^-\bar{\nu}$	8.7/11		10/11	9.3/11
ATLAS $W^+ \rightarrow l^+\nu$	13/11		13/11	14/11
ATLAS $W^+ + \bar{c}$ -jet $\sqrt{s} = 7 \text{ TeV}$			7.5/11	7.1/11
ATLAS $W^- + c$ -jet $\sqrt{s} = 7 \text{ TeV}$			1.6/11	1.8/11
Correlated χ^2	117	90	120	118
Log penalty χ^2	-14.18	-9.35	-14.13	-6.14
Total χ^2/DOF	1476/1189	1391/1161	1491/1211	1503/1239

The index i runs over all N_{data} data points, while index α runs all of the correlated error sources. Because of the multiplicative nature of both the statistical and uncorrelated systematical errors, Δ_i^2 is defined as

$$\Delta_i^2 = \delta_{i,\text{stat}}^2 T_i M_i + \delta_{i,\text{uncor}}^2 M_i^2, \quad (6)$$

which is corrected for statistical fluctuations in data by scaling with predicted values according to the recipe in Ref. [45]. This form of Δ_i^2 leads to a logarithmic term arising from likelihood transition of χ^2 . The minimization of χ^2 function against the fit parameters and the nuisance parameter is done using the package MINUIT [46], which is interfaced with XFITTER. The PDFs uncertainties arising from the uncertainties in the measurements are estimated using Hessian method [47] adopting the same tolerance $\Delta\chi^2 = 1$ as the HERA, ATLAS, and CMS groups, corresponding to 68% confidence level (CL).

IV. RESULTS

Table I lists the corresponding χ^2/DOF values for Fit 1, Fit 2, and Fit 1p plus PKU-fit showing the datasets are reasonably agreeable.

In the first place, we show the resulted strange quark PDFs and the ratio R_s , with ratio to Fit 1, extracted from ATLAS W/Z and CMS W + charm data in Fig. 1 applying Eq. (2) (Fit 1) and Eq. (3) (Fit 2) at the factorization scale $Q^2 = 1.9 \text{ GeV}^2$ and the scale at which the W boson is produced. The $x_s(x, Q^2)$ distributions display overall agreement within the uncertainties at both scales. However, it is clear to see that they differ distinctly from each other toward large x . Similarly, the R_s distributions on the bottom pictures of Fig. 1 follow the same trend that the strange quark PDF has, thus indicating that the ATLAS W/Z and CMS W + charm data have some tensions at large momentum fraction. The central distribution of strangeness from the fit of ATLAS W/Z data, Fit 1, whether in terms of parton level or the ratio R_s , is slightly

