

Vector boson production in pPb and PbPb collisions at the LHC and its impact on nCTEQ15 PDFs

A. Kusina^{1,2,a}, F. Lyonnet^{3,b}, D. B. Clark^{3,c}, E. Godat^{3,d}, T. Ježo^{4,e}, K. Kovařík^{5,f},
E. I. Olness^{3,g}, I. Schienbein^{1,h}, J. Y. Yu^{3,i}

¹ Laboratoire de Physique Subatomique et de Cosmologie, Université Grenoble-Alpes, CNRS/IN2P3, 53 Avenue des Martyrs, 38026 Grenoble, France

² Institute of Nuclear Physics Polish Academy of Sciences, 31342 Krakow, Poland

³ Southern Methodist University, Dallas, TX 75275, USA

⁴ Physik-Institut, Universität Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland

⁵ Institut für Theoretische Physik, Westfälische Wilhelms-Universität Münster, Wilhelm-Klemm-Straße 9, 48149 Münster, Germany

Received: 23 February 2017 / Accepted: 29 June 2017 / Published online: 21 July 2017

© The Author(s) 2017. This article is an open access publication

Abstract We provide a comprehensive comparison of W^\pm/Z vector boson production data in pPb and PbPb collisions at the LHC with predictions obtained using the nCTEQ15 PDFs. We identify the measurements which have the largest potential impact on the PDFs, and estimate the effect of including these data using a Bayesian reweighting method. We find this data set can provide information as regards both the nuclear corrections and the heavy flavor (strange quark) PDF components. As for the proton, the parton flavor determination/separation is dependent on nuclear corrections (from heavy target DIS, for example), this information can also help improve the proton PDFs.

Contents

1 Introduction	1
2 W^\pm/Z production at the LHC	2
2.1 Experimental data and theoretical setup	2
2.2 Comparison to proton–lead (pPb) data	3
2.3 Comparison to lead–lead data	6
2.4 W^\pm/Z cross-section correlations	7

3 Reweighting	10
3.1 Basics of PDF reweighting	10
3.2 Reweighting using CMS W^\pm rapidity distributions	13
3.3 Using asymmetries instead of differential cross sections	14
3.4 Including all the data sets	16
3.5 Comparison with EPPS16	21
4 Conclusions	23
References	24

1 Introduction

Vector boson production in hadron collisions is a well understood process and serves as one of the “standard candle” measurements at the LHC. W^\pm and Z bosons are numerous produced in heavy ion pPb and PbPb collisions at the LHC and can be used to gain insight into the structure of nuclear parton distribution functions (nPDFs). As the W^\pm and Z bosons couple weakly, their interaction with the nuclear medium is negligible which makes these processes one of the cleanest probes of the nuclear structure available at the LHC. The possibility of using vector boson production data to constrain nPDFs was previously considered [11], and this demonstrated the strong potential for the proton–lead data (especially the asymmetries) to constrain the nuclear PDFs. The current LHC measurements for W^\pm and Z production include rapidity and transverse momentum distributions for both proton–lead (pPb) and lead–lead (PbPb) collisions [1–10]. Some of these data were already used (along with jet and charged particle production data) in a recent analysis [12, 13]

^a e-mail: kusina@lpsc.in2p3.fr

^b e-mail: flyonnet@smu.edu

^c e-mail: dbclark@smu.edu

^d e-mail: egodat@smu.edu

^e e-mail: tomas.jezo@physik.uzh.ch

^f e-mail: karol.kovarik@uni-muenster.de

^g e-mail: olness@smu.edu

^h e-mail: ingo.schienbein@lpsc.in2p3.fr

ⁱ e-mail: yu@lpsc.in2p3.fr

employing a reweighting method to estimate the impact of these data on EPS09 [14] and DSSZ [15] nPDFs.¹

The LHC heavy ion W^\pm/Z data set is especially interesting as it can help to resolve the long-standing dilemma regarding the heavy flavor components of the proton PDFs. Historically, this has been an important issue as nuclear target data (especially ν -DIS) have been essential in identifying the individual parton flavors [17–20]; however, this means that the uncertainties of the heavy flavors are intimately tied to the (large) nuclear uncertainties. The LHC heavy ion W^\pm/Z data has the potential to improve this situation due to the following two key features. First, this data is in a kinematic regime where the heavier quark flavors (such as strange and charm) contribute substantially. Second, by comparing the proton W^\pm/Z data with the heavy ion results we have an ideal environment to precisely characterize the nuclear corrections. The combination of the above can not only improve the nuclear PDFs, but also the proton PDFs which are essential for any LHC study.

In this work we present predictions for vector boson production in pPb and PbPb collisions at the LHC obtained using nCTEQ15 nuclear parton distributions, and perform a comprehensive comparison to the available LHC data. We also identify the measurements which have the largest potential to constrain the nPDFs, and perform a reweighting study which allows us to estimate the effects of including these data in an nPDF fit.

The rest of the paper is organized as follows. Section 2 is devoted to predictions of vector boson production at the LHC in nuclear collisions. In particular, we provide an overview of the kinematic range probed by the W^\pm/Z data and discuss the tools we will use for the calculation. Then we present our predictions for pPb and PbPb collisions at the LHC and compare them with the experimental data and other theoretical predictions. In Sect. 3 we perform a reweighting using nCTEQ15 distributions to assess the impact of the nuclear data on the nPDFs. Finally, Sect. 4 summarizes our results and observations.

2 W^\pm/Z production at the LHC

We begin by presenting our predictions for W^\pm and Z boson production in nuclear collisions at the LHC using the recently published nCTEQ15 PDFs [21].

2.1 Experimental data and theoretical setup

For the theoretical calculations in our study we use the FEWZ (Fully Exclusive W , Z production) [22,23] program ver-

sion 2.1. Even though FEWZ can compute W and Z production with decays up to next-to-next-to-leading order, we work at next-to-leading order (NLO) to be consistent with the order of evolution of the nPDFs.²

As FEWZ is designed to handle pp or $p\bar{p}$ collisions, we have extended it so that two different PDF sets can be used for the two incoming beams as required for the pPb collisions.

For the lead PDFs we use the nCTEQ15 nPDFs [21], while we use the CT10 distributions [24] for the free protons; the only exception is the use of MSTW2008 PDFs [25] for the LHCb Z boson measurement [5] in order to match the original LHCb publication. Additionally, we compare these results with predictions calculated using nuclei made out of free-proton PDFs, and in some cases free-proton PDFs supplemented with EPS09 nuclear corrections [14].

We will consider LHC data on W^\pm and Z boson production from the ALICE, ATLAS, CMS, and LHCb experiments. The exhaustive list of data sets that we use is provided in Table 1 along with the experimental kinematical cuts implemented in the analysis. While there are measurements for both the rapidity and the transverse momentum distributions, for this study we will focus only on the rapidity measurements.

Using the transverse momentum (p_T) distributions to study the PDFs is more intricate as it requires resummations in the low p_T region where the cross section is maximal; we reserve this for a future study.

In Fig. 1 we display the kinematic space probed by the W^\pm/Z production process [26]. We translate between the $\{x_1, x_2\}$ and the $\{y, \tau\}$ variables for three values of the collider center of mass (CM) energy, \sqrt{s} . Table 2 lists the CM energy per nucleon as a function of the nominal proton beam energy which is determined from the relation:

$$\sqrt{s_{N_1 N_2}} = \sqrt{s_{pp}} \sqrt{\frac{Z_{N_1}}{A_{N_1}}} \sqrt{\frac{Z_{N_2}}{A_{N_2}}}, \quad (1)$$

where in the case of lead we have $A = 208$ and $Z = 82$. Additionally for asymmetric collisions there is a rapidity shift, δy , between the CM and the laboratory (LAB) frame:

$$\delta y = \frac{1}{2} \ln \left[\frac{E_{N_1}}{E_{N_2}} \right], \quad (2)$$

and in particular for the case of pPb collisions, $E_{Pb} = (Z_{Pb}/A_{Pb})E_p$ giving $\delta y_{pPb} = \frac{1}{2} \ln \left(\frac{82}{208} \right) \simeq -0.465$, i.e. $y_{CM} = y_{LAB} - 0.465$.

For the asymmetric case of pPb, we use the convention where x_1 is the proton momentum fraction, and x_2 is the lead momentum fraction. Thus, for pPb at large y_{CM} we have a

¹ During the publication process of this study a new global analysis including pPb LHC data has been presented [16].

² The CT10 proton PDFs used in the theoretical calculations are also at NLO.

Table 1 LHC data sets considered in this analysis

	Observable	Cuts (GeV)	Figure
pPb			
ATLAS	$d\sigma(Z \rightarrow \ell^+\ell^-)/dy_Z$ [1]	$ y_Z^{CM} < 3.5; 60 < m_{\ell^+\ell^-} < 120$	Figure 3
	$d\sigma(W^+ \rightarrow \ell^+\nu)/dy_{\ell^+}$ [2]	$p_T^{\ell^\pm} > 25; m_T^{\ell^\pm} > 40; \eta_{lab}^{\ell^\pm} < 2.4$	Figure 7a
	$d\sigma(W^- \rightarrow \ell^-\bar{\nu})/dy_{\ell^-}$ [2]	$p_T^{\ell^\pm} > 25; m_T^{\ell^\pm} > 40; \eta_{lab}^{\ell^\pm} < 2.4$	Figure 7b
CMS	$d\sigma(Z \rightarrow \ell^+\ell^-)/dy_Z$ [3]	$ \eta_{lab}^{\ell^\pm} < 2.4; 60 < m_{\ell^+\ell^-} < 120; p_T^{\ell^+(\ell^-)} > 20$	Figure 4
	$d\sigma(W^+ \rightarrow \ell^+\nu)/dy_{\ell^+}$ [4]	$p_T^{\ell^\pm} > 25; \eta_{lab}^{\ell^\pm} < 2.4$	Figure 6a
	$d\sigma(W^- \rightarrow \ell^-\bar{\nu})/dy_{\ell^-}$ [4]	$p_T^{\ell^\pm} > 25; \eta_{lab}^{\ell^\pm} < 2.4$	Figure 6b
LHCb	$\sigma(Z \rightarrow \ell^+\ell^-)$ [5]	$60 < m_{\ell^+\ell^-} < 120; p_T^{\ell^+(\ell^-)} > 20;$ $2.0 < \eta^{\ell^\pm} < 4.5; -4.5 < \eta_{\ell^\pm} < -2.0$	Figure 5
ALICE	$\sigma(W^+ \rightarrow \ell^+\nu)$ [6]	$p_T^{\ell^\pm} > 10; 2.03 < \eta_{lab}^{\ell^\pm} < 3.53;$ $-4.46 < \eta_{lab}^{\ell^\pm} < -2.96$	Figure 8a
	$\sigma(W^- \rightarrow \ell^-\bar{\nu})$ [6]	$p_T^{\ell^\pm} > 10; 2.03 < \eta_{lab}^{\ell^\pm} < 3.53;$ $-4.46 < \eta_{lab}^{\ell^\pm} < -2.96$	Figure 8b
PbPb			
ATLAS	$1/\sigma_{tot}d\sigma/dy_Z$ [7]	$66 < m_{\ell^+\ell^-} < 116; y_Z < 2.5$	Figure 9a
	A_ℓ [8]	$p_T^\ell < 25; \eta_{lab}^\ell < 2.5; m_T > 40; p_T^{miss} < 25$	Figure 10a
CMS	$1/\sigma_{tot}d\sigma/dy_Z$ [9]	$60 < m_{\ell^+\ell^-} < 120; y_Z < 2.0$	Figure 9b
	A_ℓ [10]	$p_T^\ell < 25; \eta_{lab}^\ell < 2.1; m_T > 40$	Figure 10b

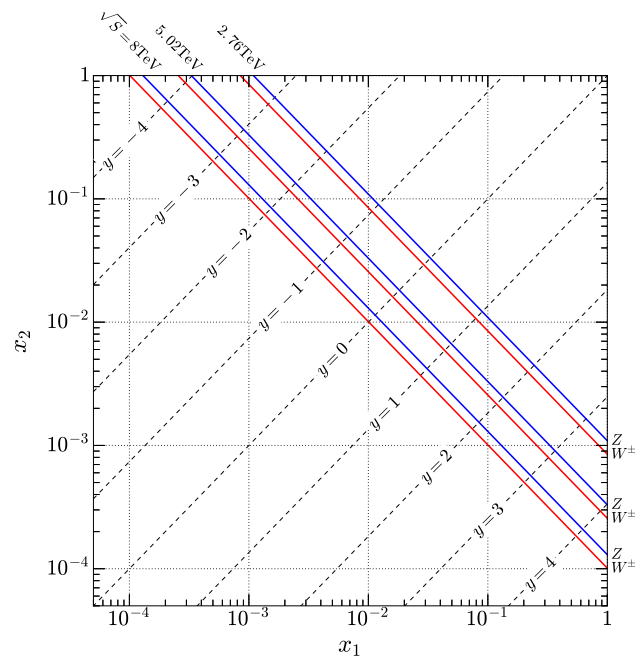


Fig. 1 The kinematic (x_1, x_2) space explored by the measurements in this study. We display lines of constant $\tau = M_V/\sqrt{s}$ where M_V is the invariant mass of the produced W^\pm/Z vector boson, as well as the appropriate rapidity shift according to Eq. (2). In Fig. 2, we display the kinematic range of the pPb data bins (central values) in the plane (y, x_2) where y is the rapidity in the CM frame of the relevant vector boson or lepton, and x_2 the lead Pb momentum fraction

Table 2 The CM energy per nucleon for pp, pPb and PbPb collisions vs. the proton beam energy in TeV units

Beam energy (TeV)	3.5	4	6.5	7
$\sqrt{s_{pp}}$	7.00	8.00	13.00	14.00
$\sqrt{s_{PbPb}}$	2.76	3.15	5.12	5.52
$\sqrt{s_{pPb}}$	4.40	5.02	8.16	8.79

large proton x_1 and a small lead x_2 ; conversely, at small y_{CM} we have a small proton x_1 and a large lead x_2 .

In Fig. 1, the pair of lines with $\sqrt{s} = 2.76$ TeV corresponds to PbPb collisions with a beam energy of 3.5 TeV per proton, and $\sqrt{s} = 5.02$ TeV corresponds to pPb collisions with a beam energy of 4 TeV per proton.

2.2 Comparison to proton–lead (pPb) data

We first consider the LHC pPb collisions at $\sqrt{s} = 5.02$ TeV. The distributions are shown in the CM frame, and include the appropriate rapidity shift according to Eq. (2). In Fig. 2, we display the kinematic range of the pPb data bins (central values) in the plane (y, x_2) where y is the rapidity in the CM frame of the relevant vector boson or lepton, and x_2 the lead

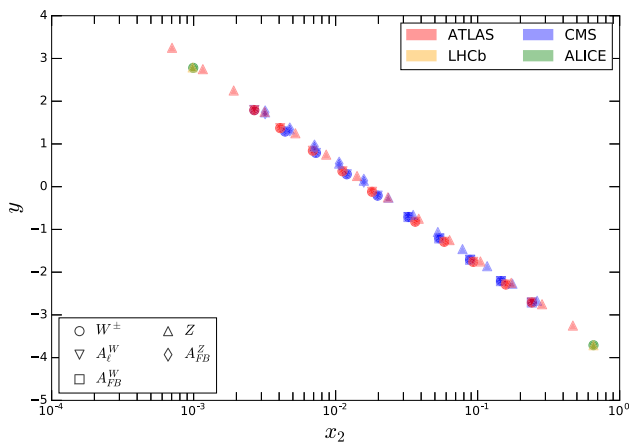


Fig. 2 Range of the pPb data used for reweighting. y is rapidity in the CM frame and x_2 is momentum of the parton from the lead beam

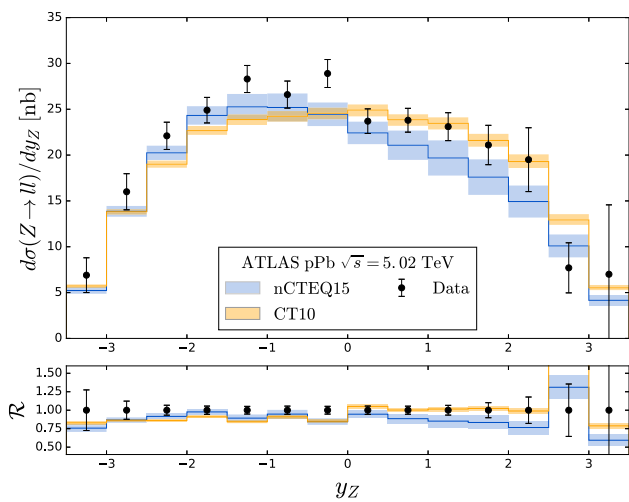


Fig. 3 ATLAS Z production in pPb collisions

parton momentum fraction. As expected, there is little data below $x \sim 10^{-3}$ and most of the constraints from these LHC data are in the low- to mid- x region.

