Impact of LHCb 13 TeV W and Z pseudo-data on the Parton Distribution Functions

Qilin Deng, Qundong Han, and Hang Yin*

Institute of Particle Physics,
Central China Normal University,
Wuhan, Hubei 430079 China

Sayipjamal Dulat†

School of Physics Science and Technology,
Xinjiang University, Urumqi, Xinjiang 830046 China

Tie-Jiun Hou‡

Department of Physics, College of Sciences,
Northeastern University, Shenyang, Liaoning 110819 China

C.–P. Yuan§

Department of Physics and Astronomy,
Michigan State University, East Lansing, MI 48824 U.S.A.

(Dated: September 8, 2020)

Abstract

We studied the potential of the LHCb 13 TeV single $W^\pm$ and $Z$ boson pseudo-data on constraining the Parton Distribution Functions (PDFs) of the proton. As an example, we demonstrated the sensitivity of the LHCb 13 TeV data, collected with an integrated luminosity of 5 fb$^{-1}$ and 300 fb$^{-1}$, respectively, to reducing the PDF uncertainty bands of the CT14HERA2 PDFs, using the error PDF updating package ePUMP. For that, the sensitivities of various experimental observables have been compared. Generally, sizable reductions in PDF uncertainties can be observed in the 300 fb$^{-1}$ data sample, particularly in the small $x$ region. The double differential cross section measurement on $Z$ boson $p_T$ and rapidity can greatly reduce the uncertainty bands of $u$ and $d$ quarks in almost all $x$ range, as compared to various single observable measurements.

* yinh@ccnu.edu.cn
CONTENTS

I. Introduction 4

II. Updating Error PDFs 6

III. Impact of the LHCb 13 TeV $W^\pm$ and $Z$ pseudo-data on CT14HERA2 PDFs 8
   A. Update from LHCb 13 TeV $W^\pm$ pseudo-data 8
   B. Update from LHCb 13 TeV $Z$ pseudo-data 10
   C. Update from LHCb 13 TeV $W^+ + W^- + Z$ pseudo-data 13
   D. Update from LHCb 13 TeV $(W^+ + W^-)/Z$ pseudo-data 16

IV. Tolerance criteria in ePump update 19

V. Conclusion 20

VI. Acknowledgments 21

References 21
I. INTRODUCTION

In the hadron colliders, most of physics analyses are highly rely on the understanding of the parton picture of hadronic beam particle, like the precision measurements of the Standard Model (SM) parameters [1–3], and new physics searches. The parton picture follows the factorization theorem of Quantum Chromodynamics (QCD). The parton distribution functions (PDFs) are nonperturbative, therefore cannot be calculated. They are function of the Bjorken-$x$ values ($x$, momentum fraction) of partons at a momentum transfer scale ($Q$), which are determined phenomenologically by a global analysis of experimental data from a wide range of physics processes, such as Deep Inelastic Scattering (DIS), Drell-Yan (DY), inclusive jets, and top quark pair production processes. The PDF dependence on $Q$ are determined by the renormalization-group based evolution equations, i.e., DGLAP equation [4–6].

Precision measurements of the single $W^\pm$ and $Z$ gauge boson production cross section \(^1\) at the CERN Large Hadron Collider (LHC) provide important tests on the QCD and the electroweak (EW) sectors of the SM. Theoretical predictions for these cross sections are available up to next-to-next-to-leading order (NNLO) in perturbative QCD [7–11], where one of the dominated systematical uncertainties comes from the PDFs. The CT14 PDFs [12] are the first CTEQ-TEA PDFs that include published results from the ATLAS, CMS, and LHCb collaborations at 7 TeV, including the $W^\pm$ and $Z$ gauge boson production cross sections and the lepton charge asymmetry measurements from the ATLAS Collaboration [13], the lepton charge asymmetry in the electron [14] and muon channels [15] from the CMS Collaboration, and the lepton charge asymmetry in the decay of $W^\pm$-bosons to an electron or a muon, and the $Z$ boson rapidity distribution from the LHCb Collaboration [16]. The ATLAS and CMS measurements primarily impose constraints on the light quark and antiquark PDFs at $x \gtrsim 0.01$. As studied in Refs. [17, 18], the LHCb 7 TeV and 8 TeV $W^\pm$ and $Z$ boson measurements, though with larger statistical uncertainties as compared to the corresponding results from the ATLAS and CMS experiments, could also impose significant constraints on $u$ and $d$ PDFs.

In the past decades, a large number of experimental results are used in the PDF global analysis, but we still have limited knowledge on the PDFs in very small and very large $x$

\(^1\) Throughout this paper, $Z$ includes both the Z boson and the virtual photon contributions.
ranges. In the single $W^\pm$ and $Z$ boson production, the $x$ value of interacting partons ($x_1$ and $x_2$) are correlated with the boson production, through its rapidity ($y$), as $y = \frac{1}{2} \ln \frac{x_1}{x_2}$. Therefore, the single $W^\pm$ and $Z$ data in the forward detector region is valuable in the PDF global analysis, as events with larger boson rapidity are produced by partons with small or large $x$. Correlations between predicted LHCb 13 TeV $Z$ boson production cross section and $u, d$-quark PDFs as a function of Bjorken-$x$ are shown in Fig. 1. As shown in the figure, the LHCb 13 TeV data is expected to have strong correlations with $u$-, $d$-quarks in the small $x$ region, indicate the LHCb 13 TeV $W^\pm$ and $Z$ data can be used to constrain the corresponding PDFs.

![Diagram](attachment:CT14HERA2_u(x,100GeV) and Z rapidity.png)

**FIG. 1:** Correlation $\cos \phi$ [19] between $u(x,Q)$-PDF (left), or $d(x,Q)$-PDF (right), and the DY differential cross-section in 18 bins of $Z$ boson rapidity, as predicted by ResBos with CT14HERA2 PDFs, at $Q = 100$ GeV. Same binning scheme as the LHCb paper [16] is used.