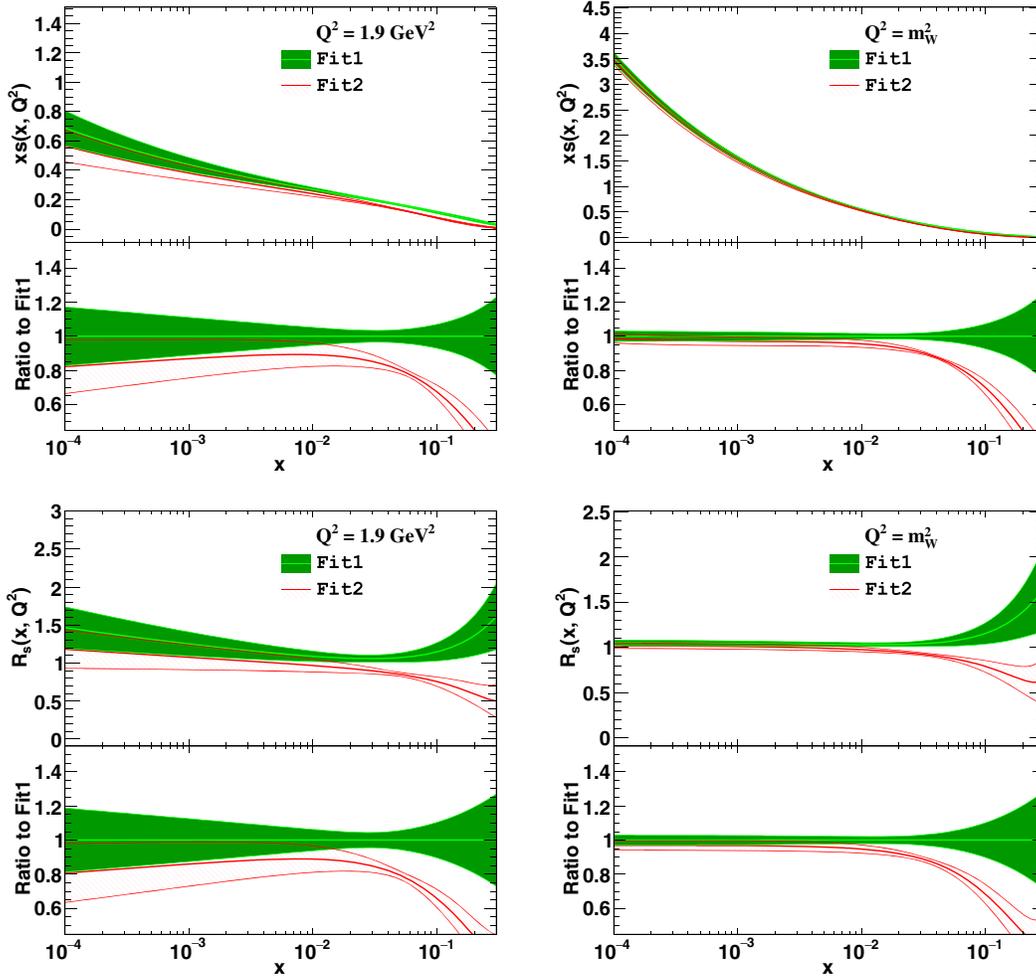


FIG. 1. The strange PDFs and strange-to-nonstrange ratio as functions of x evaluated at the factorization scale $Q^2 = 1.9 \text{ GeV}^2$ and $Q^2 = m_W^2$ from Fit 1 and Fit 2. The upper panel shows the central distribution with error bands corresponding to 68% CL, while the bottom panel shows the ratio to Fit 1.