Figures 3, 4 and 5 show our predictions for the ATLAS [1], CMS [3] and LHCb [5] Z boson production measurements, respectively. In all three cases, results obtained with the nCTEQ15 nPDFs are shown along with those obtained with a lead nucleus composed of Z protons and $A - Z$ neutrons, assuming isospin symmetry and using CT10 PDFs; the ratio of predictions over the data is shown in the lower panel. Note that the errors shown for the nCTEQ15 predictions are for nuclear uncertainties only (and only for the beam with momentum fraction x_2) which means that the PDF error of the proton beam is not accounted for.³ Furthermore, the errors shown for the pPb predictions using lead nuclei constructed from CT10 and MSTW2008 proton PDFs are only for the

³ For the symmetric case of PbPb collisions the errors on both beams are taken into account.

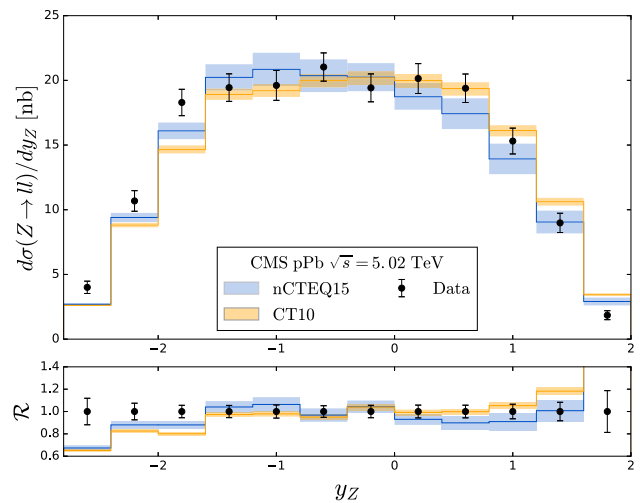


Fig. 4 CMS Z production in pPb collisions

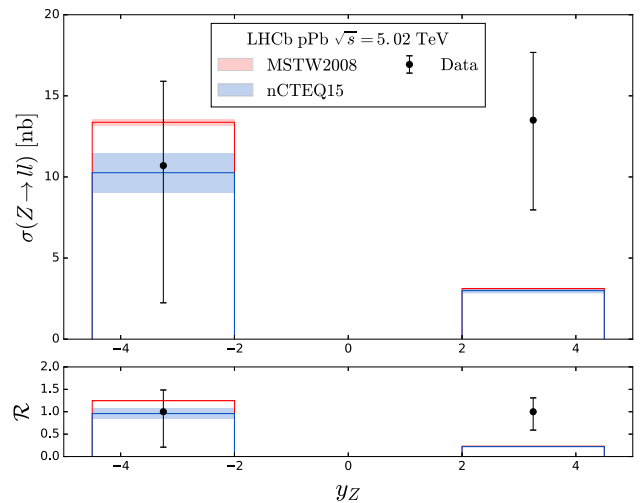


Fig. 5 LHCb Z production in pPb collisions

beam with momentum fraction x_2 . By comparing the proton uncertainties (CT10 and MSTW2008) to the nuclear uncertainties, we see that the nuclear uncertainties are much larger.

Examining Figs. 3, 4 and 5, it is interesting to note the following.

1. The data and theory are generally compatible (without significant tension) both with and without nuclear corrections; this situation may change as the experimental errors and nuclear uncertainties are reduced.
2. Focusing on the ATLAS and CMS comparison of Figs. 3 and 4, we observe that the distributions peak at negative rapidities $yz \sim -1$. Referring to Fig. 1, this corresponds to an enhancement of the $q\bar{q}$ proton–lead luminosity over the pure proton one in the x_2 region ~ 0.05 .
3. Focusing on the LHCb data of Fig. 5, we find good agreement for negative y , but large differences at pos-

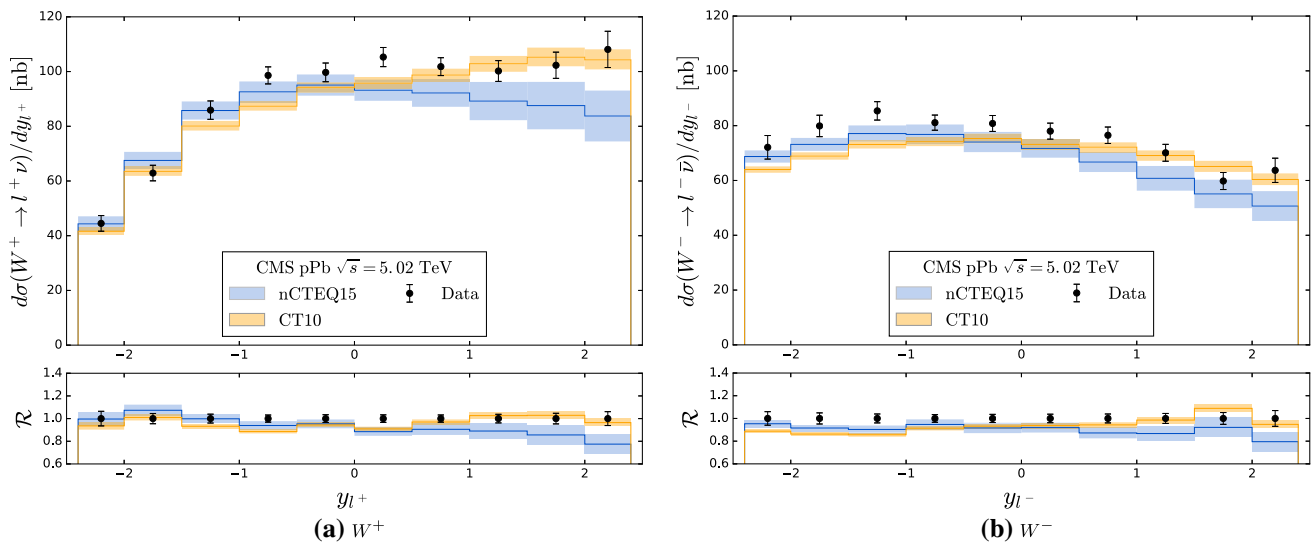


Fig. 6 CMS W^\pm production in pPb collisions at the LHC

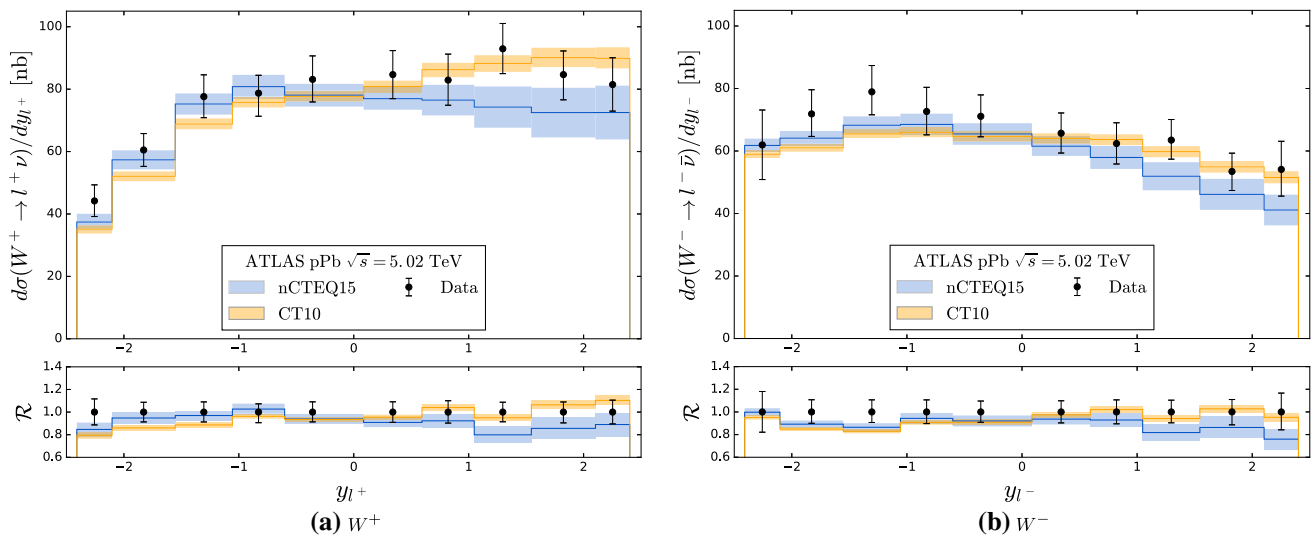


Fig. 7 ATLAS W^\pm production in pPb collisions at the LHC

itive y . Despite these differences, the large uncertainties will yield a reduced impact in our subsequent reweighting procedure.

We now turn our attention to W^+ and W^- production at the LHC. In Figs. 6, 7 and 8 we compare the data obtained by CMS [4], ATLAS [2] and ALICE [6] for W^\pm production with theoretical predictions obtained with nCTEQ15 and CT10 PDFs.

We find the W^- CMS and ATLAS data are adequately described in the negative rapidity range ($y_{\ell^-} < 0$), but the tensions grow as we move to larger rapidity. This effect is magnified for the case of W^+ where we see substantive deviations at large rapidity ($y_{\ell^+} > 1$). Referring to Fig. 1, these deviations are in the smaller x_2 region ($\sim 3 \times 10^{-3}$) where

we might expect nuclear shadowing of the $u\bar{d}$ and $d\bar{u}$ luminosities.⁴ However, this low x_2 range is unconstrained by the data currently used in nPDF fits, so these results come from an extrapolation of the larger x_2 region. It is interesting to observe that a delayed shadowing (which shifts the shadowing down to smaller x_2 values) would improve the comparison of the data with the theory in the larger y_{ℓ^\pm} region; this type of behavior was observed in the nuclear corrections

⁴ The nuclear correction factors are typically defined as the ratio of the nuclear quantity to the proton or isoscalar quantity. At large x ($\gtrsim 0.2$) in the EMC region the nuclear quantities are suppressed relative to the proton. In the intermediate region $x \sim 0.1$ we find “anti-shadowing” where the nuclear results are enhanced. Finally, at smaller x (a few $\times 10^{-2}$) we have the “shadowing” region where the nuclear results are suppressed.

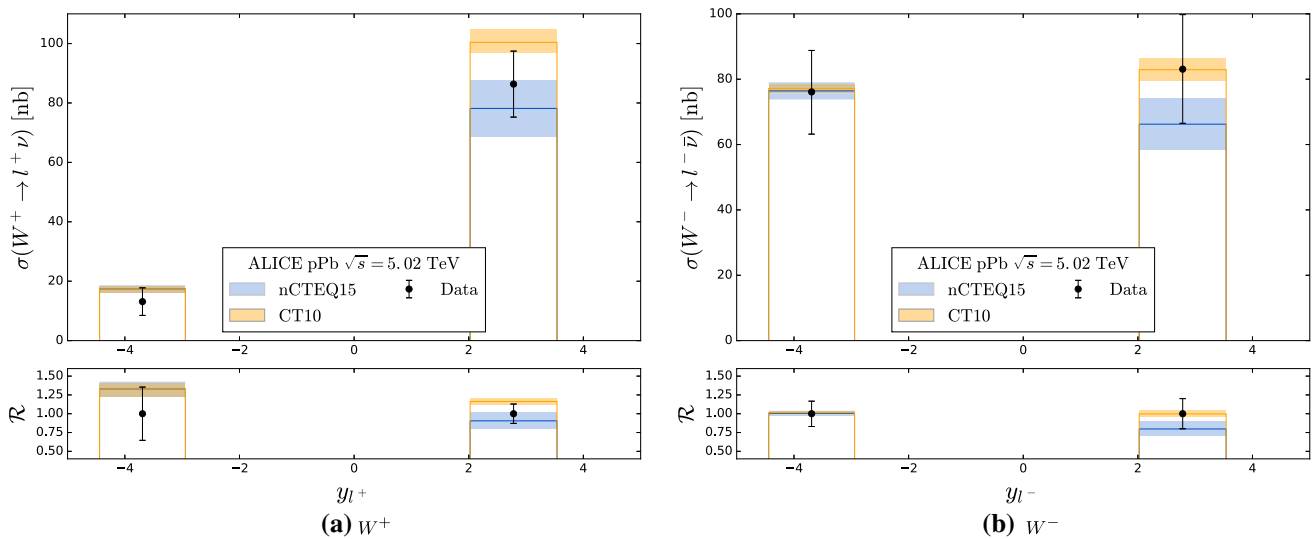


Fig. 8 ALICE W^\pm production in pPb collisions at the LHC

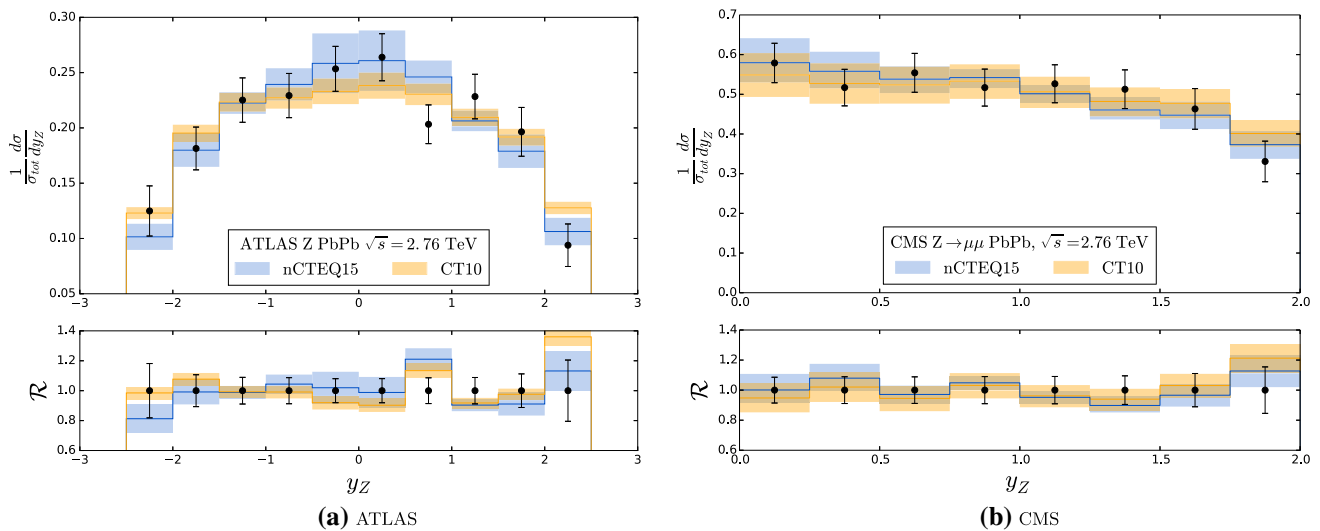


Fig. 9 Z boson production cross section normalized to the total cross section for PbPb collisions at the LHC with $\sqrt{s} = 2.76$ TeV as measured by the ATLAS and CMS collaborations. Corresponding predictions obtained with nCTEQ15 and CT10 PDFs are also shown

extracted from the neutrino-DIS charged current data [27–29]. Taking into account the errors from both the experimental data and the theoretical predictions, no definitive conclusions can be drawn at the present time. Notwithstanding, this data has the potential to strongly influence the nPDF fits, especially in the small x_2 region, if the uncertainties could be reduced.