The LHCb detector [20, 21] is a single-arm forward spectrometer designed for the study of particles containing $b$ or $c$ quarks, covering the pseudorapidity range $2 < \eta < 5$. With a high performance tracking system and a muon sub-detector, the LHCb data can be also extended to the precision EW measurements. Using $pp$ collision data collected at $\sqrt{s} = 7$ TeV, the LHCb measured the single $W^\pm$ and $Z$ boson production cross sections using both muon and electron channel [16, 22, 23], and the same measurements had been performed using $\sqrt{s} = 8$ TeV data [24–26]. These results have been used to constrain the PDFs [12, 27, 28], bring valuable information on the PDF analysis. The $\sqrt{s} = 13$ TeV $pp$ collision data sample is collected with a larger center-of-mass energy than previous publications, with more $W^\pm$ and $Z$ boson events boosted to the forward region, therefore could access even
smaller (larger) values of $x$ than the previous 7 and 8 TeV results.

The LHCb detector was running in a reduced luminosities compared to the ATLAS and CMS detectors [29], the reason is that the detector occupancy is extremely high in the forward region. In the LHC Run 2 period (2015-2018), the LHCb detector collected more than 5 fb$^{-1}$ pp collision data [30] at $\sqrt{s} = 13$ TeV. Then, an upgraded detector [31] is foreseen to allow the LHCb detector operation at a luminosity $2 \times 10^{33}$ cm$^{-2}$s$^{-1}$ in the LHC Run 3 period (2021-2023), which is five time larger instantaneous luminosity compared to previous. By the end of the LHC Run 4 period (2026-2029), the LHCb detector is expect to collect approximate 50 fb$^{-1}$ pp collision data [32]. In the future, there is a so-called the LHCb Upgrade-II phase (planned for 2031 data taking) [33], to guarantee the LHCb detector could run in an even higher luminosity ($2 \times 10^{34}$ cm$^{-2}$s$^{-1}$) [34]. After the LHCb detector Upgrade-II, by the end of the LHC operation, the LHCb detector is planning to collect a data sample corresponding to a minimum 300 fb$^{-1}$ [32]. Therefore, in this article, the pseudo-data samples used for the physics projections are set to either 5 fb$^{-1}$ or 300 fb$^{-1}$.

The article is organized as follows, in Sec. II, we discuss the error PDF updating package ePump and pseudo-data samples used in the analysis. In Sec. III, we study the impacts of the LHCb 13 TeV single $W^{\pm}$ and $Z$ boson pseudo-data on CT14HERA2 PDFs [12, 35]. In Sec. IV, we discuss the choice of tolerance criteria in ePump update. Our conclusion is given in Sec. V.

II. Updating Error PDFs

Recently, the CTEQ-TEA global analysis group has released a tool named ePump [36, 37] (short for the Error PDF Updating Method Package), which could be used to explore the impact of new data on the existing PDFs, without perform a full global analysis. The ePump method has been demonstrated in Ref. [38]. In the ePump update, two inputs are needed: a measured result from experiment and the corresponding theoretical predictions of the complete set of error PDFs. Thus far, the LHCb Collaboration only published the single $Z$ boson production cross section [39] at 13 TeV, using a small fraction of its Run 2 data sample (2015 data, with an integrated luminosity of 0.3 fb$^{-1}$). The comparisons between the LHCb 13 TeV data and ResBos prediction are shown in Fig. 2, where good agreements between data and prediction had been found. On the other hand, there is no publication
for the single $W^\pm$ boson production using the 13 TeV LHCb data. Therefore, we shall use pseudo-data in this analysis to emulate the impact of the upcoming LHCb 13 TeV data on the PDFs.

The Monte Carlo events generated with the ResBos generator [40], using the MMHT14 [27] and CT14HERA2 [12, 35] PDFs, are taken as the pseudo-data and theoretical predictions, respectively, in this work. The theoretical predictions in this work are computed using the ResBos [40] package at approximate NNLO plus next-to-next-to-leading-logarithm (NNLL) in QCD interaction, in which the canonical scales are used [41, 42]. For example, both the renormalization and factorization scales are set to be the invariant mass of the lepton pair in the Drell-Yan (DY) events.

To emulate the LHCb detector acceptance, the charged leptons ($\ell^\pm$, electrons or muons) are required to have transverse momentum ($p_T$) greater than 20 GeV/c and pseudorapidity ($\eta$) in the range of $2.0 < \eta < 4.5$. In case of the $Z$ boson events, the invariant mass of the dilepton pair is required to be in the range from 60 GeV/$c^2$ to 120 GeV/$c^2$.

![Figure 2: Comparison of DY differential cross section as a function of the $Z$ boson rapidity (left) and $p_T$ (right), between theory (ResBos) and the LHCb 13 TeV data [39]. The blue points represent the LHCb results in muon channel, and the red points in electron channel, while the black line represents the ResBos prediction.](image)

In the ePUMP study, for the $Z \rightarrow \ell^+\ell^-$ pseudo-data input, the statistical uncertainties are scaled to the 5 fb$^{-1}$ and 300 fb$^{-1}$ data sample, separately, by extrapolating the total uncertainty of the LHCb 13 TeV publication [39] to the pseudo-data sample. In the extrapolation, an assumption is made that the ratio of statistical uncertainty to systematical uncertainty will remain the same in all data samples. Similarly, for the $W^\pm \rightarrow \ell^\pm \nu$ pseudo-data sample,
uncertainties are estimated using the LHCb 8 TeV publication [25, 26].

III. IMPACT OF THE LHCb 13 TEV $W^{\pm}$ AND $Z$ PSEUDO-DATA ON CT14HERA2 PDFS

In this section, we study the impact of the LHCb 13 TeV single $W^{\pm}$ and $Z$ boson pseudo-data on the CT14HERA2 PDFs, to demonstrate the LHCb 13 TeV data sensitivity, and to further investigate valuable observable for future measurements.

Since the pseudo-data sample generated with MMHT14 PDFs is used to update the CT14HERA2 PDF sets, and there are differences between the central PDF set of MMHT14 and CT14HERA2, the central value of ePump updated PDFs is varied from the CT14HERA2 central set. In this article, we are interested in variations of PDF uncertainty, and thus we do not discuss variations of PDF central values hereafter.