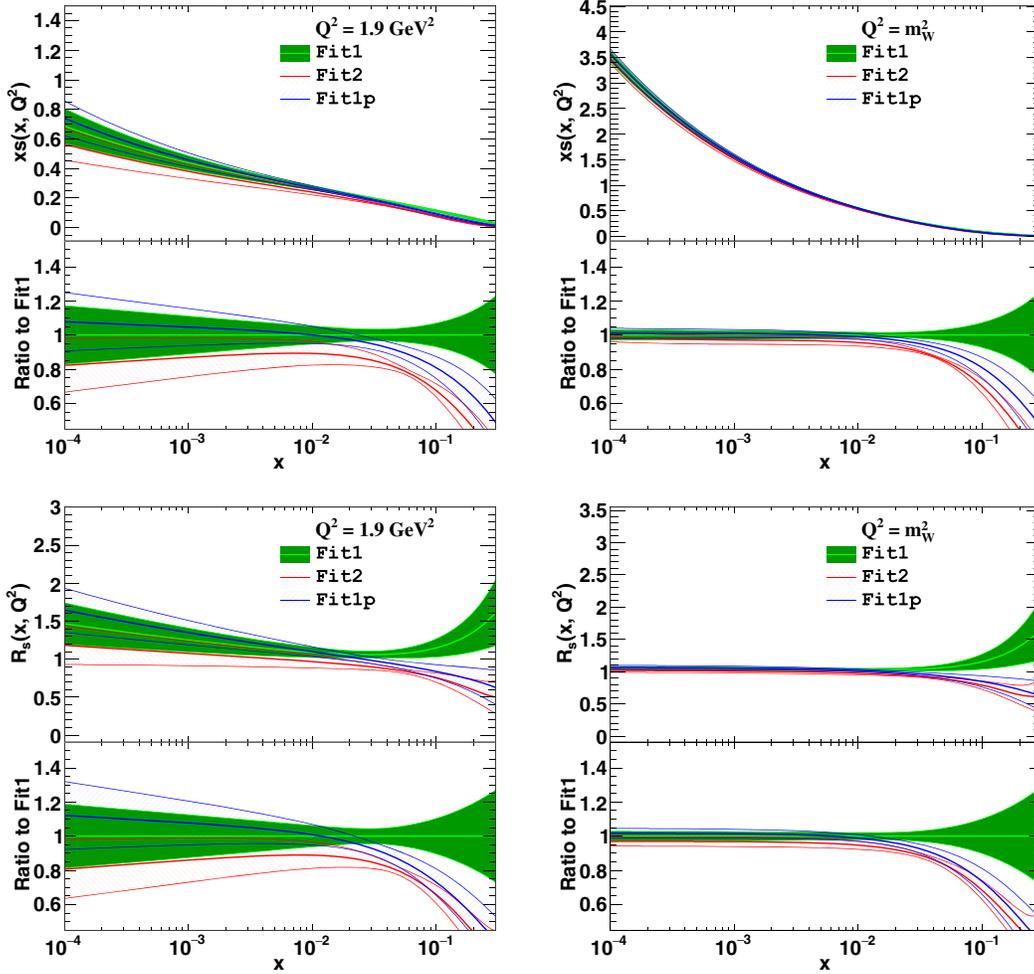


FIG. 2. The strange PDFs and strange-to-nonstrange ratio as functions of x evaluated at the factorization scale $Q^2 = 1.9 \text{ GeV}^2$ and $Q^2 = m_W^2$ from Fit 1p (blue hatched band) compared to Fit 1 and Fit 2. The upper panel shows the central distribution with error bands at 68% CL, while the bottom panel shows the ratio to Fit 1.

higher than the one obtained in Fit 2. The ratio R_s indeed has a significant x -dependent property that coincides with the results obtained by CMS analysis of $W + \text{charm}$ data [23]. The incompatibility between strange quarks is around 3σ .

So, the reason that gives rise to the different interpretations of strangeness in terms of the ratio R_s between the original ATLAS and CMS analyses lies in the different data structure or physics processes, and, possibly, the somewhat different parametrization forms with flexible parameters of PDFs also play some role; if the sea quarks are related with each other by setting $A_{\bar{u}} = A_{\bar{d}}$ and $B_{\bar{u}} = B_{\bar{d}} = B_{\bar{s}}$, they shape out the ATLAS-type R_s as the ATLAS Collaboration obtained in Refs. [15,17]. On the other hand, they produce the similar CMS-type of R_s if the parameters are set free. As for the disparity on this ratio arising toward large momentum fraction in Fit 1 and Fit 2, this is mainly because Fit 1 does not include the $W + \text{charm}$ production data that have more significance than W/Z process when it comes to the strange quark PDFs.

In the second place, the resulted strange quark PDFs and the ratio R_s of Fit 1p are depicted in Fig. 2 in comparison with Fit 1 and Fit 2. After the inclusion of ATLAS $W + \text{charm}$ data, the $xs(x, Q^2)$ distribution of Fit 1p, depicted with a blue hatched band, demonstrates decent agreement with Fit 1 and Fit 2 at $x \lesssim 10^{-2}$, while being compatible with $xs(x, Q^2)$ of Fit 2 toward large x . The R_s of Fit 1p, shown on the bottom, also manifests compatibility with that of Fit 2 with a sharp deviation from that of Fit 1 at $x \gtrsim 10^{-2}$. This compatibility is highly because the $W + \text{charm}$ data exert more constraints on the strange quark PDFs in Fit 1. Based on these results, we can safely state that the observed disparity in the interpretations of strangeness by the ATLAS and CMS groups can be ascribed mainly to the different usage of data from the different physical processes. The outcome would be consistent if this factor were handled properly with more flexible parametrization forms.