Finally, the ALICE data (Fig. 8) currently have large uncertainties, and we expect they will have a minimal impact on the reweighting.

2.3 Comparison to lead–lead data

We now consider the LHC PbPb collisions at $\sqrt{s} = 2.76$ TeV. As these beams are symmetric, we now have $y_{CM} = y_{lab}$.

Again, we will use nCTEQ15 [21] and CT10 [24] PDFs for the theoretical predictions. Results from ATLAS and CMS collaborations are available in the form of either event yields (Z boson production) or charge asymmetries (A_ℓ).

In Fig. 9a, b we present the comparison of the ATLAS [7] and CMS [9] data with theoretical predictions with nCTEQ15 and CT10 PDFs. Note that the differential cross sections have been normalized to the total cross section. The PbPb data generally exhibits no tension as the distributions are well described across the kinematical range; however, this is in part due to the large uncertainties due to two nuclei in the initial state.

The measurement of charge asymmetries can provide strong constraints on the PDF fits as many of the systematic

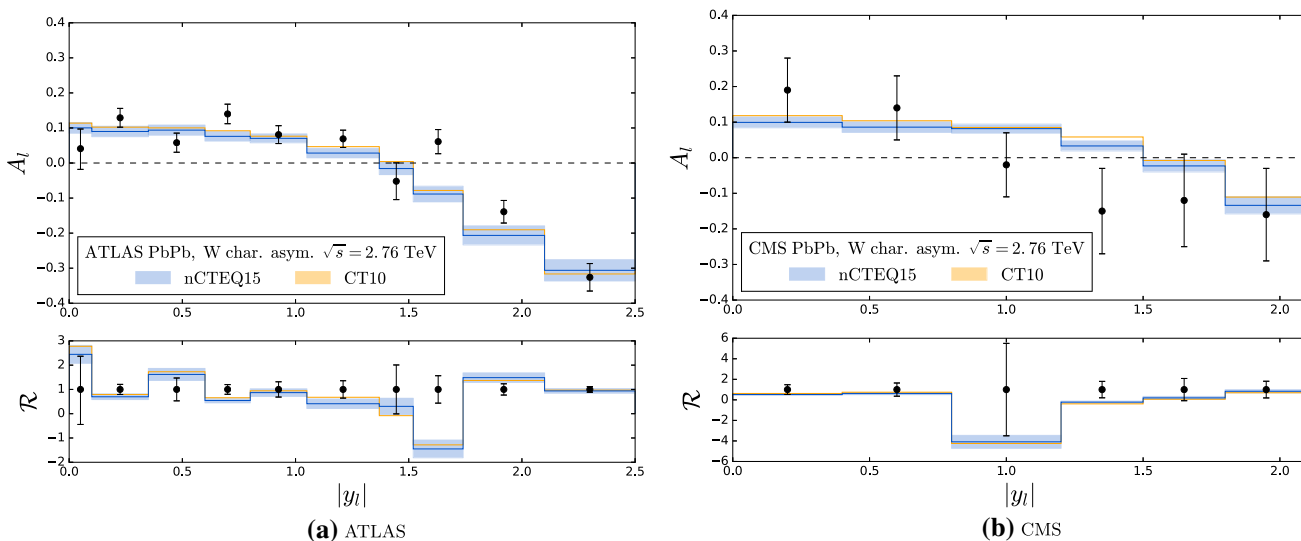


Fig. 10 W charge asymmetry for PbPb collisions at the LHC with $\sqrt{s} = 2.76$ TeV as measured by the ATLAS and CMS collaborations. Corresponding predictions obtained with nCTEQ15 and CT10 PDFs are also shown

uncertainties cancel in such ratios. In Fig. 10 we compute the lepton ($\ell = [\mu, e]$) charge asymmetry $A_\ell(y_\ell)$:

$$A_\ell(y_\ell) = \frac{dN(W^+ \rightarrow \ell^+ \nu_\ell) - dN(W^- \rightarrow \ell^- \bar{\nu}_\ell)}{dN(W^+ \rightarrow \ell^+ \nu_\ell) + dN(W^- \rightarrow \ell^- \bar{\nu}_\ell)} \quad (3)$$

for W^+ and W^- bosons as measured by the ATLAS [8] and CMS [10] experiments. Unfortunately, it appears that the dependence on the nuclear corrections largely cancels out in the ratio as the nuclear nCTEQ15 result is indistinguishable from the CT10 proton result. Hence, these charge asymmetry ratios cannot constrain the nuclear corrections at the present time.

2.4 W^\pm/Z cross-section correlations

In order to analyze our results more quantitatively, it is very useful to look at PDF correlations. In particular, we are interested in assessing the importance of the strange quark in our results. We first review some standard definitions before presenting our analysis.

The definition of the correlation cosine of two PDF-dependent observables X and Y is [30]

$$\begin{aligned} \cos \phi &= \frac{\nabla X \cdot \nabla Y}{\Delta X \Delta Y} \\ &= \frac{1}{4\Delta X \Delta Y} \sum_i (X_i^{(+)} - X_i^{(-)}) (Y_i^{(+)} - Y_i^{(-)}), \end{aligned} \quad (4)$$

where ΔX is the PDF error of the corresponding observable. For the nCTEQ15 PDFs this corresponds to the symmetric error given by

$$\Delta X = \frac{1}{2} \sqrt{\sum_i^N (X_i^{(+)} - X_i^{(-)})^2}. \quad (5)$$

$X_i^{(\pm)}$ is the observable evaluated along the \pm error PDF eigenvector i , and the summation runs over all eigenvector directions.

In our case we are interested in observables $X, Y \in \{\sigma_Z, \sigma_{W^+}, \sigma_{W^-}\}$. Here, we focus on the planes formed by the (W^+, W^-) and the (Z, W^\pm) boson production cross sections to visualize the correlations.

Figure 11 shows the correlations of the W^+ and W^- production cross sections for pPb collisions at the LHC in comparison with the CMS and ATLAS measurements. Similarly, in Fig. 12 we display the results for Z and W^\pm bosons. The results are shown for three different rapidity regions, $y < -1$, $|y| < 1$, $y > 1$, and for several PDFs sets. For the proton side we always use the CT10 PDFs and for the lead side we examine three results: (i) nCTEQ15, (ii) CT10, and (iii) CT10 PDFs supplemented by the nuclear corrections from EPS09 (CT10+EPS09). Finally, the central predictions are supplemented with uncertainty ellipses illustrating correlations between the cross sections. The ellipses are calculated in the following way [30]:

$$\begin{aligned} X &= X_0 + \Delta X \cos \theta, \\ Y &= Y_0 + \Delta Y \cos(\theta + \phi), \end{aligned} \quad (6)$$

where X, Y represent PDF-dependent observables, X_0 (Y_0) is the observable calculated with the central PDF, ΔX (ΔY) is defined in Eq. (5), ϕ is the correlation angle defined in Eq. (4), and θ is a parameter ranging between 0 and 2π .

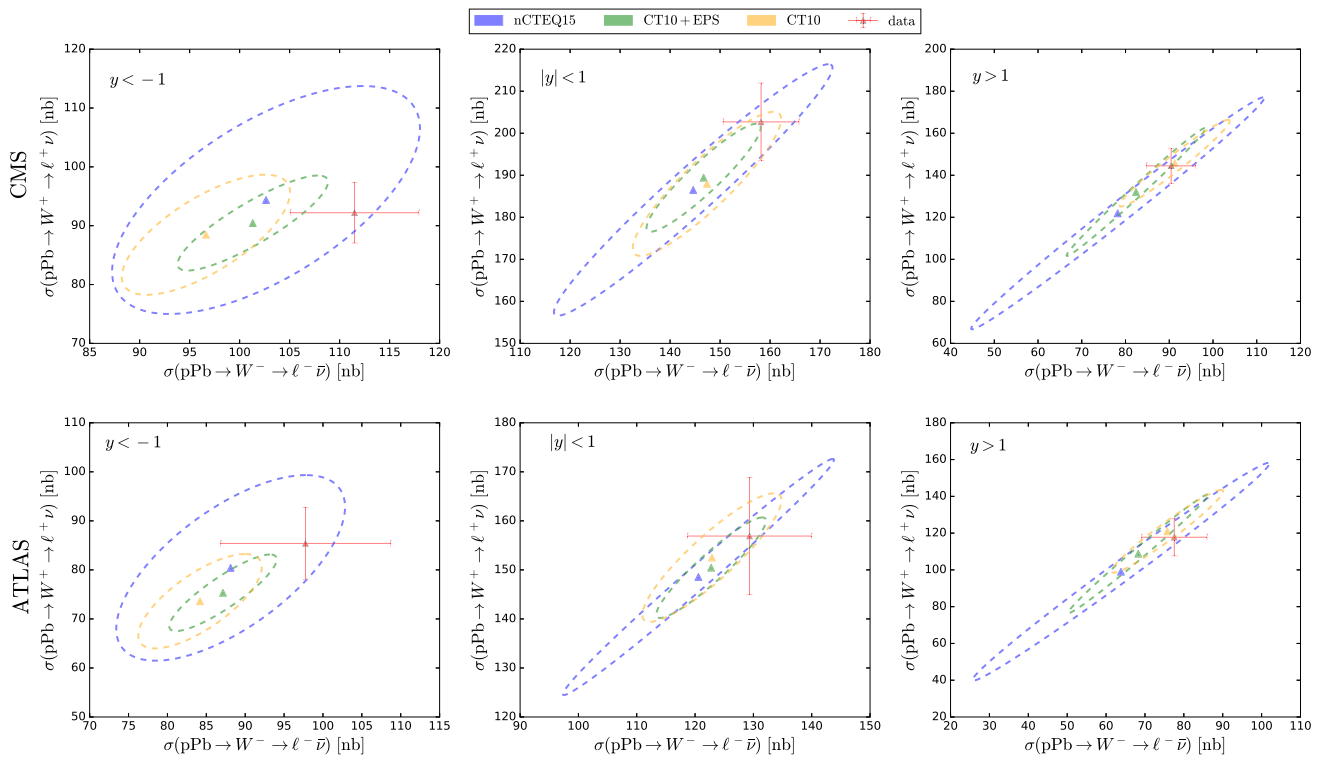


Fig. 11 Correlations between W^+ and W^- cross sections calculated with different PDFs overlaid with the corresponding LHC data from CMS and ATLAS

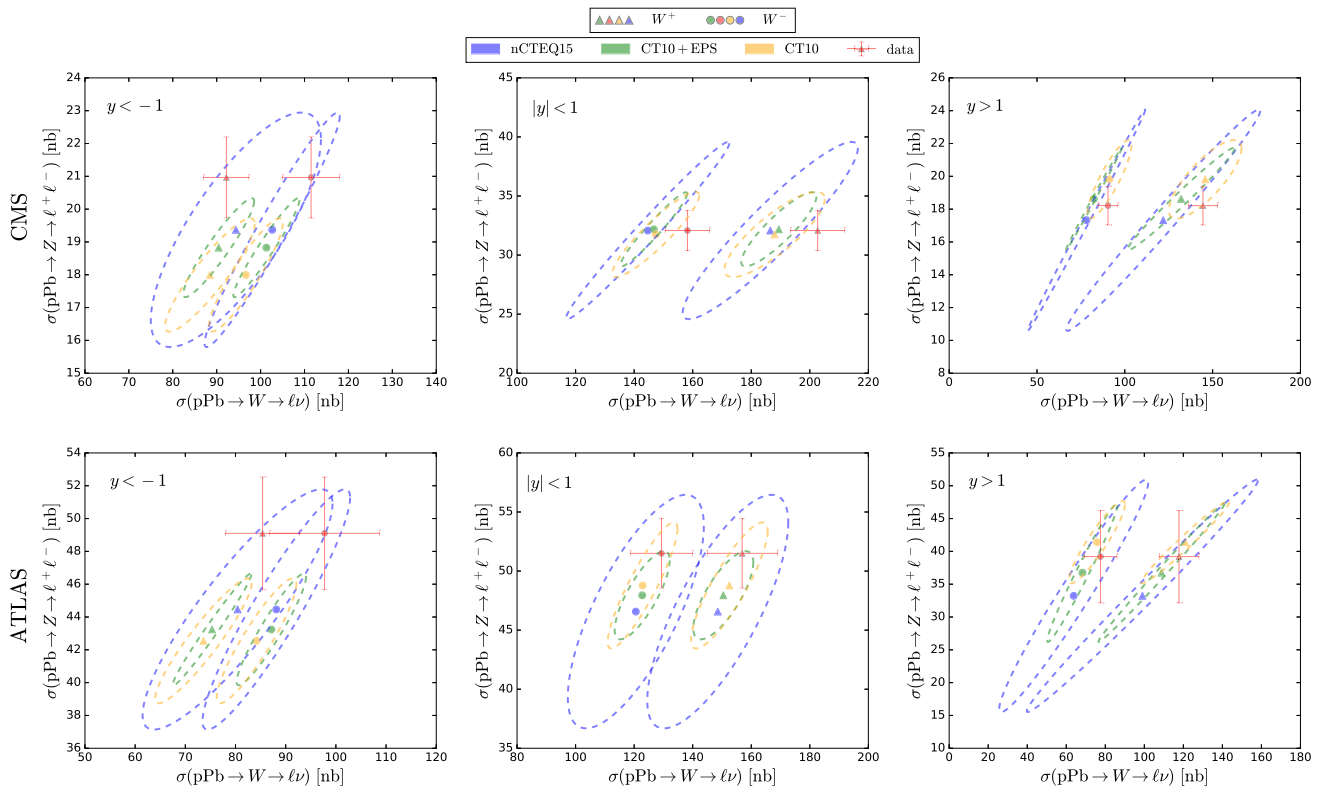


Fig. 12 Correlations between Z and W^+ / W^- cross sections calculated with different PDFs overlaid with the corresponding LHC data from CMS and ATLAS

From Figs. 11 and 12 one can generally observe that the ellipses for the different PDF sets overlap. Furthermore, the central predictions for all three PDF sets lie in the overlapping area of the three ellipses. However, a trend can be observed as a function of the rapidity:

1. For negative rapidities ($y < -1$), the central predictions from the nuclear PDFs (nCTEQ15, EPS09) are closer to the experimental data as they yield larger cross sections than the uncorrected (proton) CT10 PDFs. This can be understood because the lead x_2 values probed in this rapidity bin lie in the region $x_2 \sim 10^{-1}$ where the nPDFs are enhanced due to anti-shadowing (cf., Fig. 9 in Ref. [21]). Due to the larger uncertainties associated with the nCTEQ15 predictions, the ATLAS and CMS cross sections lie within the 1σ ellipse. Conversely, the measured data lie outside the uncorrected (proton) CT10 error ellipsis.
2. For the central rapidity bin ($|y| < 1$), the predictions from all three PDF sets lie generally very close together. In this case, the probed x_2 values lie in the range $0.007 \leq x_2 \leq 0.05$, which is in the transition zone from the anti-shadowing to the shadowing region. We find the LHC W^+ and W^- cross sections in Fig. 11 tend to lie above the theory predictions. Examining the Z cross section of Fig. 12, we find the CMS data agrees closely with the theory predictions, while the ATLAS data is larger by approximately 1σ .
3. For the positive rapidity bin ($y > 1$), we find the central predictions from CT10 match the W^\pm data very closely, but slightly overshoot the Z data. The nuclear PDFs (nCTEQ15, EPS09) undershoot the W^\pm data by a bit more than 1σ , but agree with the Z cross section within 1σ . Here, the probed x_2 values are $\lesssim 0.007$; in this region the lead PDFs are poorly constrained and the corresponding cross sections are dependent on extrapolations of the PDF parameterization in this region.