A. Update from LHCb 13 TeV $W^{\pm}$ pseudo-data

In the $W^{\pm}$ boson leptonic decay, there is a neutrino and a charged lepton in the final state, where the neutrino will escape from detector, only the charged lepton can be detected in a hadron collider experiment. This feature makes a $W^{\pm}$ boson analysis become complicated, since irreducible background contribution is difficult to be modeled. On the other hand, the single $W^{\pm}$ production rate is one order of magnitude larger than that of the $Z$ boson at the LHCb. If we could model the background properly for the $W^{\pm}$ events, such sample with large statistics could allow us to perform many precision measurements. In this study, we used the charged lepton pseudorapidity distribution ($\eta$) as an observable in the ePump update, with same binning scheme as the previous LHCb publications [25, 26].

After the ePump update, the updated quark PDF distribution is compared with the default one of CT14HERA2. The $d$ quark PDF distribution with its uncertainty is shown in Fig. 3, which indicates the LHCb 13 TeV $W^{\pm}$ boson data have large impact on $d$ quark PDF, especially in the small $x$ range from $10^{-5}$ to $10^{-3}$. With a 300 fb$^{-1}$ LHCb 13 TeV $W^{\pm}$ data sample, the $d$ quark PDF uncertainty can be reduced by a factor of 30% around $x = 10^{-3}$. While the LHCb 13 TeV $W^{\pm}$ boson data have smaller impact on the $u$ quark PDF as compared to the $d$ quark PDF. But with a 300 fb$^{-1}$ data sample, the LHCb 13 TeV
$W^\pm$ boson data have impact on the $u$ quark PDF in the small $x$ region.

FIG. 3: PDF uncertainties associated with $d$ quark, as a function of $x$, in the CT14HERA2 PDFs and ePUMP updated new PDFs. The denominator is the central value of each PDF set. The blue (red) line represents central value of PDF ratio before (after) ePUMP update, the blue band represents the CT14HERA2 PDF uncertainty, the red shaded band represents the updated PDF uncertainty. The charged lepton pseudorapidity distributions of $W^\pm$ events are used as inputs for ePUMP update. (top-left) shows $d$-quark result using 5 fb$^{-1}$ $W^+$ events, (top-right) shows $d$-quark result using 5 fb$^{-1}$ $W^-$ events, (bottom-left) shows $d$-quark result using 300 fb$^{-1}$ $W^+$ events, (bottom-right) shows $d$-quark result using 300 fb$^{-1}$ $W^-$ events.

The impacts of the LHCb 13 TeV $W^+/W^-$ data, the ratio of $W^+$ and $W^-$ event rates, on the PDF ratios $d/u$ and $\bar{d}/\bar{u}$ are shown in Fig. 4. As we see, even with a 5 fb$^{-1}$ data sample, the LHCb 13 TeV $W^+/W^-$ data can already reduce the $d/u$ PDF uncertainty. Most of the improvements are concentrated in the small $x$ region, from $10^{-5}$ to $10^{-3}$. The 300 fb$^{-1}$ LHCb 13 TeV data sample could further reduce the uncertainties of both the PDF ratios $d/u$ and $\bar{d}/\bar{u}$ (by about $\sim 20\%$) in the small $x$ region, as well as some noticeable improvements...
in the large $x$ region. In the current PDF global fitting, the DIS data provide the largest constraint on the PDF ratio $d/u$, cf. Ref. [37]. In the future, the LHCb data could provide additional information on the PDF ratios $d/u$ and $\bar{d}/\bar{u}$.

![Error bands of $d(x,Q)/u(x,Q)$](error_bands.png)

**FIG. 4:** PDF uncertainties associated with $d/u$ (left) and $\bar{d}/\bar{u}$ (right), as a function of $x$, in the CT14HERA2 PDFs and the ePump updated PDFs. The charged lepton pseudorapidity distributions of $W^+/W^-$ events are used as the input for the ePump update; (top-left) shows $d/u$ result using a 5 fb$^{-1} W^+/W^-$ data sample, and (top-right) for $\bar{d}/\bar{u}$; (bottom-left) shows $d/u$ result using a 300 fb$^{-1} W^+/W^-$ data sample, and (bottom-right) for $\bar{d}/\bar{u}$.

### B. Update from LHCb 13 TeV $Z$ pseudo-data

The single $Z$ boson leptonic decay has two charged leptons in the final state, these two charged leptons have large transverse momentum, and are isolated in the detector. Based on these features, the $Z \to \ell^+\ell^-$ events are easy to be reconstructed and identified in a hadron collider, with small background contamination. Therefore, the $Z \to \ell^+\ell^-$ channel is one of
best channels to perform the precision EW measurements.

We used the $Z$ boson rapidity distribution as an observable for the ePUMP update, explored other observables that could be used in the future PDF fitting, and proposed a novel way to present $Z$ boson production measurement that provides more valuable information for PDF fitting. In this study, a binning scheme similar to the previous LHCb publication [39] is used.

The updated PDF results are shown in Fig. 5 for $d$ quark. As shown in the figure, with $5\,\text{fb}^{-1}$ of data sample, the LHCb 13 TeV single $Z$ boson data is not as powerful as $W^\pm$ data, mainly due to its smaller event rate. While with $300\,\text{fb}^{-1}$, the $Z$ boson data has impact on $d$ quark PDFs in small $x$ region, from $10^{-5}$ to $10^{-2}$.

The impacts from the LHCb 13 TeV $300\,\text{fb}^{-1}$ single $Z$ boson data on $d/u$ and $\bar{d}/\bar{u}$ are shown in Fig. 6, where the LHCb $Z$ boson data could reduces the $d/u$ PDF uncertainty in the small $x$ region.

FIG. 5: PDF uncertainties associated with $d$ quark, as a function of $x$, in the CT14HERA2 PDFs and ePUMP updated new PDFs. The denominator is the central value of each PDF set. The rapidity distribution of $Z$ boson events is used as input of the ePUMP update. (left) shows $d$-quark result using $5\,\text{fb}^{-1}$ data, (right) shows $d$-quark result using $300\,\text{fb}^{-1}$ data.