The combined usage of the inclusive W/Z measurements with the $W + \text{charm}$ data and W -lepton charge asymmetry provides a better sensitivity on the strangeness

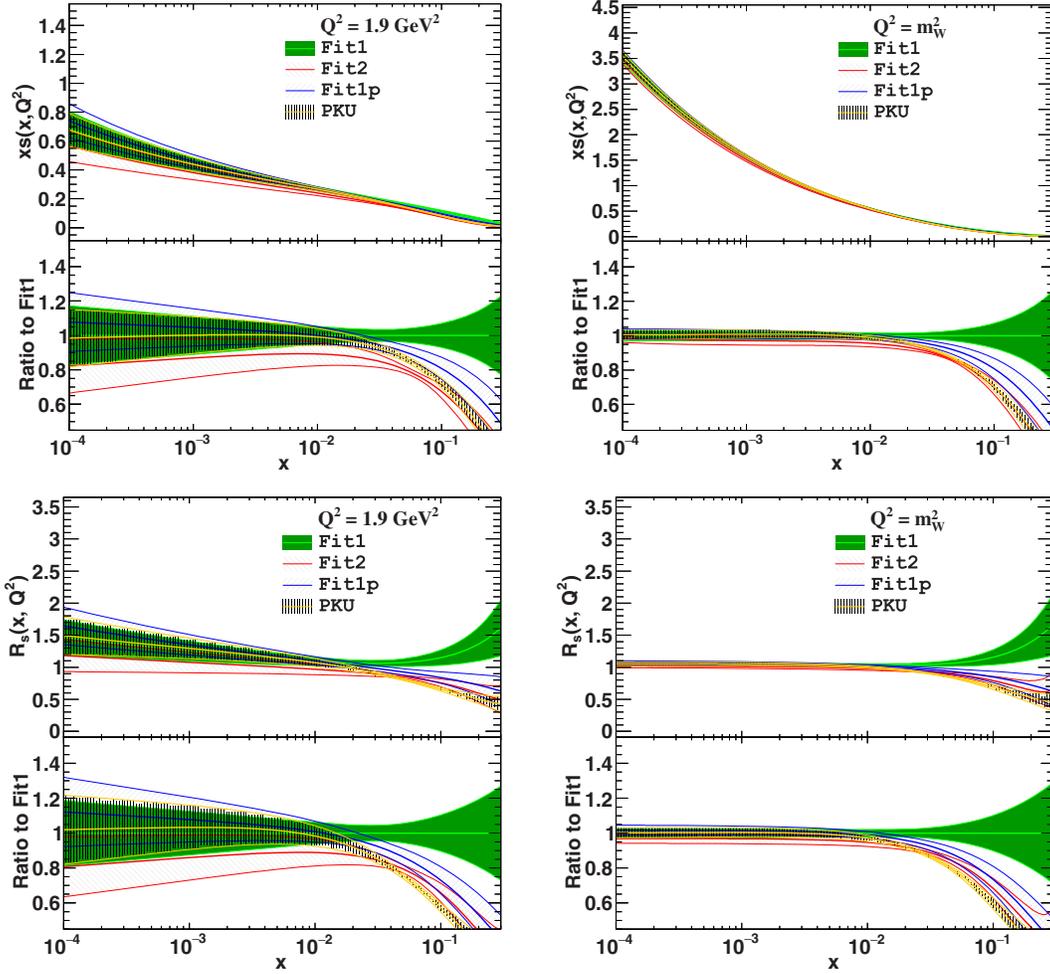


FIG. 3. The $xs(x, Q^2)$ and R_s distributions evaluated at $Q^2 = 4 \text{ GeV}^2$ as functions of x for PKU-fit in comparison with the previous results. The bands represent 68% CL of experimental uncertainties.