Interpreting the above set of results appears complicated, so we will try and break the problem down into smaller components. We now compute the same results as above, but using only two flavors (one family) of quarks: $\{u, d\}$; specifically, these plots are produced by zeroing the heavy flavor components (s, c, b), but keeping (u, d) and the gluon. For the Z production this eliminates the $s\bar{s}$ and (the smaller) $c\bar{c}$ contributions, while for W^+/W^- production it is the $\bar{s}c/s\bar{c}$ contribution which drives the change. While the charm PDF does play a role in the above (the bottom contribution is minimal), $c(x)$ is generated radiatively by the process $g \rightarrow c\bar{c}$ (we assume no intrinsic component); thus, it is essentially determined by the charm mass value and the gluon PDF. In contrast, the ‘‘intrinsic’’ nature of the strange PDF leads to its

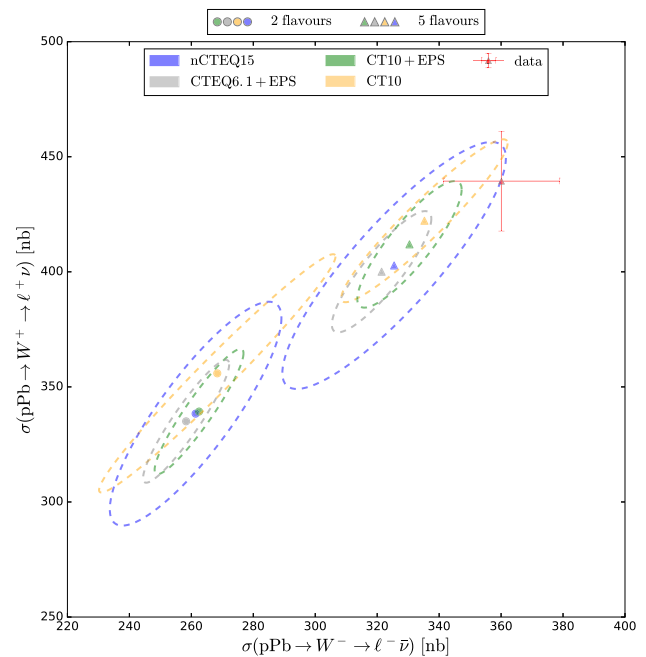


Fig. 13 Comparison of correlations between W^+ and W^- cross sections for the case when only one family of quarks $\{u, d\}$ is included and when all families are accounted for. We show here results for nCTEQ15, EPS09+CT10, EPS09+CTEQ6.1 and CT10 PDFs overlaid with the CMS data

comparably large uncertainties. For example, if we compare the free-proton PDF baselines (CTEQ6.1, CT10), the strange quark exhibits substantial differences while the charm (and bottom) distributions are quite similar; this pattern then feeds into the nPDFs. Therefore, the strange quark PDF will be the primary focus of the following discussion.

In Figs. 13 and 14 we compare the five flavor and two flavor results using nCTEQ15, CT10+EPS09, CTEQ6.1+EPS09, and CT10. We have added CTEQ6.1+EPS09 as CTEQ6.1 was the baseline used for the EPS09 fit.

Examining Fig. 13, the shift of the two flavor results compared to the five flavor results can be as large as $\sim 30\%$ and reflects, the contributions of the strange and charm quarks.

For the five flavor case (\blacktriangle), the calculations are scattered to the low side of the data in both W^+ and W^- . The CT10 result is the closest to the data, but due to the larger uncertainties of nCTEQ15, the data point is within range of both of their ellipses. We also observe that the CT10+EPS09 and CTEQ6.1+EPS09 results bracket the nCTEQ15 value; again, this is due to the very different strange PDF associated with CT10 and CTEQ6.1.

For the two flavor case (\bullet), all the nuclear results (nCTEQ15, CT10+EPS09, CTEQ6.1+EPS09) coalesce, and they are distinct from the non-nuclear result (CT10). This pattern suggests that the nuclear corrections of nCTEQ15 and EPS09 for the $\{u, d\}$ flavors are quite similar, and the spread observed in the five flavor case comes from differ-

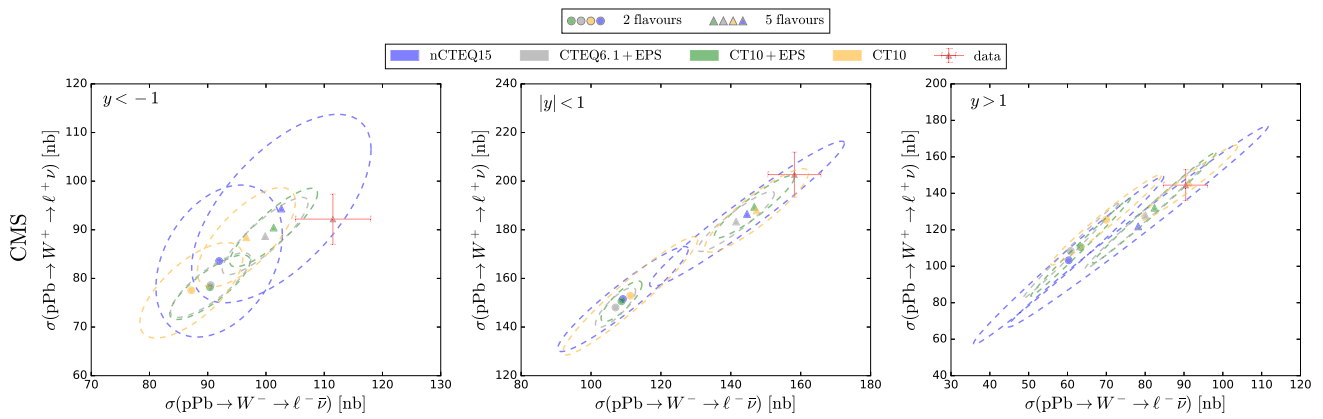


Fig. 14 Same as Fig. 13 but divided into rapidity bins

ences of $s(x)$ in the underlying base PDF. Thus we infer that the difference between the nuclear results and the proton result accurately represents the nuclear corrections for the two flavor case (for $\{u, d\}$), but for the five flavor case it is a mix of nuclear corrections and variations of the underlying sea quarks.

Figure 14 displays the same information as Fig. 13 except it is divided into rapidity bins. As we move from negative y to positive y we move from high x where the nPDFs are well constrained to small x where the nPDFs have large uncertainties (*cf.*, Fig. 2). Thus, it is encouraging that at $y < -1$ we uniformly find the nuclear predictions yield larger cross sections than the proton results (without nuclear corrections) and thus lie closer to the LHC data.

Conversely, for $y > 1$ we find the nuclear predictions yield smaller cross sections than the proton results. The comparison with the LHC data varies across the figures, but this situation suggests a number of possibilities.

First, the large nPDF uncertainties in this small x_2 region could be improved using the LHC data.

Second, the lower nPDF cross sections are partly due to the nuclear shadowing in the small x region; if, for example, this shadowing region were shifted to even lower x values, this would increase the nuclear results. Such a shift was observed in Refs. [27–29] using charged current neutrino-DIS data, and this would move the nuclear predictions of Fig. 11 at $y > 1$ toward the LHC data.

Finally, we note that measurements of the strange quark asymmetry [31] indicate that $s(x) \neq \bar{s}(s)$ which is unlike what is used in the current nPDFs; this would influence the W^\pm/Z cross sections separately as (at leading order) [26] $W^+ \sim \bar{s}c$, $W^- \sim s\bar{c}$, and $Z \sim \bar{s}s$. As the strange PDF has a large impact on the W^\pm/Z measurements, this observation could provide incisive information on the individual s and \bar{s} distributions.

These points are further exemplified in Fig. 15 which displays W^\pm production for both two and five flavors as a function of lepton rapidity y_{ℓ^\pm} . For large y_{ℓ^\pm} , (small lead x_2)

the CT10 proton result separates from the collective nuclear results; presumably, this is due to the nuclear shadowing at small x_2 . Again, we note that in this small x_2 region there are minimal experimental constraints and the nPDFs come largely from extrapolation at higher x_2 values. Additionally, by comparing the two and five flavor results, we clearly see the impact of the heavier flavors, predominantly the strange quark PDF.

Furthermore, different strange quark PDFs in the baseline PDFs compared in Figs. 11 and 12, make it challenging to distinguish nuclear effects from different strange quark distributions. Thus, we find that the extraction of the nuclear corrections is intimately intertwined with the extraction of the proton strange PDF, and we must be careful to separately distinguish each of these effects. Fortunately, the above observations can help us disentangle these two effects.

3 Reweighting

In this section we perform a reweighting study to estimate the possible impact of the W^\pm/Z data on nCTEQ15 lead PDFs. For this purpose we will use only the pPb data sets.

We refrain from using PbPb data as typically the agreement of these data with current nPDFs is much better (in part due to the large uncertainties), so the impact in the reweighting analysis will be minimal. Secondly the factorization in lead–lead collisions is not firmly established theoretically [32], so that the interpretation may be complicated.

3.1 Basics of PDF reweighting

In this section we summarize the PDF reweighting technique and provide formulas for our specific implementation of this method. Additional details can be found in the literature [33–37].

In preparation for the reweighting, we need to convert the nCTEQ15 set of Hessian error PDFs into a set of PDF repli-

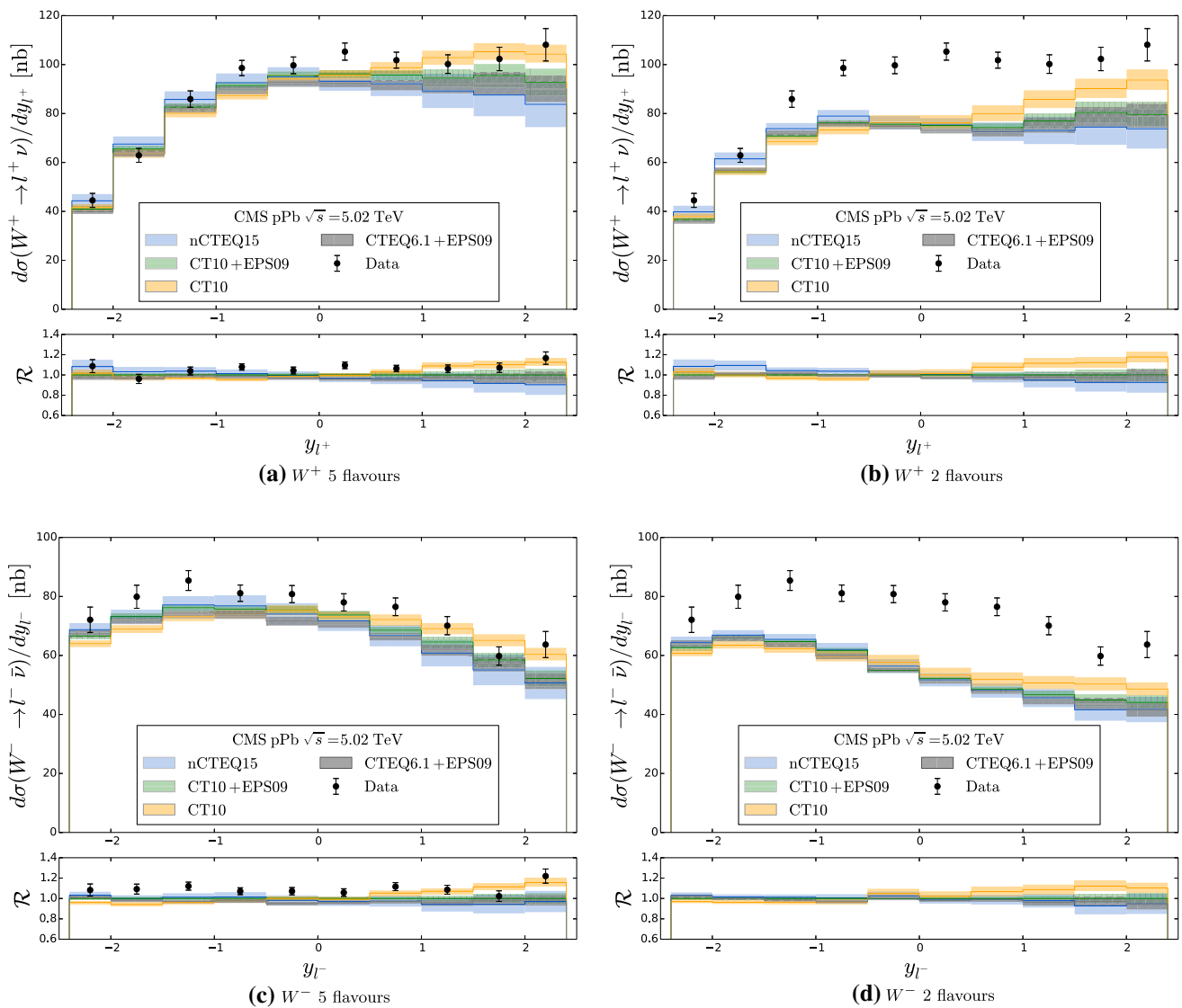


Fig. 15 Rapidity distributions for W^\pm cross sections measured by CMS compared with predictions from the nCTEQ15, EPS09+CT10, EPS09+CTEQ6.1 and CT10 PDFs. Figures **a** and **c** show the results for five flavors, while Figs. **b** and **d** show the two flavors results

cas [12,38] which serve as a representation of the underlying probability distribution. The PDF replicas can be defined by a simple formula,⁵

$$f_k = f_0 + \sum_{i=1}^N \frac{f_i^{(+)} - f_i^{(-)}}{2} R_{ki}, \tag{7}$$

where f_0 represents the best fit (central) PDF, $f_i^{(+)}$ and $f_i^{(-)}$ are the plus and minus error PDFs corresponding to the eigenvector direction i , and N is the number of eigenvectors defining the Hessian error PDFs. Finally, R_{ki} is a random number from a Gaussian distribution centered at 0 with standard devi-

⁵ A detailed discussion of the construction of replicas from Hessian PDF sets in the case of asymmetric errors can be found in Ref. [39].

ation of 1, which is different for each replica (k) and each eigen-direction (i).