We have also explored the sensitivity of the $Z$ boson $p_T$, lepton $\cos \theta^*$ (defined in Collins-Soper frame [43]), and $Z$ boson rapidity distributions measured at the LHCb to further constrain the PDFs. We consider their impacts one at a time in the ePUMP update. As shown in Fig. 7, as we see each observable has a slightly different impact on $u$ ($d$) quark PDFs in all $x$ region, as expected.
FIG. 6: PDF uncertainties associated with $d/u$ (left) and $\bar{d}/\bar{u}$ (right), as a function of $x$, in the CT14HERA2 PDFs and ePUMP updated new PDFs. The denominator is the central value of each PDF set. The rapidity distribution of $Z$ boson events is used as input of the ePUMP update. (left) shows $d/u$ quark result using 300 fb$^{-1}$ data, (right) shows $\bar{d}/\bar{u}$ quark result using 300 fb$^{-1}$ data.

Below, we propose a better way to extract useful information from the LHCb 13 TeV $Z$ data, by performing a multi-dimensional analysis. With more $pp$ collision data collected by the LHCb detector in the future, it is feasible to perform $Z$ boson production cross section measurement with a multi-dimension binning, like double-, triple-differential cross section measurements. Comparisons of updated PDF uncertainties with different number of (input) experimental observables are shown in Fig. 8. As shown in the figure, we compared the impacts of the LHCb 13 TeV $Z$ boson pseudo-data on PDFs by performing a single differential ($Z$ boson $p_T$, which is labeled as ‘1D’ in the figure), a double-differential ($Z$ boson $p_T$ and $Z$ boson rapidity, which is labeled as ‘2D’ in the figure), and a triple-differential ($Z$ boson $p_T$, $Z$ boson rapidity, and lepton $\cos \theta^*$, and which is labeled as ‘3D’ in the figure) cross section measurements. We found that with limited statistics of $Z$ boson events (5 fb$^{-1}$ data), the multi-dimensional measurements cannot significantly improve the PDF determination, as compared to one-dimensional measurements on $Z$ boson $p_T$, lepton $\cos \theta^*$, and $Z$ boson rapidity, respectively. While using a 300 fb$^{-1}$ data sample, the multi-dimensional measurement has better constraints on the PDFs, in all $x$ range. The triple-differential cross section gives the best constraints on $u$ and $d$ quark PDFs, in all $x$ range. The improvements of performing ‘2D’ to ‘3D’ measurement is not as strong as that from ‘1D’ to ‘2D’ measurement. In the experimental point of view, a triple-differential cross section measurement could has limited
statistics in extreme phase space, such as boundary of observables. Furthermore, it is complicated to calculate correlated systematic uncertainties in a ‘3D’ measurement, compared to a ‘2D’ measurement. Therefore, in the future with large data sample, a double-dimensional $Z$ boson cross section measurement (a double-differential $Z$ boson $p_T$ and $Z$ boson rapidity) is feasible and recommended, which could provides more valuable information in PDF fitting than a single- or a triple-dimensional measurement.

In the double- and triple-differential cross section measurements, the binning scheme for $Z$ boson $p_T$, rapidity ($y_{\ell\ell}$), and lepton cos $\theta^*$ are defined as:

- $0 < p_T < 250$: \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 17.5, 20, 22, 25, 28, 33, 40, 50, 100, 250\} GeV/c.
- $2 < y_{\ell\ell} < 4.5$: \{2.00, 2.14, 2.28, 2.42, 2.56, 2.69, 2.83, 2.97, 3.11, 3.25, 3.39, 3.53, 3.67, 3.81, 3.94, 4.08, 4.22, 4.36, 4.50\}.
- $-1 < \cos \theta^* < 1$: \{-1, 0, 1\}.

C. Update from LHCb 13 TeV $W^+ + W^- + Z$ pseudo-data

In general, global fitting of PDFs should use all available experimental data as inputs. Therefore, we checked the impact from the LHCb 13 TeV data on the PDF fitting, including
FIG. 8: PDF uncertainties associated with $u$ quark (left) and $d$ quark (right), as a function of $x$, in the CT14HERA2 PDFs and ePump updated new PDFs. The denominator is the central value of each PDF set. The single $Z$ boson multi-dimensional (1D, 2D, and 3D) differential cross section is used in the ePump update. (top-left) shows $u$-quark result using 5 fb$^{-1}$ data, (top-right) shows $d$-quark result using 5 fb$^{-1}$ data, (bottom-left) shows $u$-quark result using 300 fb$^{-1}$ data, (bottom-right) shows $d$-quark result using 300 fb$^{-1}$ data.

both single $W^\pm$ boson and $Z$ boson data samples. In reality, as $W^\pm$ and $Z$ boson results from one experiment are measured with same data sample, therefore, many systematical uncertainties are correlated. In any PDF fitting, correlation matrices between different observable must be provided to avoid potential data bias. In this study, without detector level simulated events, we cannot calculate the correlation matrices between the single $W^\pm$ and $Z$ boson measurements. So we assume in this study that there is no correlation between the LHCb 13 TeV $W^\pm$ and $Z$ pseudo-data. In the study, the $W^\pm$ and $Z$ boson single differential cross section results are used as the inputs of the ePump update, which are the charged lepton $\eta$ distribution of $W^\pm$ boson events and the rapidity distribution of $Z$ bosons.

The updated PDF results are shown in Fig. 9 for $u$-, $d$-, $c$-quark and gluon PDFs, and
the $d/u$ and $\bar{d}/\bar{u}$ ratio results are shown in Fig. 10, respectively. Based on these figures, the following features are found:

- The largest improvement is on the $d$-quark PDFs. The uncertainty of $d$-quark PDFs can be improved significantly by the LHCb 13 TeV $W^\pm/Z$ data in all $x$ region. Special in small $x$ region $10^{-5} < x < 10^{-2}$, the uncertainty would be reduced by a factor of 60% at $x \sim 10^{-3}$.

- The uncertainty of $u$-, $s$-, $c$-quark and gluon PDFs can be reduced in all $x$ region, and significant improvements in very small and larger $x$ region are expected.