than used separately. Having clarified the discrepancy and its origin between ATLAS and CMS interpretations of the strange quark PDFs, a fairly exact distribution of the strange quark PDFs obtained through the PKU-fit is shown in Fig. 3 in comparison with Fit 1, Fit 2, and Fit 1p at the initial and W mass scale. It is clear that the strange quark PDFs of PKU-fit are still in accordance with that of Fit 2 and Fit 1p throughout the entire x while agreeing with Fit 1 only at $x \lesssim 10^{-2}$ and deviating from it toward large x . As Fit 1p and Fit 2 display, the PKU-fit also manifests a smoothly x -dependent behavior of R_s at $Q^2 = 1.9 \text{ GeV}^2$, which is declining faster than other fits toward large x . The ratio R_s distribution at $Q^2 = m_W^2$ tends to unity at $x \lesssim 0.02$, suggesting $(s + \bar{s}) = (\bar{u} + \bar{d})$. Generally, we can confirm that the strange quark PDFs, from both the ATLAS W/Z and CMS $W + \text{charm}$ data, are fairly consistent within uncertainties in smaller momentum fraction regions and manifest a tension toward large momentum fraction. This tension fades away when the same physical processes are used with flexible parameterization forms. The strange

quark density of the proton is only suppressed at $x > 0.023$ and enhanced toward smaller x .

As confirmed, the inclusion of $W + \text{charm}$ data together with W/Z from both the ATLAS and CMS experiments produce well-constrained strangeness. We calculate the cross section of $W + \text{charm}$ production of both the ATLAS and CMS experiments with the extracted PDF sets to see how much the strange quark PDFs from PKU-fit improve the theoretical results. This is pictured in Figs. 4 and 5 compared with the measurements. The predictions of both the ATLAS and CMS $W + \text{charm}$ measurements agree with the measurements quite well. However, it is worth noting that the calculation obtained from Fit 1 is apparently bigger than the measurements at $|\eta| \approx 2$. The deviations at the large pseudorapidity have a relation with strange quark density at the corresponding momentum fraction x . This means that the deviation appears when there is an increase in the strange quark density. Understandably, it is the large- x distribution of the strange quark that leads to the deviation. This tells us that the

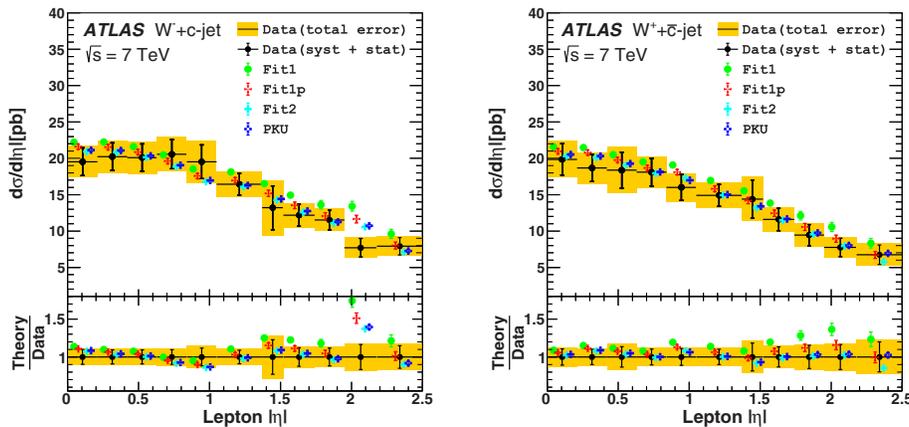


FIG. 4. The measured differential cross section of ATLAS $W +$ charm production as a function of lepton $|\eta|$ compared to predictions computed with the resulted PDF sets: $W^- + c$ -jet (left panel) and $W^+ + \bar{c}$ -jet (right panel). The measurements are shown with black error bars for systematic + statistical uncertainty and with filled bands for total uncertainty (including uncorrelated uncertainty). The predictions are obtained with PDF uncertainties at 68% CL.