After producing the replicas, we can calculate the average and variance of any PDF-dependent observable as moments of the probability distribution:

$$\langle \mathcal{O} \rangle = \frac{1}{N_{\text{rep}}} \sum_{k=1}^{N_{\text{rep}}} \mathcal{O}(f_k),$$

$$\delta \langle \mathcal{O} \rangle = \sqrt{\frac{1}{N_{\text{rep}}} \sum_{k=1}^{N_{\text{rep}}} (\mathcal{O}(f_k) - \langle \mathcal{O} \rangle)^2}. \tag{8}$$

In particular, it can be done for the PDFs themselves; we should be able to reproduce our central PDF f_0 by the average $\langle f \rangle$, and the (68% c.l.) Hessian error bands $\Delta f =$

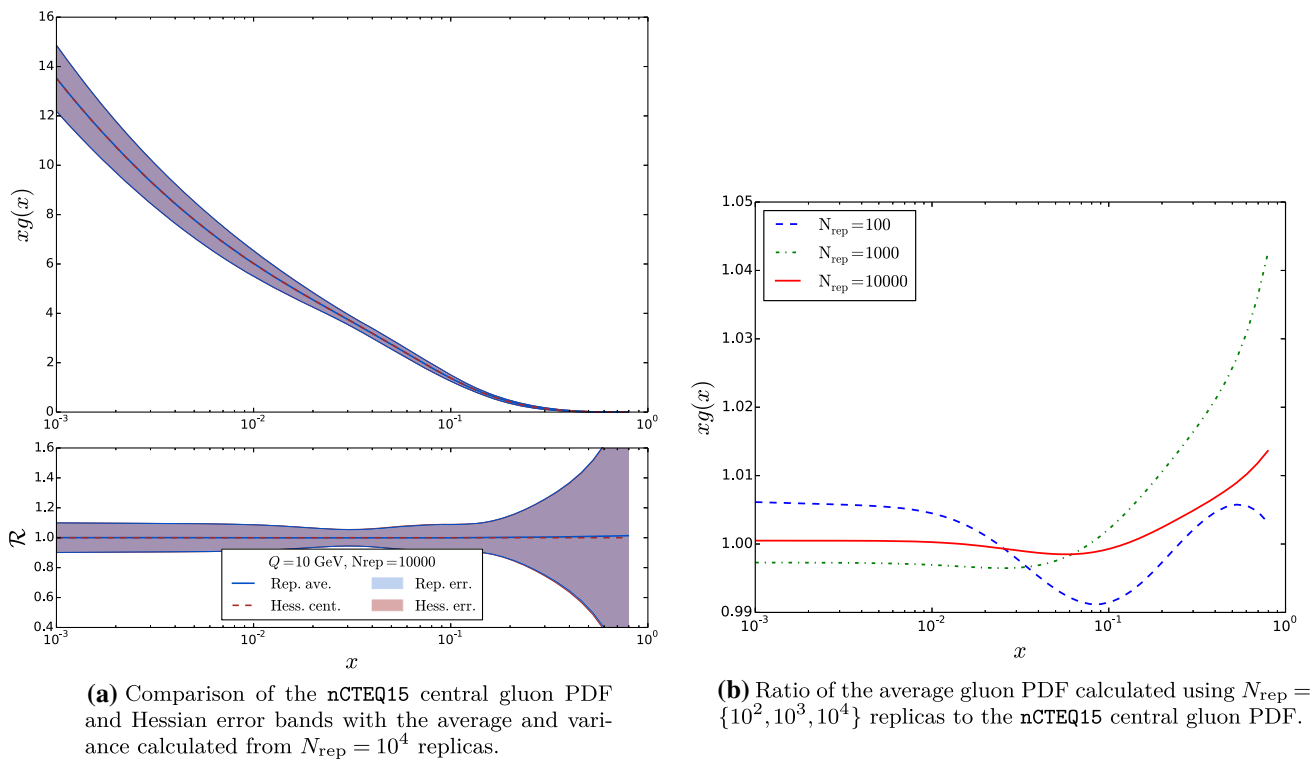


Fig. 16 Comparison of the nCTEQ15 Hessian gluon distribution and its reproduction in terms of replicas at a scale of $Q = 10 \text{ GeV}$

$\frac{1}{2} \sqrt{\sum_i^N (f_i^{(+)} - f_i^{(-)})^2}$ by the corresponding variance $\delta \langle f \rangle$. Of course, the precision at which we are able to reproduce Hessian central PDFs and corresponding uncertainties depends on how well we reproduce the underlying probability distribution, and this will depend on the number of replicas, N_{rep} , we use. In the following we use $N_{\text{rep}} = 10^4$, which allows for a very good reproduction of both central and error PDFs (within $\sim 0.1\%$ or better).

We note here that since the nCTEQ15 error PDFs correspond to the 90% confidence level (c.l.) we need to convert the obtained uncertainties such that they correspond to the 68% c.l.⁶ The conversion is done using the following approximate relation between the 68% c.l. and 90% c.l. Hessian uncertainties: $\Delta_{90}^H \mathcal{O} \approx 1.645 \Delta_{68}^H \mathcal{O}$.

In Fig. 16 we perform the above exercise and determine if our procedure is self consistent. Specifically, in Fig. 16a we display the central value and uncertainty bands for the original gluon PDF and those generated from the replicas; they are indistinguishable. Additionally, in Fig. 16b we demonstrate the convergence of the average of replicas to the central Hessian PDF for $N_{\text{rep}} = \{10^2, 10^3, 10^4\}$. For $N_{\text{rep}} = 10^4$ the central gluon is reproduced to better than 1% except at the highest x values. This is certainly a sufficient accuracy considering the size of the PDF

errors. Even the $N_{\text{rep}} = 10^2$ and $N_{\text{rep}} = 10^3$ replicas yield good results except at larger x ($\gtrsim 0.1$) where the PDFs are vanishing and the uncertainties are large. Since our computational cost will be mostly dictated by the number of Hessian error PDFs, we will use $N_{\text{rep}} = 10^4$ to get a better representation of the underlying probability distribution.

Having defined the replicas we can apply the reweighting technique to estimate the importance of a given data set on our current PDFs. The idea is based on Bayes theorem, which states that the posterior distribution representing the probability of a hypothesis (new probability distribution representing the PDFs if we would perform a fit including the new data set we are using in the reweighting) is a product of the prior probability (PDFs without the new data set) and an appropriate likelihood function. This allows us to assign a weight to each of the replicas generated earlier according to Eq. (7).

In the context of Hessian PDFs using a global tolerance criterion the appropriate weight definition is given by a modified Giele–Keller expression [12,33,36,37]⁷

$$w_k = \frac{e^{-\frac{1}{2} \chi_k^2 / T}}{\frac{1}{N_{\text{rep}}} \sum_i^{N_{\text{rep}}} e^{-\frac{1}{2} \chi_i^2 / T}}, \tag{9}$$

⁶ The 68% c.l. is necessary to correspond with the variance of the PDF set defined below.

⁷ In the context of Monte Carlo PDF sets a NNPDF weight definition should be used [35].

where T is the tolerance criterion used when defining Hessian error PDFs⁸ and χ_k^2 represents the χ^2 of the data sets considered in the reweighting procedure for a given replica k . The pPb W^\pm and Z data do not provide correlated errors (the published errors are a sum of statistical and systematic errors in quadrature)⁹ so it is sufficient for our analysis to use a basic definition of the χ^2 function given by

$$\chi_k^2 = \sum_j \frac{(D_j - T_j^k)^2}{\sigma_j^2}, \tag{10}$$

where index j runs over all data points in the data set(s), N_{data} is the total number of data points, D_j is the experimental measurement at point j , σ_j is the corresponding experimental uncertainty, and T_j^k is the corresponding theoretical prediction calculated with PDFs given by replica k .

With the above prescription we can now calculate the weights needed for the reweighting procedure. The expectation value and variance of any PDF-dependent observable can now be computed in terms of weighted sums:

$$\begin{aligned} \langle \mathcal{O} \rangle_{\text{new}} &= \frac{1}{N_{\text{rep}}} \sum_{k=1}^{N_{\text{rep}}} w_k \mathcal{O}(f_k), \\ \delta \langle \mathcal{O} \rangle_{\text{new}} &= \sqrt{\frac{1}{N_{\text{rep}}} \sum_{k=1}^{N_{\text{rep}}} w_k (\mathcal{O}(f_k) - \langle \mathcal{O} \rangle_{\text{new}})^2}. \end{aligned} \tag{11}$$

For our reweighting analysis we will only use the pPb data sets. Because the uncertainty of the nuclear PDFs dominates the proton PDFs, it is sufficient to only vary the lead PDFs. Consequently, the pPb cross sections are linear in the lead uncertainties, and we can compute the reweighting by evaluating cross sections only on the Hessian error PDFs (32+1 in the case of nCTEQ15) instead of the individual replicas ($N_{\text{rep}} = 10^4$)

$$\sigma_k = f^{\text{P}} \otimes \hat{\sigma} \otimes \left[f_0^{\text{Pb}} + \sum_i \frac{f_i^{\text{Pb}(+)} - f_i^{\text{Pb}(-)}}{2} R_{ki} \right]. \tag{12}$$

A similar decomposition can be used for pp or PbPb data to reduce the number of necessary evaluations. However, because of the quadratic dependence on the PDFs, the reduction is smaller and does not necessarily lead to lower computational costs.

We will compare the χ^2 for each experiment calculated with the initial PDFs (before reweighting) and with the PDFs

⁸ In the case of the nCTEQ15 PDFs, the tolerance criterion is $T = 35$, which corresponds to a 90% c.l., the detailed explanation on how it was defined can be found in appendix A of [21]. The tolerance factor used in this analysis corresponds to the 68% c.l. which we obtain by rescaling the above: $T \approx 35/1.645^2 \sim 13$.

⁹ In our analysis we also add the normalization errors in quadrature to the statistical and systematic ones.

after the reweighting procedure; this will allow us to estimate the impact of each individual data set. We do this using the following formula:

$$\chi^2 = \sum_j \frac{(D_j - \langle T_j \rangle)^2}{\sigma_j^2}, \tag{13}$$

where $\langle T_j \rangle$ is a theory prediction calculated as an average over the (reweighted or not-reweighted) replicas according to Eq. (11) (with or without weights).

Finally, the effectiveness of the reweighting procedure can be (qualitatively) estimated by computing the *effective number of replicas* defined as [35]:

$$N_{\text{eff}} = \exp \left[\frac{1}{N_{\text{rep}}} \sum_{k=1}^{N_{\text{rep}}} w_k \ln(N_{\text{rep}}/w_k) \right]. \tag{14}$$

N_{eff} provides a measure of how many of the replica sets are effectively contributing to the reweighting procedure. By definition, N_{eff} is restricted to be smaller than N_{rep} . However, when $N_{\text{eff}} \ll N_{\text{rep}}$ it indicates that there are many replicas whose new weight (after the reweighting procedure) is sufficiently small for them to provide a negligible contribution to the updated probability density. This typically happens when the new data is not compatible with the data used in the original fit, or if the new data introduces substantial new information; in both cases, the procedure becomes ineffective and a new global fit is recommended.

3.2 Reweighting using CMS W^\pm rapidity distributions

As an example, we consider the reweighting using the CMS W^\pm production data from pPb collisions [4]. In this example we use rapidity distributions of charged leptons originating from the decay of both W^+ and W^- bosons with $N_{\text{rep}} = 10^4$ replicas leading to $N_{\text{eff}} = 5913$.

In Fig. 17 we display the distribution of the weights obtained from the reweighting procedure. We see that the magnitudes of the weights are reasonable; they extend up to ~ 9 with a peak at the lowest bin. It will be useful to compare this distribution with later results using different observables and data sets.

In Fig. 18 we show the comparison of the data to theory before and after the reweighting procedure.¹⁰ As expected, we see that after the reweighting procedure the description of

¹⁰ We note here the difference of PDF uncertainties compared to the plots presented in Sect. 2; this is caused by the fact that now we use the 68% c.l. errors whereas in Sect. 2 we have used the 90% c.l. errors that are provided with the nCTEQ15 PDFs. This holds for all plots in Sect. 3.

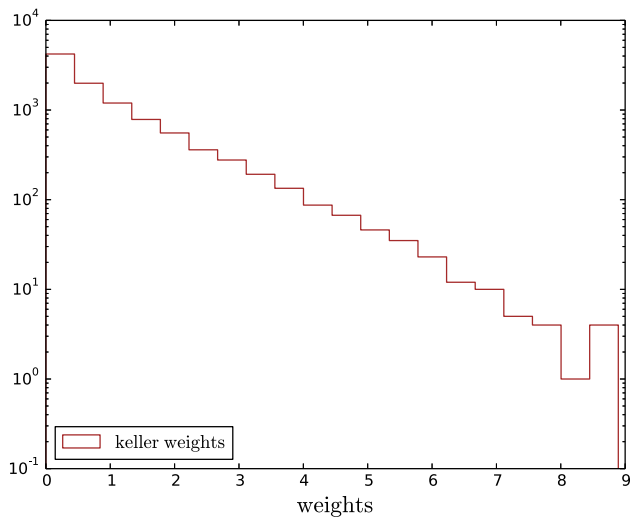


Fig. 17 Weight distribution after reweighting using rapidity distributions of charged leptons from CMS W^\pm production data

the data is improved. This is true for both the W^+ (left figure) and the W^- (right figure) cases. We can quantify the improvement of the fit by examining the χ^2/N_{data} for the individual distributions. For the W^+ case, the χ^2/N_{data} is improved from 5.07 before reweighting to 3.23 after reweighting. Similarly, for W^- the χ^2/N_{data} is improved from 4.57 to 3.44. The amount of change due to the reweighting procedure should be proportional to the experimental uncertainties of the incorporated data; this is the same as we would expect from a global

fit. For W^\pm production investigated here, the uncertainties are quite substantial, and the effects are compounded by the lack of correlated errors.

Finally, we show the effect of the reweighting on the PDFs themselves. In Fig. 19, we display PDFs for the up quark and gluon at a scale of $Q = 10\text{ GeV}$. We can see that the reweighting has the largest effects in the low x region, and this holds also for the other flavors as well. Generally the effects at intermediate and large x values are limited, with the exception of the gluon which is poorly constrained and exhibits a substantial change for large x .

In Figs. 18 and 19, in addition to the reweighting results, we also show results calculated using the Hessian profiling method [37]. The Hessian profiling should agree precisely with our reweighting calculations, and this can serve as an independent cross-check of our results. Indeed, in the figures we observe that the profiling exactly matches the reweighted results. In the following figures we will display only the reweighting results, but in all presented cases we have checked that these two methods agree.