- The uncertainties of $d/u$ and $\bar{d}/\bar{u}$ ratios can be significantly reduced in all $x$ range, even only with 5 fb$^{-1}$ data. In very larger $x$ region, the LHCb 13 TeV data could have large impact on the $d/u$ ratio.

- The LHCb 13 TeV $W^\pm$ and $Z$ data also has large impact on the $\bar{u}$- and $\bar{d}$-quark PDFs, mainly in the small $x$ region.

As to the $d/u$ and $\bar{d}/\bar{u}$ ratios, the future LHCb 13 TeV data will provide the most important constraints on them. In Fig. 10, the 300 fb$^{-1}$ LHCb 13 TeV pseudo-data provides valuable constraint on $d/u$ ratio in the very large $x$ region ($> 0.5$) and $\bar{d}/\bar{u}$ in $x > 0.2$. In these regions, the LHCb data would be the only clean data, which is free of nuclear corrections as needed when describing the low energy Drell-Yan data to constrain $d/u$.

As can be seen based upon a leading-order quark-parton model analysis, this ratio is expected to have especially pronounced sensitivity to the $x$ dependence of the PDF ratio $\bar{d}/\bar{u}$. 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 SeaQuest/E906 experiment at Fermilab [45], from which results are expected soon. The LHCb data could be used to check the impact from SeaQuest [46] result on $\bar{d}/\bar{u}$ in the large $x$ region. In Fig. 11, we compare the theoretical prediction based on the update CT14HERA2 PDFs (with the 300 fb$^{-1}$ LHCb 13 TeV combined $W^\pm$ and $Z$ pseudo-data as
FIG. 9: PDF uncertainties associated with $u$-quark (left) and $d$-quark (right), as a function of $x$, in the CT14HERA2 PDFs and ePUMP updated new PDFs. The denominator is the central value of each PDF set. The 300 fb$^{-1}$ single $W^\pm$ and $Z$ boson differential cross section are taken as ePUMP input. (top-left) shows $u$-quark result, (top-right) shows $d$-quark result, (bottom-left) shows gluon result, (bottom-right) shows $c$-quark result.

input) and the original CT14HERA2 PDFs. It illustrates that the LHCb 13 TeV data could further constrain the deuteron-to-proton ratio, $\sigma_{pd}/2\sigma_{pp}$, in the large $x$ region.

D. Update from LHCb 13 TeV $(W^+ + W^-)/Z$ pseudo-data

At tree level, the $Z$ boson is produced via $q\bar{q}$ annihilation, where $q$ could be $u$, $d$, $c$, $s$, and $b$. While the $W^\pm$ boson could produced via $u\bar{s}$ and $\bar{u}s$. Therefore, the ratio of $W^\pm$ distribution to that of $Z$ boson could be sensitive to the $\frac{s + \bar{s}}{u + \bar{u}}$ at the first order [19].

With a uniform binning (18 pseudorapidity/rapidity bins, from 2.0 to 4.5), a $(W^+ + W^-)/Z$ ratio in each bin is calculated, where muon pseudorapidity of $W^\pm$ boson and rapidity of $Z$ boson are used. Correlations between the predicted LHCb 13 TeV $(W^+ + W^-)/Z$ ratio
FIG. 10: PDF uncertainties associated with $d/u$ (left) and $\bar{d}/\bar{u}$ (right), as a function of $x$, in the CT14HERA2 PDF sets and ePUMP updated new PDF sets. The denominator is the central value of each PDF set. The single $W^\pm$ and $Z$ boson differential cross section are taken as ePUMP input. (top-left) shows $d/u$ result using 5 fb$^{-1}$ data, (top-right) shows $\bar{d}/\bar{u}$ result using 5 fb$^{-1}$ data, (bottom-left) shows $d/u$ result using 300 fb$^{-1}$ data, (bottom-right) shows $\bar{d}/\bar{u}$ result using 300 fb$^{-1}$ data.

and $\frac{s+\bar{s}}{u+d}$ are shown in Fig. 12. With calculated $(W^+ + W^-)/Z$ ratio as ePUMP input, we checked the impact on $\frac{s+\bar{s}}{u+d}$ from the LHCb single $W^\pm$ and $Z$ data, as shown in Fig. 12. With 5 fb$^{-1}$ LHCb 13 TeV pseudo-data as input, the $\frac{W^+ + W^-}{Z}$ data does not have visible impact on the PDF ratio $\frac{s+\bar{s}}{u+d}$. While with 300 fb$^{-1}$ LHCb 13 TeV pseudo-data as input, the impact on the PDF ratio $\frac{s+\bar{s}}{u+d}$ becomes significantly in the $x$ range of $10^{-2}$ to $10^{-1}$, which could be used to precisely determine the strange quark PDFs in the future. As expected, larger correlations between $(W^+ + W^-)/Z$ and $\frac{s+\bar{s}}{u+d}$ are seen in the same $x$ range.
FIG. 11: Theoretical predictions based on the updated CT14HERA2 (red band) and original CT14HERA2 (blue band) for the fixed-target Drell-Yan cross section, $\sigma_{pd}/2\sigma_{pp}$, in the region of larger $x_2 \gtrsim 0.1$ to be probed by the SeaQuest experiment [45] at Fermilab. For comparison, the higher-$x_2$ portion of the older E866 data [44] (blue diamonds) is also presented here.

FIG. 12: Left panel shows $\cos \phi$ between the PDF ratio $\frac{s+\bar{s}}{u+\bar{d}}$ and the $(W^+ + W^-)/Z$ ratio. Lines with different color represent different bins, same index as Fig. 1 is used. Right panel shows PDF uncertainties associated with $\frac{s+\bar{s}}{u+\bar{d}}$ distribution, using the 300 fb$^{-1}$ data, as a function of $x$, in the CT14HERA2 PDFs and ePUMP updated new PDFs. The denominator is the central value of each PDF set. The single $\frac{W^+ + W^-}{Z}$ boson differential cross section are taken as ePUMP input.
The tolerance criterion (i.e., the choice of total $\Delta \chi^2$ value in a global analysis) is an important parameter in PDF fitting.