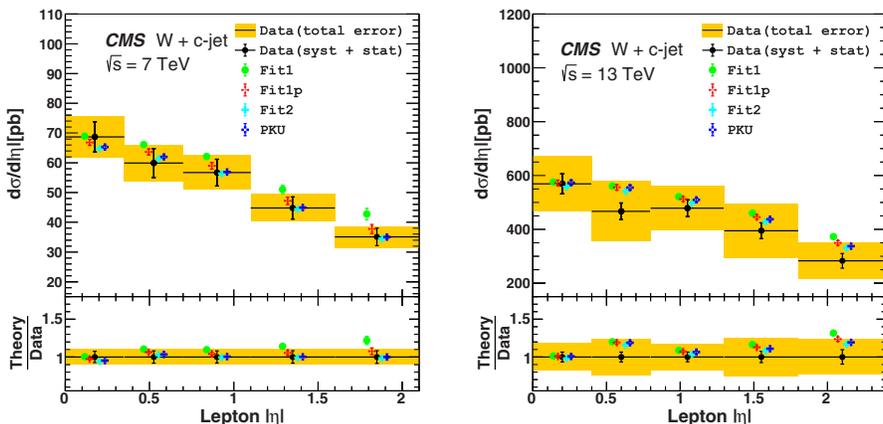


FIG. 5. The measured differential cross section of CMS $W +$ charm production as a function of lepton $|\eta|$ compared to predictions computed with the resulted PDF sets: $\sqrt{s} = 7$ TeV (left panel) and $\sqrt{s} = 13$ TeV (right panel). The measurements are shown with black error bars for systematic + statistical uncertainty and with filled bands for total uncertainty (including correlated uncertainty). The predictions are obtained with PDF uncertainties at 68% CL.

uncertainty of the strange quark density in Fit 1 by fitting only the ATLAS W/Z data is large at large momentum fraction.

V. CONCLUSIONS

The shapes of the strange quark PDFs of the proton extracted from LHC data were interpreted in different ways: the ATLAS statement of fully enhanced strangeness and the CMS statement of the suppressed one which supports the same idea obtained from fixed target neutrino-nuclear collision experiments. In this paper, through four rounds of analyses of the HERA I + II $e^\pm p$, LHC W/Z , $W +$ charm production, and W -lepton charge asymmetry data, i.e., Fit 1, Fit 2, Fit 1p, and PKU-fit, we have investigated the exact distribution of the strange quark and its ratio to nonstrange sea quarks implementing the ATLAS

and CMS analyses setup with free sea quark parameters using the XFITTER program and resolved tensions between ATLAS and CMS data and finally extracted fairly well-shaped strange quark densities. At first, both the ATLAS W/Z and CMS $W +$ charm data are proved, through Fit 1 and Fit 2 fits, to be partially consistent with the relaxed sea quark parameters in terms of strange quark and its ratio to nonstrange sea quark distributions at $x \lesssim 10^{-2}$. The inconsistency toward large x between strange quark density and the ratio R_s from the ATLAS and CMS collaborations suggests that there is tension between ATLAS 7 TeV W/Z and CMS 7 and 13 TeV $W +$ charm data. It is shown in Fit 1p that this tension can be eliminated by fitting both the W/Z and $W +$ charm data simultaneously. The shapes of the strange quark PDFs and the ratio obtained in our analysis are purely data driven. There is no handmade

constraint on the fit parameters. The $W + \text{charm}$ cross section is calculated with the obtained PDFs, and deviations from measurement at the large lepton $|\eta|$ are observed. The PDFs obtained from only the HERA I + II $e^\pm p$ and ATLAS W/Z measurement produces larger cross section values than the $W + \text{charm}$ data at the large lepton $|\eta|$, suggesting a large uncertainty of the proton strangeness from fitting only the ATLAS 7 TeV W/Z data at large x . When the HERA I + II $e^\pm p$, ATLAS W/Z , CMS $W + \text{charm}$, and W -lepton data are used in combination, further strengthened by ATLAS $W + \text{charm}$ measurement, with a more flexible parametrization form being applied, analyzed in PKU-fit, the resulted distribution of the strange quark density gets improved, and the ratio R_s shows more

x -dependent property. The strangeness is gradually suppressed at larger x , but it is highly enhanced toward smaller x with its magnitude becoming comparable to that of the light-flavor up and down sea quarks.