3.3 Using asymmetries instead of differential cross sections

In this section we will re-investigate the reweighting analysis from the previous section employing the CMS W^\pm production data. Instead of using rapidity distributions (as in the previous section), we will use two types of asymmetries which

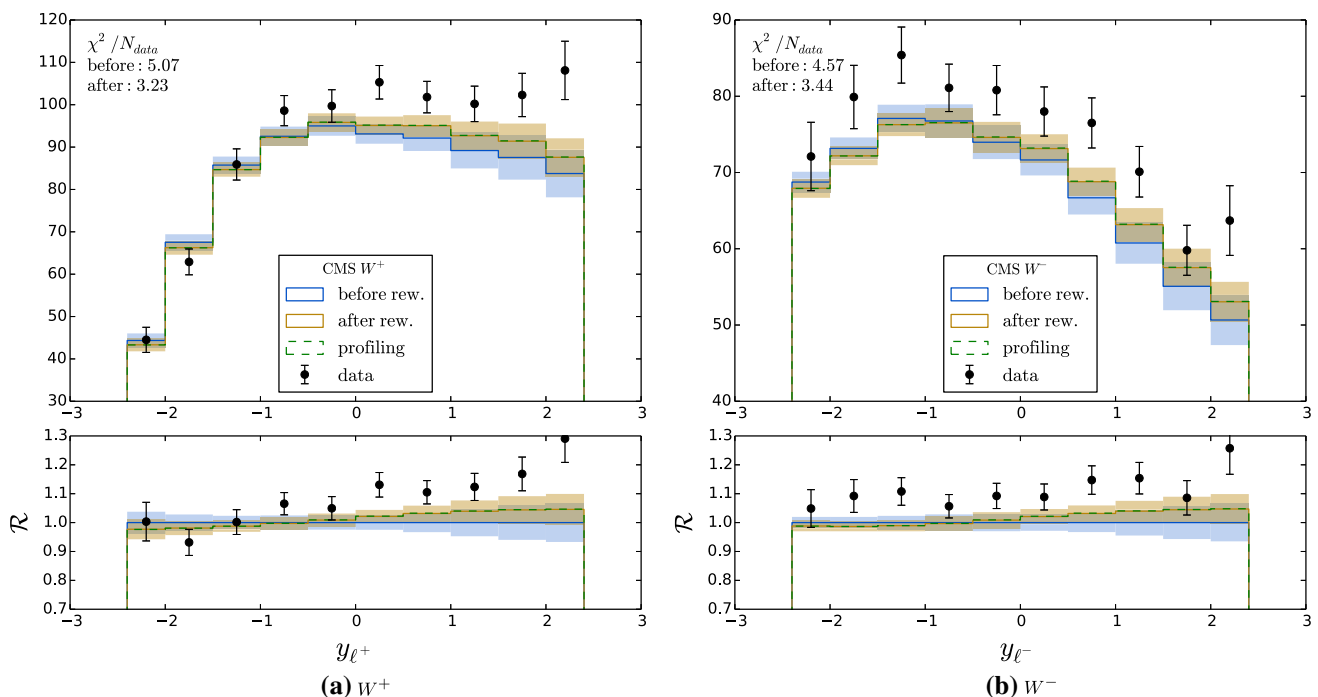


Fig. 18 Comparison of data and theory before and after the reweighting procedure for the rapidity distributions of charged leptons in W^\pm production measured by CMS [4]

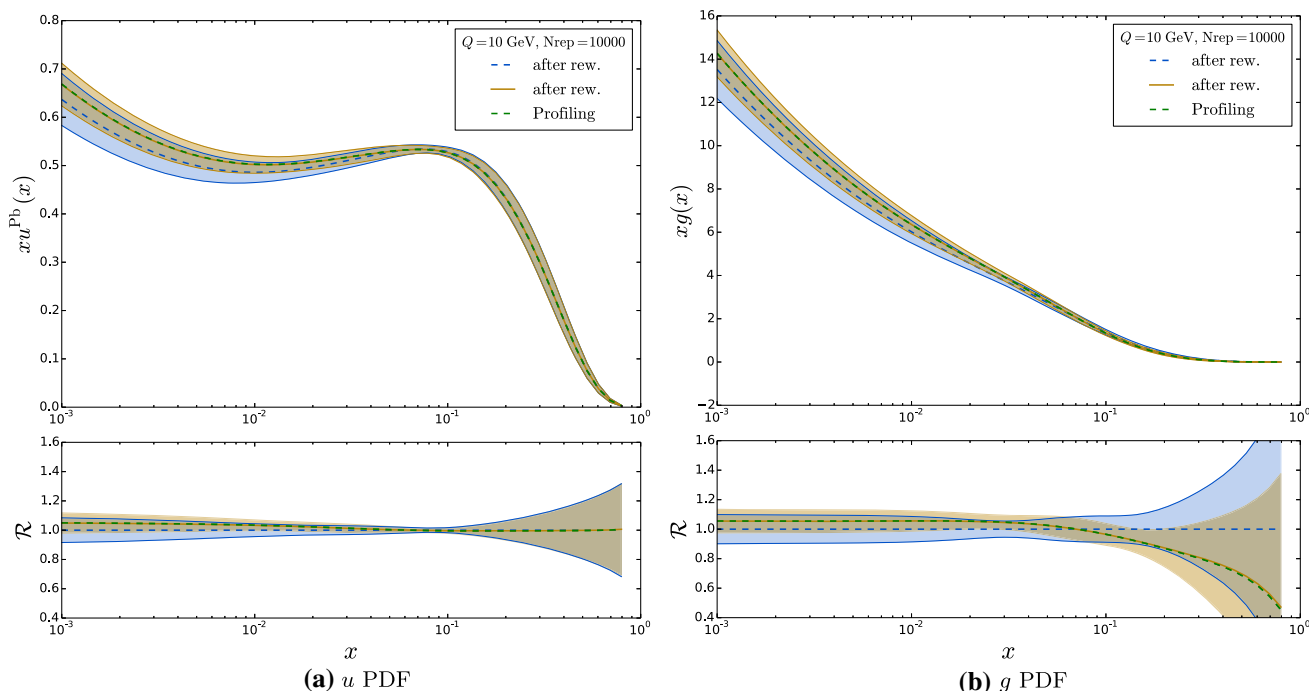


Fig. 19 Comparison of PDFs at $Q = 10$ GeV before and after the reweighting procedure using only the CMS rapidity distribution data set. The lower plots show the ratio compared to the central (average) distribution before the reweighting

are constructed with the charged leptons. The lepton charge asymmetry is

$$A_\ell(y_\ell) = \frac{dN_{l^+} - dN_{l^-}}{dN_{l^+} + dN_{l^-}}, \tag{15}$$

and is defined per bin in the rapidity of the charged lepton where N_{l^\pm} represents the corresponding number of observed events in a given bin. For the purpose of the theory calculation, N_{l^\pm} will be replaced by the corresponding cross section in a given bin.

It is useful to consider the expression for the charge asymmetry at leading order in the parton model assuming a diagonal CKM matrix:

$$A_\ell = \frac{u(x_1)\bar{d}(x_2) + \bar{d}(x_1)u(x_2) + c(x_1)\bar{s}(x_2) + \bar{s}(x_1)c(x_2) - \bar{u}(x_1)d(x_2) - d(x_1)\bar{u}(x_2) - \bar{c}(x_1)s(x_2) - s(x_1)\bar{c}(x_2)}{u(x_1)\bar{d}(x_2) + \bar{d}(x_1)u(x_2) + c(x_1)\bar{s}(x_2) + \bar{s}(x_1)c(x_2) + \bar{u}(x_1)d(x_2) + d(x_1)\bar{u}(x_2) + \bar{c}(x_1)s(x_2) + s(x_1)\bar{c}(x_2)}. \tag{16}$$

Here, the partons with momentum fraction x_1 are in the proton, and those with momentum fraction x_2 are inside the lead. At large negative rapidities (small x_1 , large x_2), we have $f(x_1) = \bar{f}(x_1)$ for all parton flavors ($f = u, d, s, c$) and the expression for the asymmetry simplifies to the following form:

$$A_\ell \rightarrow \frac{d(x_1)u_v(x_2) - u(x_1)d_v(x_2) - c(x_1)s_v(x_2) + s(x_1)c_v(x_2)}{d(x_1)u_v(x_2) + u(x_1)d_v(x_2) + c(x_1)s_v(x_2) + s(x_1)c_v(x_2)}. \tag{17}$$

Assuming $c_v(x_2) = c(x_2) - \bar{c}(x_2) = 0$ and $s_v(x_2) = s(x_2) - \bar{s}(x_2) = 0$, as it is the case in all the existing nPDF sets, the expression further simplifies

$$A_\ell \rightarrow \frac{d(x_1)u_v(x_2) - u(x_1)d_v(x_2)}{d(x_1)u_v(x_2) + u(x_1)d_v(x_2)} \simeq \frac{u_v(x_2) - d_v(x_2)}{u_v(x_2) + d_v(x_2)}. \tag{18}$$

In the last equation, we have used the fact that the small x_1 up and down PDFs are very similar [40]. Since the $d_v^{Pb}(x_2) > u_v^{Pb}(x_2)$, we expect the asymmetry to be negative at large negative rapidities. One can also observe that the asymmetry calculated with either $n_f = 2$ or $n_f = 5$ will be the same.

A non-zero strange asymmetry ($s(x_2) > \bar{s}(x_2)$) would lead to a decrease of the A_ℓ asymmetry, thereby improving the description of the CMS data.

Conversely, at large positive rapidities (small x_2 , large x_1), we have $f(x_2) = \bar{f}(x_2)$ for all parton flavors ($f = u, d, s, c$) and the expression for the asymmetry becomes

$$A_\ell \rightarrow \frac{u_v(x_1)d(x_2) - d_v(x_1)u(x_2) + c_v(x_1)s(x_2) - s_v(x_1)c(x_2)}{u_v(x_1)d(x_2) + d_v(x_1)u(x_2) + c_v(x_1)s(x_2) + s_v(x_1)c(x_2)}. \tag{19}$$

Again, assuming $c(x_1) = \bar{c}(x_1)$ and $s(x_1) = \bar{s}(x_1)$, this expression further simplifies to

$$A_\ell \rightarrow \frac{u_v(x_1)d(x_2) - d_v(x_1)u(x_2)}{u_v(x_1)d(x_2) + d_v(x_1)u(x_2)} \simeq \frac{u_v(x_1) - d_v(x_1)}{u_v(x_1) + d_v(x_1)}, \tag{20}$$

where we have again used $u(x_2) \simeq d(x_2)$ at small x_2 . Since $u_v(x_1) > d_v(x_1)$ in the proton, we expect a positive asymmetry in the kinematic region of large positive rapidities. Furthermore, the reweighting of the nuclear PDFs will have very little impact on the charge asymmetry in this limit even if the precision of the data will increase in the future.

Another asymmetry used by CMS is the forward–backward asymmetry. This is defined as a ratio of the number of events in the forward and backward region in a given rapidity bin:

$$A_{\text{FB}}^\pm(y_\ell) = \frac{dN_{l^\pm}(+y_{\text{lab}})}{dN_{l^\pm}(-y_{\text{lab}})}. \tag{21}$$

This asymmetry is defined separately for the W^+ and W^- cases. It can also be combined into a single quantity, the forward–backward asymmetry of charge-summed W bosons:

$$A_{\text{FB}}(y_\ell) = \frac{dN_{l^+}(+y_{\text{lab}}) + dN_{l^-}(+y_{\text{lab}})}{dN_{l^+}(-y_{\text{lab}}) + dN_{l^-}(-y_{\text{lab}})}. \tag{22}$$

This is the quantity we will use for our analysis in this section.

We now use the asymmetries of Eqs. (15) and (22) to perform a reweighting of the nCTEQ15 lead PDFs. These asymmetries are just combinations of the rapidity distributions used in Sect. 3.2, and if both are employed at the same time they should encode similar information to the rapidity distributions themselves. In the literature it is sometimes argued that the asymmetries are more sensitive to the PDFs and in turn are better suited to performing PDF fits [4, 11, 12]. We will empirically check this statement by comparing reweighting predictions using rapidity distributions and the above mentioned asymmetries.

In the following, we present the results of the reweighting using the lepton charge asymmetry and forward–backward asymmetry of charge-summed W bosons. In this case, the effective number of replicas is $N_{\text{eff}} = 7382$.

The distribution of weights is displayed in Fig. 20, and we can see that compared to the reweighting using directly the rapidity distributions (Fig. 17), the weights are smaller extending only to around ~ 2.7 and more evenly distributed.

In Fig. 21 we show a comparison of data and theory before and after the reweighting procedure. In the case of the charge asymmetry we do not see a large improvement, but this is not surprising as there is already good agreement between the data and theory before the reweighting. We note that the

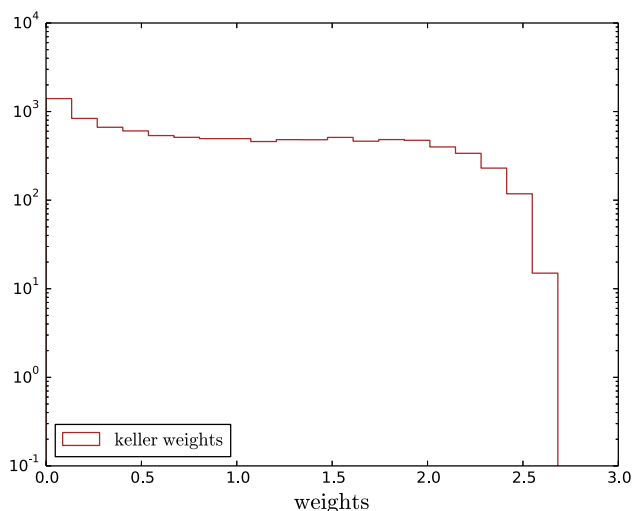


Fig. 20 Weight distribution after reweighting using the lepton charge asymmetry and forward–backward asymmetry from CMS W^\pm production data

χ^2/N_{data} before the reweighting is 1.44 and 1.27 after the reweighting.

In the case of the forward–backward asymmetry the initial agreement between data and theory is not as good and the corresponding improvement is much larger; χ^2/N_{data} changes from 4.03 to 1.31.

We now show the effect of the reweighting procedure on the PDFs. In Fig. 22 we display the PDFs for the up quark and gluon at a scale of $Q = 10 \text{ GeV}$. We can see that in both cases the effect is limited to the low x region and does not exceed few percent. The results for other flavors are similar, and overall the asymmetries with the current experimental uncertainties seem to have rather small effect on the nPDFs.

In particular it seems that using asymmetry ratios yields a reduced impact, at least compared to the rapidity distributions of Sect. 3.2. This is possibly due to the fact that much of the information on the nuclear corrections is lost when constructing the ratios. However, asymmetries can be still useful to explore the very forward and backward regions of the rapidity distributions (corresponding to higher/lower x values) where experimental uncertainties are typically large but can cancel in the ratios.

3.4 Including all the data sets

Due to large experimental uncertainties, the effect of individual data sets presented in Sect. 2 on the lead PDFs is rather limited. The largest constraint is obtained from the CMS W^\pm data [4] (Sect. 3.2), and from (preliminary) ATLAS W^\pm data [2]. In order to maximize the effects on the PDFs, we now employ all proton–lead data sets from Table 1 to perform the reweighting of the nCTEQ15 lead PDFs. Note that we use both the rapidity distributions and the asymmetries;

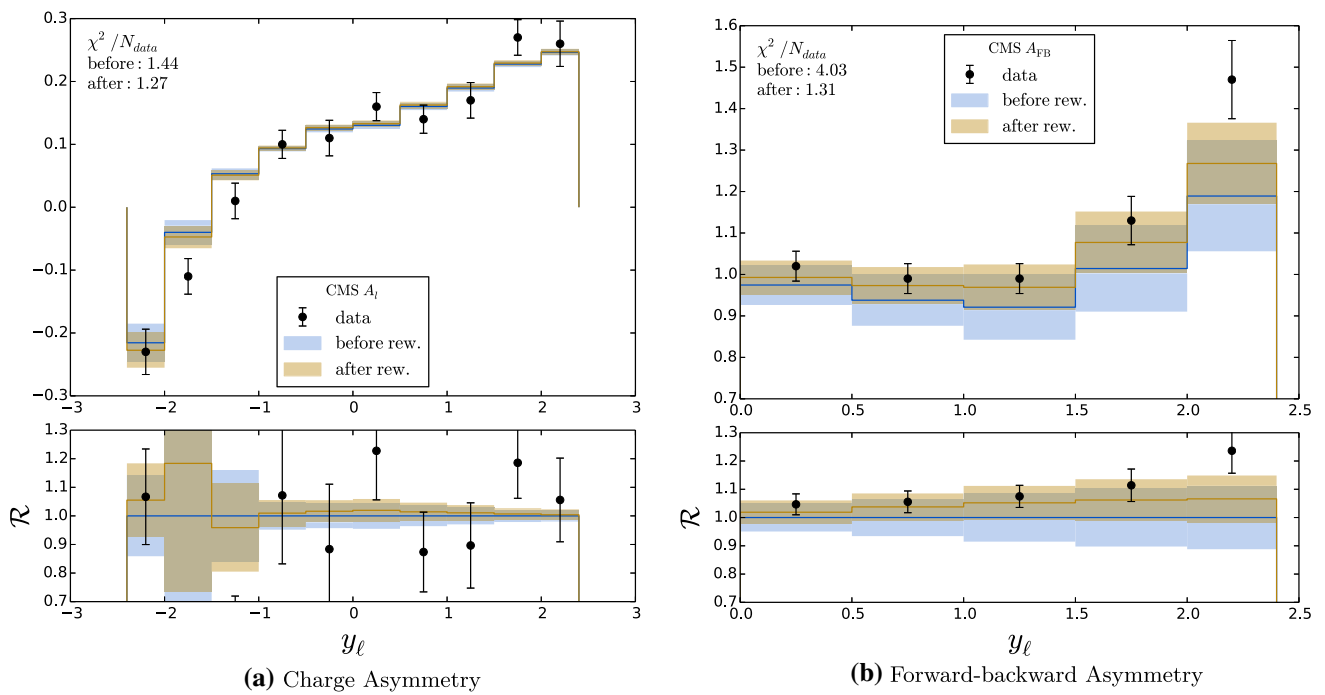


Fig. 21 Comparison of data and theory before and after the reweighting procedure using the charge asymmetry (*left*) and the forward-backward asymmetry (*right*) from W^\pm production data by CMS [4]