It was extensively discussed in Ref. [47] that in order to best reproduce CT14HERA2 global fit, one should use dynamical tolerance in ePump. If the tolerance is set to be $\Delta \chi^2 = 1$ at 68% confidence level (CL), or equivalently, $(1.645)^2$ at 90% CL, instead of using a dynamical tolerance, a very large weight is effectively assigned to the new input data, when updating the existing PDFs in the CT PDF global analysis framework.

To illustrate differences between the ePump updated PDFs using a dynamical tolerance and a fixed tolerance with $\Delta \chi^2 = (1.645)^2$, we used the single LHCb 13 TeV $Z$ data as the ePump input, the result is shown in Fig. 13. The impact from the new data are enhanced when update the PDFs with $\Delta \chi^2 = (1.645)^2$, which could introduce bias in the new updated PDF set.

This conclusion also holds for using MMHT2014 [27] and PDF4LHC15 [48] PDFs in profiling analysis to study the impact of a new (pseudo-) data on updating the existing PDFs.

**FIG. 13**: PDF uncertainties associated with $d$-quark distribution, as a function of $x$, in the CT14HERA2 PDFs and ePump updated new PDFs. The denominator is the central set of each PDF sets. The rapidity distribution of $Z$ boson events (300 fb$^{-1}$) is used as input of the ePump update. (left) shows result with fixed tolerance of $\Delta \chi^2 = (1.645)^2$ at the 90% CL; (right) shows result with dynamical tolerance.
V. CONCLUSION

With detector instrumented in the forward region, the $pp$ collision data collected by the LHCb detector provides essential and complementary information for a global analysis of experimental data to determine the Parton Distribution Functions (PDFs) of the proton, as compared to data collected by the ATLAS and CMS detectors.

In this article, we studied the potential of the LHCb 13 TeV single $W^\pm$ and $Z$ boson pseudo-data on constraining the PDFs in proton. As an example, we demonstrated the sensitivity of the LHCb 13 TeV data, collected with an integrated luminosity of $5 \text{ fb}^{-1}$ and $300 \text{ fb}^{-1}$, respectively, to reducing the PDF uncertainty bands of the CT14HERA2 PDFs. We have also investigated the sensitivities of various experimental observables.

The large impact from the LHCb 13 TeV data on various quark flavor PDFs in all $x$ region had been seen, and significant contributions in small $x$ ($< 10^{-3}$) region are expected. Particularly, the $d$ and $\bar{d}$ quark PDF uncertainties are reduced dramatically, $\sim 60\%$ improvement at momentum fraction ($x$) around $10^{-3}$. Due to its large event rate, the LHCb 13 TeV $W^+/W^-$ data can already further reduce the $d/u$ PDF uncertainty, even with only a $5 \text{ fb}^{-1}$ data sample, as shown in Fig. 4. Most of the improvements are concentrated in the small $x$ region, from $10^{-5}$ to $10^{-3}$. The $300 \text{ fb}^{-1}$ LHCb 13 TeV data sample could further reduce the uncertainties of both the PDF ratios $d/u$ and $\bar{d}/\bar{u}$ (by about $\sim 20\%$) in the small $x$ region, as well as some noticeable improvements in the large $x$ region which is currently dominated by DIS data in global analysis.

Although with a smaller event rate as compared to the $W^\pm$ events, the LHCb 13 TeV $Z$ data can also provide important constraints on PDFs, particularly when considering a double-differential distribution (as a function of $Z$ boson $p_T$ and rapidity), as the experimental observable for updating the existing PDFs. We have compared the impact of the LHCb 13 TeV $Z$ boson pseudo-data on PDFs by performing a single differential ($Z$ boson $p_T$), a double-differential ($Z$ boson $p_T$ and $Z$ boson rapidity), and a triple-differential ($Z$ boson $p_T$, $Z$ boson rapidity, and lepton $\cos \theta^*$) cross section measurements. We found that with limited statistics of $Z$ boson events ($5 \text{ fb}^{-1}$ data), the multi-dimensional measurements cannot significantly improve the PDF determination, as compared to one-dimensional measurements on $Z$ boson $p_T$, lepton $\cos \theta^*$, and $Z$ boson rapidity, respectively. While using a $300 \text{ fb}^{-1}$ data sample, the multi-dimensional measurement has better constraints on the
PDFs, in all $x$ range.

It is evidently that the combined data sample of $W^\pm$ and $Z$ boson events can provide further constraints on PDFs. Examining the ratio of the event rates $\frac{W^+ + W^-}{Z}$ could directly probe the PDF ratio $\frac{s+\bar{s}}{u+d}$, as previously noted in Ref. [19]. With 5 fb$^{-1}$ LHCb 13 TeV pseudo-data as input, the $\frac{W^+ + W^-}{Z}$ data does not have visible impact on the PDF ratio $\frac{s+\bar{s}}{u+d}$. While with 300 fb$^{-1}$ LHCb 13 TeV pseudo-data as input, the impact on the PDF ratio $\frac{s+\bar{s}}{u+d}$ becomes significantly in the $x$ range of $10^{-2}$ to $10^{-1}$, which could be used to precisely determine the strange quark PDFs in the future. The above features suggest that the LHCb single $W^\pm$ and $Z$ data taken at the LHC 13 TeV will provide very important and unique information in a global analysis, complementary to the ATLAS and CMS results. With an integrated luminosity of 300 fb$^{-1}$ collected data in the future, the impact from the LHCb 13 TeV data on the PDF fitting could be enhanced significantly, by performing a multi-differential cross section measurement.

Before concluding our work, we also pointed out the important role of tolerance criteria in the PDF updating. As discussed in Ref. [47], one should use dynamical tolerance in ePUMP. Setting tolerance to be 1 at 68% CL, or equivalently, $(1.645)^2$ at 90% CL, will greatly overestimate the impact of a given new data set when updating the existing PDFs in the CT PDF global analysis framework. This conclusion also holds for using MMHT2014 [27] and PDF4LHC15 [48] PDFs in profiling analysis to study the impact of a new (pseudo-) data on updating the existing PDFs.