ACKNOWLEDGMENTS

We thank the XFITTER team for providing useful help and discussions on the usage of their awesome program. We are also so grateful to Katerina Lipka from DESY for providing the corresponding grid file of CMS $W + \text{charm}$ production at $\sqrt{s} = 13$ TeV for the corresponding theoretical calculation. This work is supported by the National Natural Science Foundation of China (Grant No. 12075003).

-
- [1] E. D. Bloom, D. H. Coward, H. C. DeStaabler, J. Drees, G. Miller, L. W. Mo, R. E. Taylor, M. Breidenbach, J. I. Friedman, G. C. Hartmann, and H. W. Kendall, *Phys. Rev. Lett.* **23**, 930 (1969).
 - [2] M. Breidenbach, J. I. Friedman, H. W. Kendall, E. D. Bloom, D. H. Coward, H. C. DeStaabler, J. Drees, L. W. Mo, and R. E. Taylor, *Phys. Rev. Lett.* **23**, 935 (1969).
 - [3] L. Del Debbio *et al.* (NNPDF Collaboration), *J. High Energy Phys.* **03** (2007) 039.
 - [4] S. Dulat, T. J. Hou, J. Gao, M. Guzzi, J. Huston, P. Nadolsky, J. Pumplin, C. Schmidt, D. Stump, and C. P. Yuan, *Phys. Rev. D* **93**, 033006 (2016).
 - [5] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne, *Eur. Phys. J. C* **75**, 204 (2015).
 - [6] S. Alekhin, J. Blumlein, S. Klein, and S. Moch, *Phys. Rev. D* **81**, 014032 (2010).
 - [7] A. M. Cooper-Sarkar (ZEUS and H1 Collaborations), *Proc. Sci.*, EPS-HEP2011 (2011) 320 [arXiv:1112.2107].
 - [8] U. Baur, F. Halzen, S. Keller, M. L. Mangano, and K. Riesselmann, *Phys. Lett. B* **318**, 544 (1993).
 - [9] E. Farhi and R. L. Jaffe, *Phys. Rev. D* **30**, 2379 (1984).
 - [10] M. Goncharov *et al.* (NuTeV Collaboration), *Phys. Rev. D* **64**, 112006 (2001).
 - [11] D. Mason *et al.* (NuTeV Collaboration), *Phys. Rev. Lett.* **99**, 192001 (2007).
 - [12] O. Samoylov *et al.* (NOMAD Collaboration), *Nucl. Phys.* **B876**, 339 (2013).
 - [13] G. Aad *et al.* (ATLAS Collaboration), *Phys. Rev. Lett.* **109**, 012001 (2012).
 - [14] H. Abramowicz *et al.* (H1 and ZEUS Collaborations), *Eur. Phys. J. C* **75**, 580 (2015).
 - [15] M. Aaboud *et al.* (ATLAS Collaboration), *Eur. Phys. J. C* **77**, 367 (2017).
 - [16] A. M. Cooper-Sarkar and K. Wichmann, *Phys. Rev. D* **98**, 014027 (2018).
 - [17] G. Aad *et al.* (ATLAS Collaboration), *J. High Energy Phys.* **07** (2021) 223..
 - [18] S. Bailey, T. Cridge, L. A. Harland-Lang, A. D. Martin, and R. S. Thorne, *Eur. Phys. J. C* **81**, 341 (2021).
 - [19] G. Aad *et al.* (ATLAS Collaboration), *J. High Energy Phys.* **05** (2014) 068.