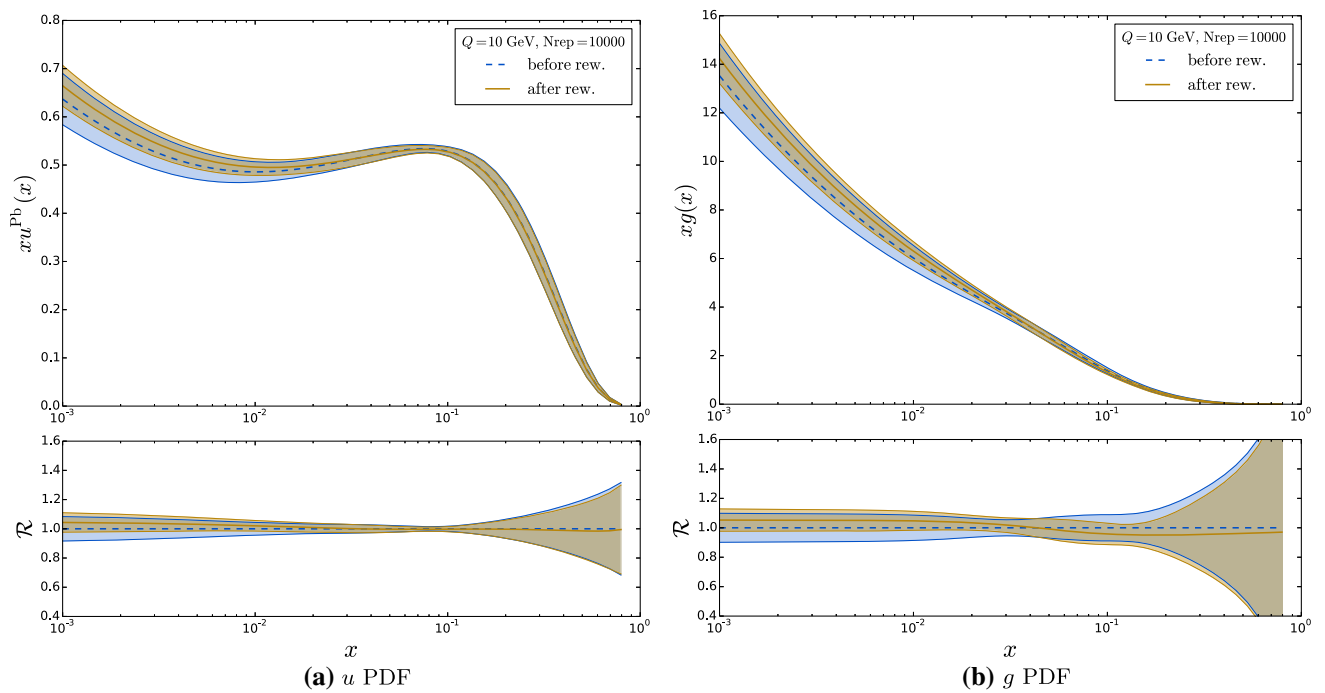


Fig. 22 Comparison of PDFs at a scale of $Q = 10$ GeV before and after the reweighting procedure using asymmetries measured by CMS [4]. The lower plots show the ratio compared to the central (average) distribution before the reweighting

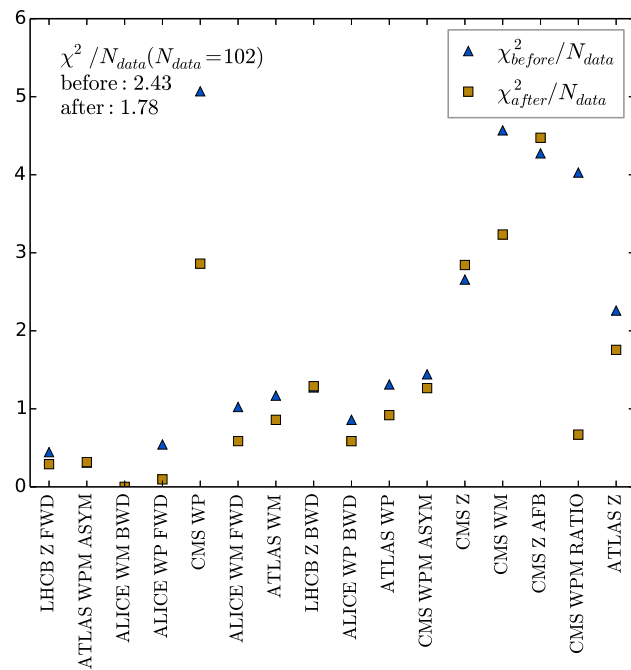


Fig. 23 χ^2 per experiment before and after the reweighting procedure using all LHC pPb data

although this can be regarded as “double counting”, it is a common practice in proton PDF analyses, e.g. [19].

As the impact of the reweighting on the theory predictions for ALICE W^\pm production data [6], LHCb Z data [5] and both ATLAS [1] and CMS [3] Z production data is very small, we will not show the corresponding comparisons of theory predictions before and after the reweighting. We do note that in the majority of these cases the χ^2/N_{data} has improved indicating that the data sets are compatible, cf. Fig. 23. However, the initial χ^2 for these data sets was already very small which reflects the large experimental uncertainties of these data sets and their limited constraining power on the nPDFs.

We start by examining the distribution of weights of the new replicas which is displayed in Fig. 24. We see that the distribution is steeply falling in a similar manner to the one from Fig. 17 obtained using only CMS W^\pm rapidity distributions, but it extends to higher values of ~ 17 . These results are not very surprising as the CMS W^\pm data set is the one introducing the most constraints. We also note that the reweighting procedure results in the effective number of replicas $N_{\text{eff}} = 3603$, which is around 40% of the number of initial replicas. This suggests that the reweighting procedure should still yield reliable results. Now we turn to the comparison of data with the theory predictions before and after the reweighting procedure. In Fig. 25 we show the predictions for the CMS W^\pm data [4], and in Fig. 26 we show the corresponding predictions for the ATLAS W^\pm data [2]. We can see that in both cases we observe an improvement

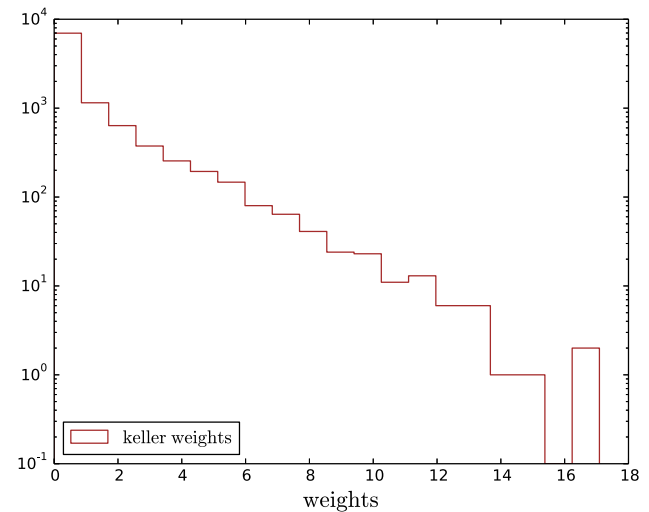


Fig. 24 Weight distribution after reweighting using all LHC pPb data on W/Z production

in the data description that is confirmed by the corresponding χ^2/N_{data} values (see figures). The χ^2 values tell us also that the largest effect comes from the CMS data which has smaller errors and for which the initial description (before the reweighting) was worse than in the ATLAS case.

Furthermore, comparing the values of χ^2/N_{data} for the CMS W^\pm data after the reweighting using all data sets and using only CMS data (Sect. 3.2) we see further improvement of χ^2/N_{data} when more data is included. This shows that the different data sets are compatible with each other and that they pull the results in the same direction.

In addition, we show in Fig. 23 the χ^2/N_{data} before and after the reweighting for each of the experiments, as well as the χ^2/N_{data} combining all 102 data points from the different experiments. This highlights the fact that the CMS W^\pm measurement yields the largest impact on the PDFs out of all the considered data sets.

Finally, in Figs. 27, 28, and 29 we present the effects of the reweighting on the $\{u, d, \bar{u}, \bar{d}, g, s\}$ distributions in lead for a scale $Q = 10 \text{ GeV}$. The effects are similar when looking at different scales. From the figures we can see that changes in the PDFs are generally affecting the low- x distributions, and to a lesser extent the moderate to high- x distributions.

When considering the ratios of PDFs, the effects of the reweighting appear to be quite substantial at large x , especially for the gluon; however, as is evident from looking at the plots of the PDFs directly, they are approaching zero at large x so the impact for physical observables is minimal.

Furthermore, when interpreting the results of the reweighting analysis it is important to remember that this method can only estimate the effects a given data set might have on the PDFs; it is not equivalent to a full fit. For example, a reweighting analysis cannot be used to explore new param-

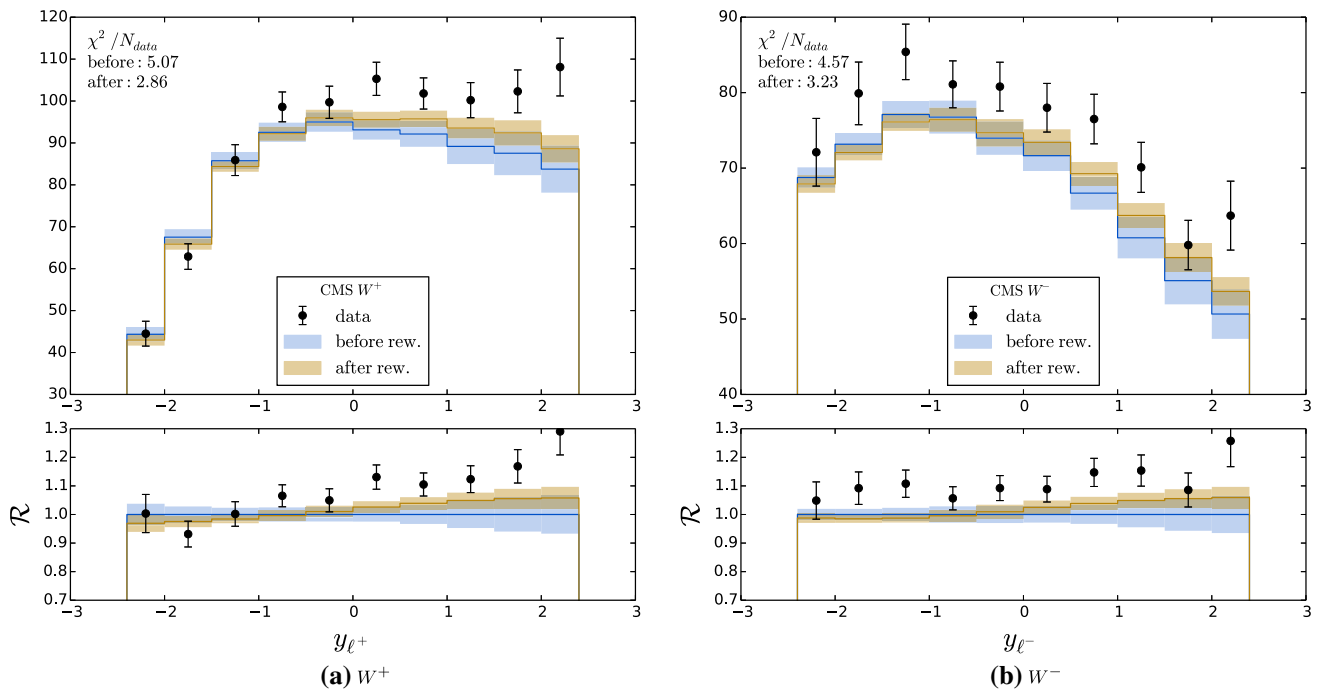


Fig. 25 Comparison of data and theory before and after the reweighting procedure using all LHC pPb data. The results for the CMS W^+ (left) and W^- (right) distributions are shown

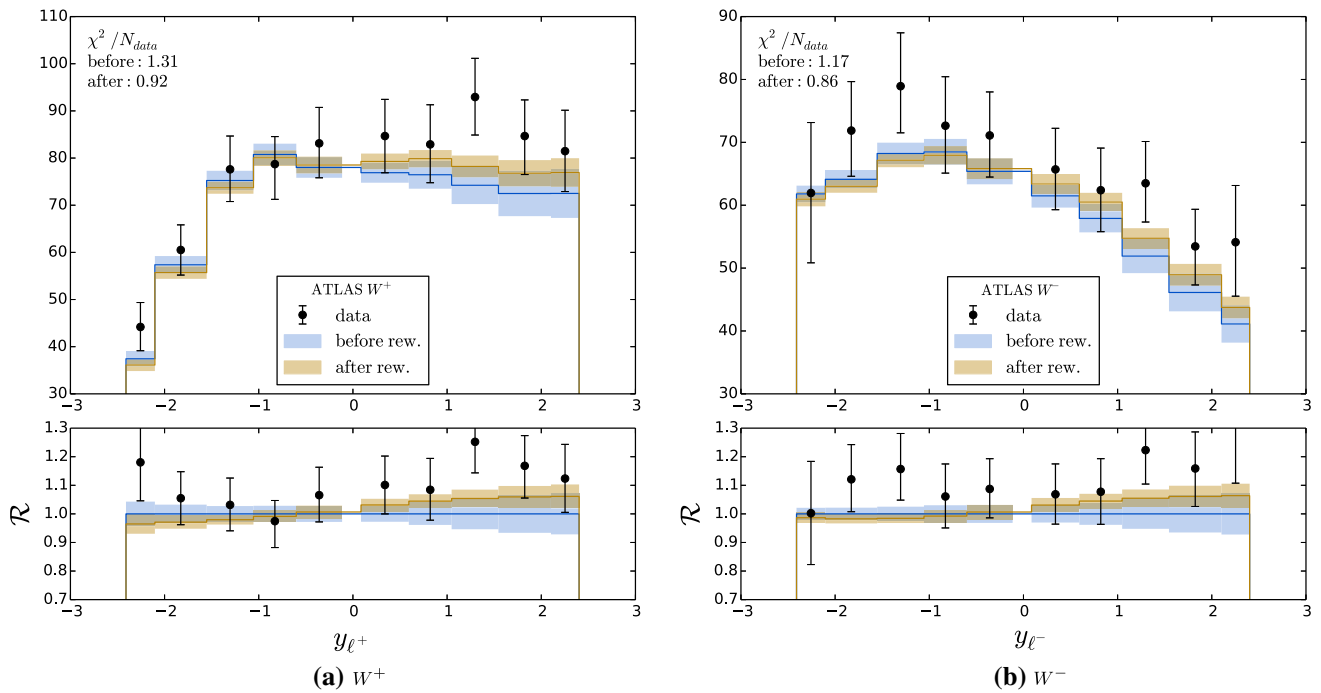


Fig. 26 Comparison of data and theory before and after the reweighting procedure using all LHC pPb data. The results for the ATLAS W^+ (left) and W^- (right) distributions are shown