VI. ACKNOWLEDGMENTS

This work is supported by the National Natural Science Foundation of China under the Grant No. 11875142, No. 11965020, and by the U.S. National Science Foundation under Grant No. PHY-2013791. C.-P. Yuan is also grateful for the support from the Wu-Ki Tung endowed chair in particle physics.

[1] CDF Collaboration, T. Aaltonen et al., Precise measurement of the $W$ boson mass with the CDF II detector, Phys. Rev. Lett. 108, 151803 (2012), arXiv:1203.0275.
[2] **D0** Collaboration, V. M. Abazov et al., Measurement of the $W$ Boson Mass with the D0 Detector, *Phys. Rev. Lett.* **108**, (2012) 151804, arXiv:1203.0293.

[3] **ATLAS** Collaboration, M. Aaboud et al., Measurement of the $W$-boson mass in $pp$ collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector, *Eur. Phys. J. C* **78**, (2018) 110, Erratum: *Eur. Phys. J. C* **78**, (2018) 898, arXiv:1701.07240.

[4] V. N. Gribov and L. N. Lipatov, Deep inelastic $ep$ scattering in perturbation theory, *Sov. J. Nucl. Phys.* **15**, (1972) 438–450.

[5] Yuri L. Dokshitzer, Calculation of the Structure Functions for Deep Inelastic Scattering and $e^{+}e^{-}$ Annihilation by Perturbation Theory in Quantum Chromodynamics, *Sov. Phys. JETP* **46**, (1977) 641–653.

[6] G. Altarelli and G. Parisi, Asymptotic Freedom in Parton Language, *Nucl. Phys. B* **126**, (1977) 298–318.

[7] P. J. Rijken and W. L. van Neerven, Order $\alpha_s^2$ contributions to the Drell-Yan cross-section at fixed target energies, *Phys. Rev.* **D51** (1995) 44–63, arXiv:hep-ph/9408366.

[8] R. Hamberg, W. L. van Neerven, and T. Matsuura, A complete calculation of the order $\alpha_s^2$ correction to the Drell-Yan $K$ factor, *Nucl. Phys. B**359** (1991) 343–405. Erratum: *Nucl. Phys. B**464**, (2002) 403.

[9] R. V. Harlander and W. B. Kilgore, Next-to-next-to-leading order Higgs production at hadron colliders, *Phys. Rev. Lett.* **88** (2002) 201801, arXiv:hep-ph/0201206.

[10] W. L. van Neerven and E. B. Zijlstra, The $O(\alpha_s^2)$ corrected Drell-Yan $K$ factor in the DIS and MS scheme, *Nucl. Phys. B**382**, (1992) 11. Erratum: *Nucl. Phys. B**680**, (2004) 513.

[11] C. Anastasiou, L. Dixon, K. Melnikov, and F. Petriello, High precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at NNLO, *Phys. Rev. D**69**, (2004) 094008, arXiv:hep-ph/0312266.

[12] S. Dulat, T. J. Hou, J. Gao, M. Guzzi, J. Huston, P. Nadolsky, J. Pumplin, C. Schmidt, D. Stump, and C. P. Yuan, New parton distribution functions from a global analysis of quantum chromodynamics, *Phys. Rev. D**93**, (2016) 033006, arXiv:1506.07443.

[13] **ATLAS** Collaboration, G. Aad et al., Measurement of the inclusive $W^{\pm}$ and $Z/\gamma^*$ cross sections in the electron and muon decay channels in $pp$ collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector, *Phys. Rev. D**85**, (2012) 072004, arXiv:1109.5141.
[14] CMS Collaboration, S. Chatrchyan et al., Measurement of the electron charge asymmetry in inclusive \( W \) production in \( pp \) collisions at \( \sqrt{s} = 7 \) TeV, Phys. Rev. Lett. 109 (2012) 111806, arXiv:1206.2598.

[15] CMS Collaboration, S. Chatrchyan et al., Measurement of the muon charge asymmetry in inclusive \( pp \rightarrow W + X \) production at \( \sqrt{s} = 7 \) TeV and an improved determination of light parton distribution functions, Phys. Rev. D90 (2014) 032004, arXiv:1312.6283.

[16] LHCb Collaboration, R. Aaij et al., Inclusive \( W \) and \( Z \) production in the forward region at \( \sqrt{s} = 7 \) TeV, JHEP 06 (2012) 058, arXiv:1204.1620.

[17] S. Alekhin, J. Blumlein, and S. Moch, The ABM parton distributions tuned to LHC data, Phys. Rev. D89 (2014) 054028, arXiv:1310.3059.

[18] J. Gao, and L. Harland-Lang, and J. Rojo, The Structure of the Proton in the LHC Precision Era, Phys. Rept. 742 (2018) 1–121, arXiv:1709.04922.

[19] P. M. Nadolsky, H.-L. Lai, Q.-H. Cao, J. Huston, J. Paplin, D. Stump, W.-K. Tung, and C. P. Yuan, Implications of CTEQ-TEA global analysis for collider observables, Phys. Rev. D78 (2008) 013004, arXiv:0802.0007.

[20] LHCb Collaboration, A. A. Alves, Jr. et al. The LHCb Detector at the LHC, JINST 3 (2008) S08005.

[21] LHCb Collaboration, R. Aaij et al. LHCb Detector Performance, Int. J. Mod. Phys. A 30, (2015) 1530022, arXiv:1412.6352.

[22] LHCb Collaboration, R. Aaij et al., Measurement of the cross-section for \( Z \rightarrow e^+e^- \) production in \( pp \) collisions at \( \sqrt{s} = 7 \) TeV, JHEP 1302, (2013) 106, arXiv:1212.4620.

[23] LHCb Collaboration, R. Aaij et al., Measurement of the forward \( W \) boson cross-section in \( pp \) collisions at \( \sqrt{s} = 7 \) TeV, JHEP 1412, (2014) 079, arXiv:1408.4354.

[24] LHCb Collaboration, R. Aaij et al., Measurement of forward \( Z \rightarrow e^+e^- \) production at \( \sqrt{s} = 8 \) TeV, JHEP 1505, (2015) 109, arXiv:1503.00963.