 - [20] S. Chatrchyan *et al.* (CMS Collaboration), *Phys. Rev. D* **90**, 032004 (2014).
 - [21] V. Khachatryan *et al.* (CMS Collaboration), *Eur. Phys. J. C* **76**, 469 (2016).
 - [22] S. Chatrchyan *et al.* (CMS Collaboration), *J. High Energy Phys.* **02** (2014) 013.
 - [23] A. M. Sirunyan *et al.* (CMS Collaboration), *Eur. Phys. J. C* **79**, 269 (2019).
 - [24] N. Yalkun and S. Dulat, *Chin. Phys. C* **43**, 123101 (2019).
 - [25] C. Schmidt, J. Pumplin, C. P. Yuan, and P. Yuan, *Phys. Rev. D* **98**, 094005 (2018).
 - [26] T. J. Hou, J. Gao, T. J. Hobbs, K. Xie, S. Dulat, M. Guzzi, J. Huston, P. Nadolsky, J. Pumplin, C. Schmidt *et al.*, *Phys. Rev. D* **103**, 014013 (2021).
 - [27] F. Faura, S. Iranipour, E. R. Nocera, J. Rojo, and M. Ubiali, *Eur. Phys. J. C* **80**, 1168 (2020).
 - [28] S. Alekhin *et al.* (HERAFitter Collaboration), *Eur. Phys. J. C* **75**, 304 (2015).
 - [29] C. Gong and B.-Q. Ma, *Phys. Rev. C* **97**, 065207 (2018).
 - [30] S. J. Brodsky, I. Schmidt, and J. J. Yang, *Phys. Rev. D* **70**, 116003 (2004).
 - [31] V. N. Gribov and L. N. Lipatov, *Sov. J. Nucl. Phys.* **15**, 438 (1972).
 - [32] V. N. Gribov and L. N. Lipatov, *Sov. J. Nucl. Phys.* **15**, 675 (1972).
 - [33] Y. L. Dokshitzer, *Sov. Phys. JETP* **46**, 641 (1977).
 - [34] G. Altarelli and G. Parisi, *Nucl. Phys.* **B126**, 298 (1977).
 - [35] S. Fanchiotti, B. A. Kniehl, and A. Sirlin, *Phys. Rev. D* **48**, 307 (1993).
 - [36] M. Botje, *Comput. Phys. Commun.* **182**, 490 (2011).
 - [37] R. S. Thorne and R. G. Roberts, *Phys. Rev. D* **57**, 6871 (1998).
 - [38] R. S. Thorne, *Phys. Rev. D* **73**, 054019 (2006).
 - [39] R. S. Thorne, *Phys. Rev. D* **86**, 074017 (2012).

- [40] J. M. Campbell and R. K. Ellis, *Phys. Rev. D* **60**, 113006 (1999).
- [41] J. M. Campbell, R. K. Ellis, and C. Williams, *J. High Energy Phys.* **07** (2011) 018.
- [42] T. Carli, D. Clements, A. Cooper-Sarkar, C. Gwenlan, G. P. Salam, F. Siegert, P. Starovoitov, and M. Sutton, *Eur. Phys. J. C* **66**, 503 (2010).
- [43] S. Alekhin, J. Blumlein, and S. Moch, *Eur. Phys. J. C* **78**, 477 (2018).
- [44] S. J. Brodsky and B.-Q. Ma, *Phys. Lett. B* **381**, 317 (1996).
- [45] F. D. Aaron *et al.* (H1 and ZEUS Collaborations), *J. High Energy Phys.* **01** (2010) 109.
- [46] F. James and M. Roos, *Comput. Phys. Commun.* **10**, 343 (1975).
- [47] J. Pumplin, D. Stump, R. Brock, D. Casey, J. Huston, J. Kalk, H. L. Lai, and W. K. Tung, *Phys. Rev. D* **65**, 014013 (2001).