[25] LHCb Collaboration, R. Aaij et al., Measurement of forward \( W \) and \( Z \) boson production in \( pp \) collisions at \( \sqrt{s} = 8 \) TeV, JHEP 1601, (2016) 155, arXiv:1511.08039.

[26] LHCb Collaboration, R. Aaij et al., Measurement of forward \( W \rightarrow e\nu \) production in \( pp \) collisions at \( \sqrt{s} = 8 \) TeV, JHEP 1610, (2016) 030, arXiv:1608.01484.

[27] L. A. Harland-Lang, A. D. Martin, P. Motylinski and R. S. Thorne, Parton distributions in the LHC era: MMHT 2014 PDFs, Eur. Phys. J. C 75, (2015) 204, arXiv:1412.3989.
[28] **NNPDF** Collaboration, R. D. Ball *et al.*, Parton distributions for the LHC Run II, *JHEP* **1504** (2015) 040, [arXiv:1410.8849](https://arxiv.org/abs/1410.8849).

[29] **LHCb** Collaboration, R. Aaij *et al.*, Precision luminosity measurements at LHCb, *JINST* **9**, (2014) P12005, [arXiv:1410.0149](https://arxiv.org/abs/1410.0149).

[30] **LHCb** Collaboration, R. Aaij *et al.*, Observation of CP Violation in Charm Decays, *Phys. Rev. Lett.* **122**, (2019) 11803, [arXiv:1903.08726](https://arxiv.org/abs/1903.08726).

[31] **LHCb** Collaboration, I. Bediaga *et al.*, Framework TDR for the LHCb Upgrade: Technical Design Report, CERN-LHCC-2012-007, LHCb-TDR-12.

[32] **LHCb** Collaboration, R. Aaij *et al.*, Physics case for an LHCb Upgrade II - Opportunities in flavour physics, and beyond, in the HL-LHC era, LHCb-PUB-2018-009, CERN-LHCC-2018-027, [arXiv:1808.08865](https://arxiv.org/abs/1808.08865).

[33] **LHCb** Collaboration, R. Aaij *et al.*, Expression of Interest for a Phase-II LHCb Upgrade: Opportunities in flavour physics, and beyond, in the HL-LHC era, CERN-LHCC-2017-003.

[34] I. Efthymiopoulos *et al.*, LHCb Upgrades and operation at 10^{34} cm^{-2}s^{-1} luminosity –A first study, CERN-ACC-NOTE-2018-0038.

[35] T. J. Hou, S. Dulat, J. Gao, M. Guzzi, J. Huston, P. Nadolsky, J. Pumplin, C. Schmidt, D. Stump, C. P. Yuan CTEQ-TEA parton distribution functions and HERA Run I and II combined data, *Phys. Rev. D* **95**, (2017) 034003, [arXiv:1609.07968](https://arxiv.org/abs/1609.07968).

[36] C. Schmidt, J. Pumplin, and C. P. Yuan, Updating and optimizing error parton distribution function sets in the Hessian approach, *Phys. Rev. D* **98**, (2018) 094005, [arXiv:1806.07950](https://arxiv.org/abs/1806.07950).

[37] T. J. Hou, Z. Yu, S. Dulat, C. Schmidt and C. P. Yuan, Updating and optimizing error parton distribution function sets in the Hessian approach. II., *Phys. Rev. D* **100**, (2019) 114024, [arXiv:1907.12177 [hep-ph]](https://arxiv.org/abs/1907.12177).

[38] C. Willis, R. Brock, D. Hayden, T. J. Hou, J. Isaacson, C. Schmidt and C. P. Yuan, New method for reducing parton distribution function uncertainties in the high-mass Drell-Yan spectrum, *Phys. Rev. D* **99**, (2019) 054004, [arXiv:1809.09481](https://arxiv.org/abs/1809.09481).

[39] **LHCb** Collaboration, R. Aaij *et al.*, Measurement of the forward Z boson production cross-section in pp collisions at √s = 13 TeV, *JHEP* **1609**, (2016) 136, [arXiv:1607.06495](https://arxiv.org/abs/1607.06495).

[40] C. Balazs and C. P. Yuan, Soft gluon effects on lepton pairs at hadron colliders, *Phys. Rev. D* **56**, (1997) 5558, [arXiv:hep-ph/9704258](https://arxiv.org/abs/hep-ph/9704258).
[41] F. Landry, R. Brock, P. M. Nadolsky and C. Yuan, Tevatron Run-1 $Z$ boson data and Collins-Soper-Sterman resummation formalism, Phys. Rev. D 67, (2003) 073016, arXiv:hep-ph/0212159 [hep-ph].

[42] P. Sun, J. Isaacson, C. P. Yuan and F. Yuan, Nonperturbative functions for SIDIS and Drell–Yan processes, Int. J. Mod. Phys. A 33, (2018) 11, 1841006, arXiv:1406.3073 [hep-ph].

[43] J. C. Collins and D. E. Soper, Angular Distribution of Dileptons in High-Energy Hadron Collisions, Phys. Rev. D 16, (1977) 2219.

[44] NuSea Collaboration, R. S. Towell et al., Improved measurement of the $\bar{d}/\bar{u}$ asymmetry in the nucleon sea, Phys. Rev. D64 (2001) 052002, arXiv:hep-ex/0103030.

[45] SeaQuest Collaboration, C. A. Aidala et al., The SeaQuest Spectrometer at Fermilab, Nucl. Instrum. Meth. A 930 (2019) 49–63, arXiv:1706.09990.

[46] A. Tadepalli, Light anti-quark flavor asymmetry in the nucleon sea and the nuclear dependence of anti-quarks in nuclei at the SeaQuest Experiment, FERMILAB-THESIS-2019-12.

[47] T. J. Hou, J. Gao, T. J. Hobbs, K. Xie, S. Dulat, M. Guzzi, J. Huston, P. Nadolsky, J. Pumplin, C. Schmidt, I. Sitiwaldi, D. Stump, and C. P. Yuan, New CTEQ global analysis of quantum chromodynamics with high-precision data from the LHC, arXiv:1912.10053.

[48] J. Butterworth et al., PDF4LHC recommendations for LHC Run II, J. Phys. G 43 (2016) 023001, arXiv:1510.03865.