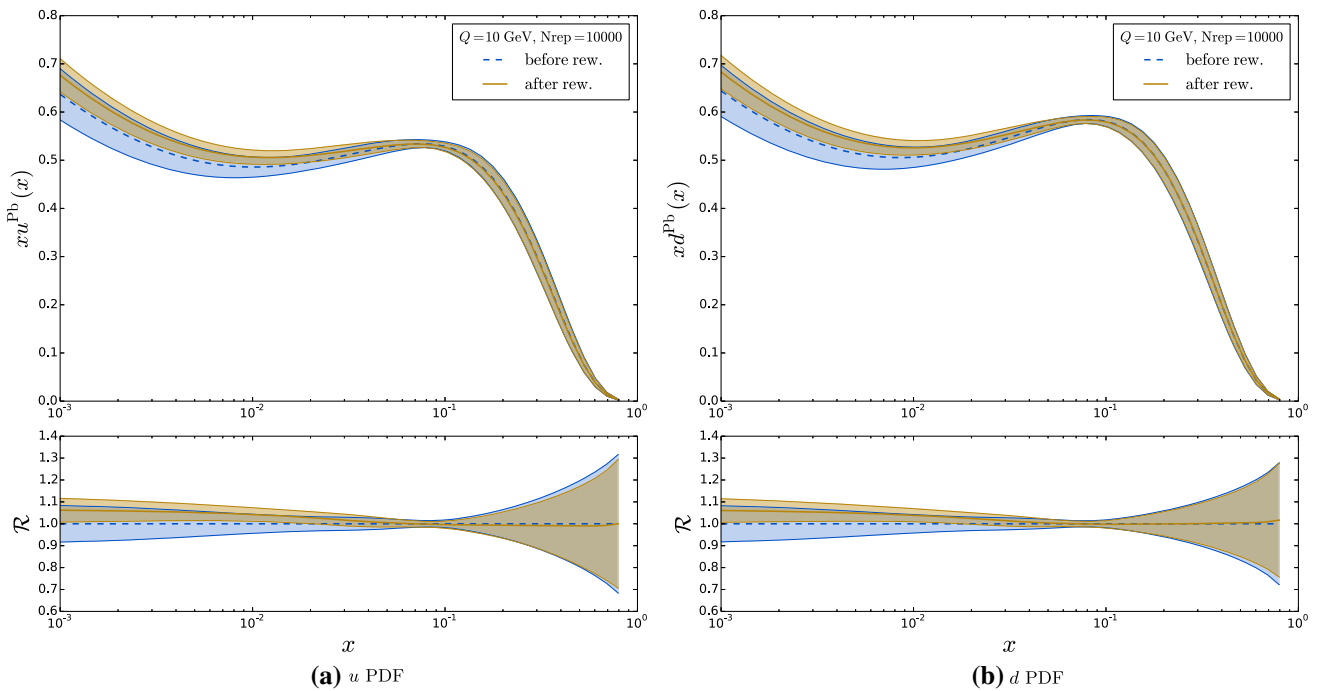


Fig. 27 u and d PDFs before and after the reweighting using all LHC pPb data sets

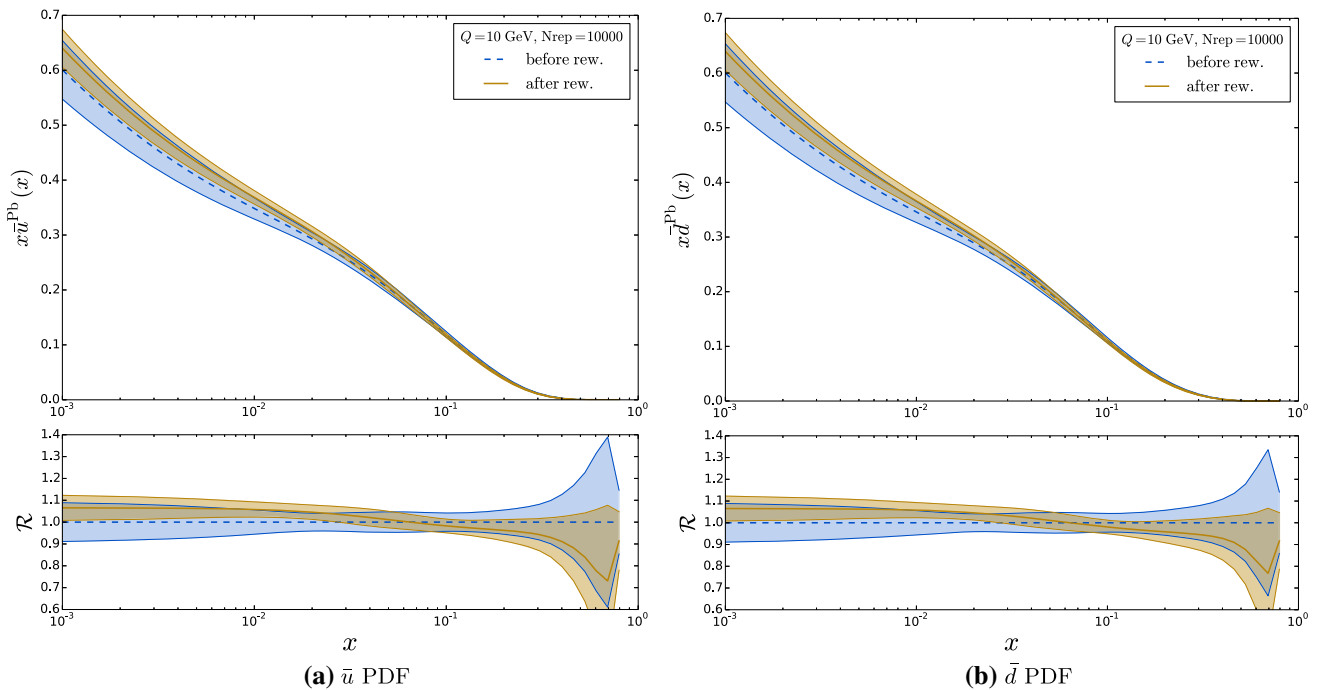


Fig. 28 \bar{u} and \bar{d} PDFs before and after the reweighting using all LHC pPb data sets

ters or other dimensions that are not already spanned by the original PDF uncertainty basis. In particular, this study has shown us that the strange quark PDF can play an important role in the LHC pPb production of W/Z . As our current $s(x)$ is parameterized proportional to $\bar{u}(x) + \bar{d}(x)$, this restricts

our ability to vary the strange PDF independently,¹¹ hence, an independent fit (in progress) is needed to better the impact of this data on the nPDFs.

¹¹ This point was explored in more detail in Ref. [41].

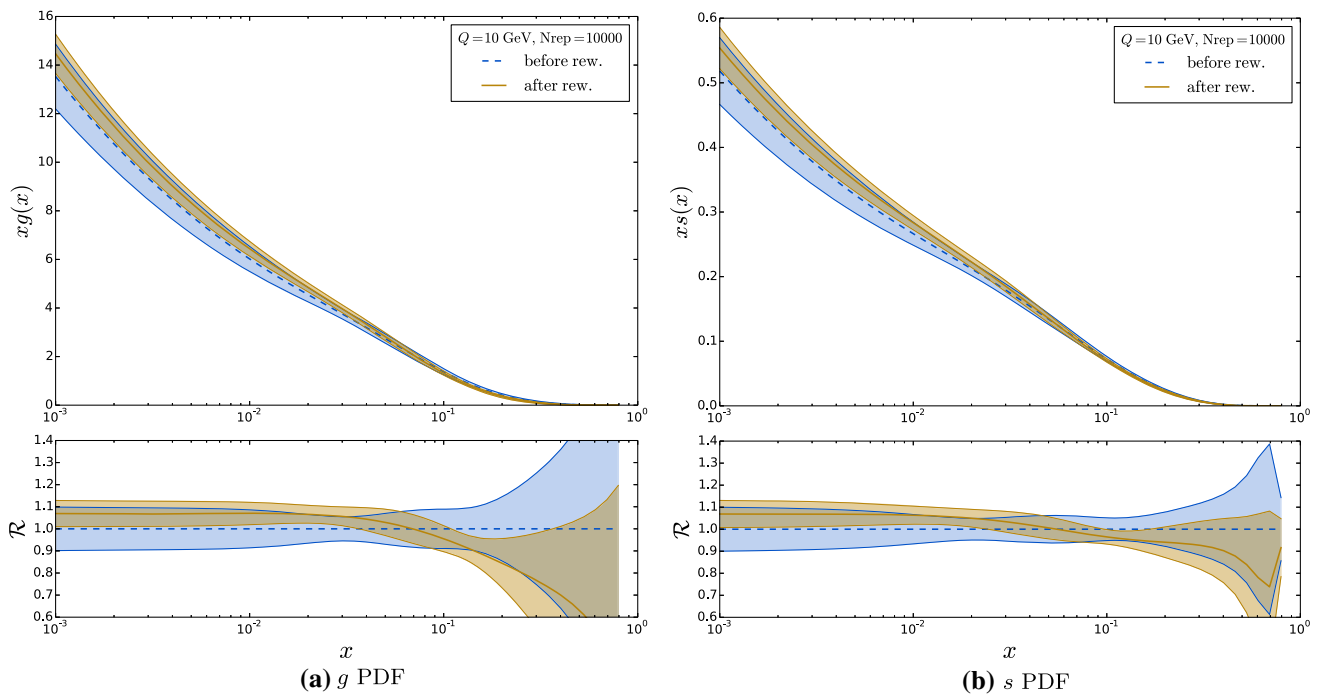


Fig. 29 g and s PDFs before and after the reweighting using all LHC pPb data sets

3.5 Comparison with EPPS16

During the course of our analysis, a new global fit including LHC data (EPPS16 [16]) has been released. This gives us an opportunity to compare the results of our reweighting study with these new PDFs. We note here that this is a qualitative comparison as the data sets used in these two studies are different. Another important difference is that the EPPS16 fit has more parameters to describe the sea-quark PDFs as compared to the nCTEQ15 analysis; this provides EPPS16 additional flexibility to accommodate all the considered data. As mentioned earlier, our reweighting of nCTEQ15 cannot compensate for our more restrictive parametrization, so this must be considered when evaluating these comparisons.

In Figs. 30 and 31 we present a comparison of $u, d, \bar{u}, \bar{d}, g$ and s for the nCTEQ15 PDFs before and after the reweighting, with the EPPS16 distributions at the scale of 80 GeV. There are a number of trends which emerge.

1. In the low x region, the reweighted nCTEQ15 PDFs approach the EPPS16 distributions; for the g and s PDFs, the central values are very close. The effect of the reweighting appears mostly in this region where (prior to the LHC data) there were minimal constraints on the PDFs. Therefore, adding the LHC data is able to significantly adjust the PDFs in this region.
2. In the intermediate x range ($\sim 3 \times 10^{-2}$), the central values of the EPPS16 and both reweighted and initial

nCTEQ15 PDFs coincide, and their uncertainty bands are also similar (except for the strange quark). This region was previously constrained by pre-LHC data, and we observe minimal changes in this region.

3. On the contrary, where x is large, the differences are more important with no consistent pattern. This is a challenging region as the absolute value of the PDFs is small, and the nCTEQ15 parameterization may not be sufficiently flexible to accommodate the new data. Additionally, the inclusion of certain data sets in the EPPS16 analysis (such as the CHORUS ν -Pb data [42]) can have a significant impact.

Finally, we also see that the EPPS16 PDFs have consistently larger uncertainty bands (especially at low x). As the nCTEQ15 uncertainty bands in this region are essentially extrapolated from larger x results, the EPPS16 uncertainties are probably a more realistic assessment. The issue of PDF parameterization is a perennial challenge for the nuclear PDFs as there is less data and more degrees of freedom as compared to the proton PDFs. The common solution is to impose assumptions on the nPDF parameters, or to limit the flexibility of the parameterization, and thereby underestimate the uncertainty. These issues highlight the importance of including this new LHC data in the nPDF analyses as they not only will help determine the central fits, but also provide for more reliable error estimation.

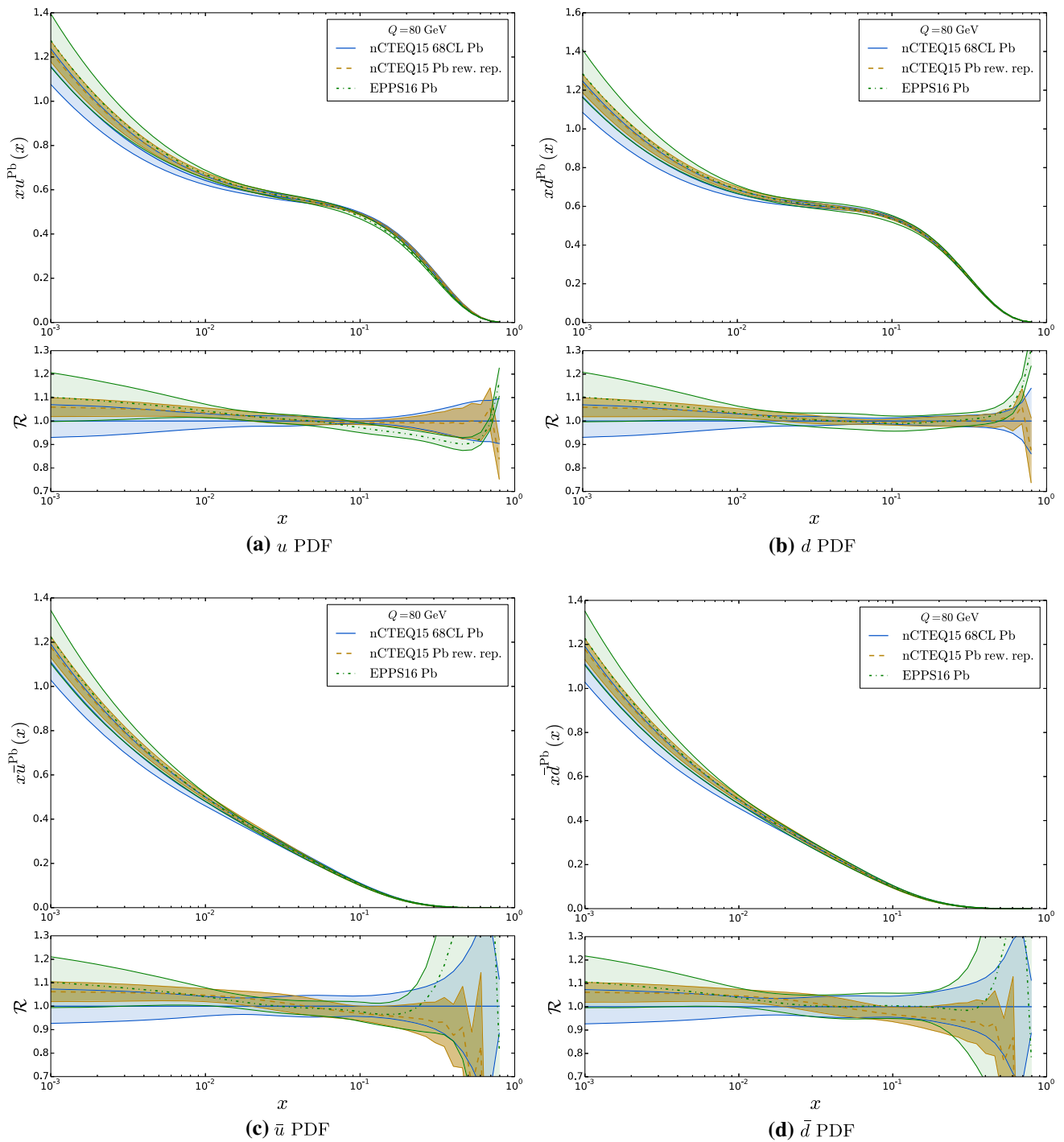


Fig. 30 Comparison of the nCTEQ15 PDFs before and after the reweighting using all pPb data sets with the EPPS16 PDFs including LHC data. The EPPS16 error bands include only the nuclear errors

(unlike what is provided in LHAPDF where also the proton baseline errors are included) and they are calculated using symmetric formula

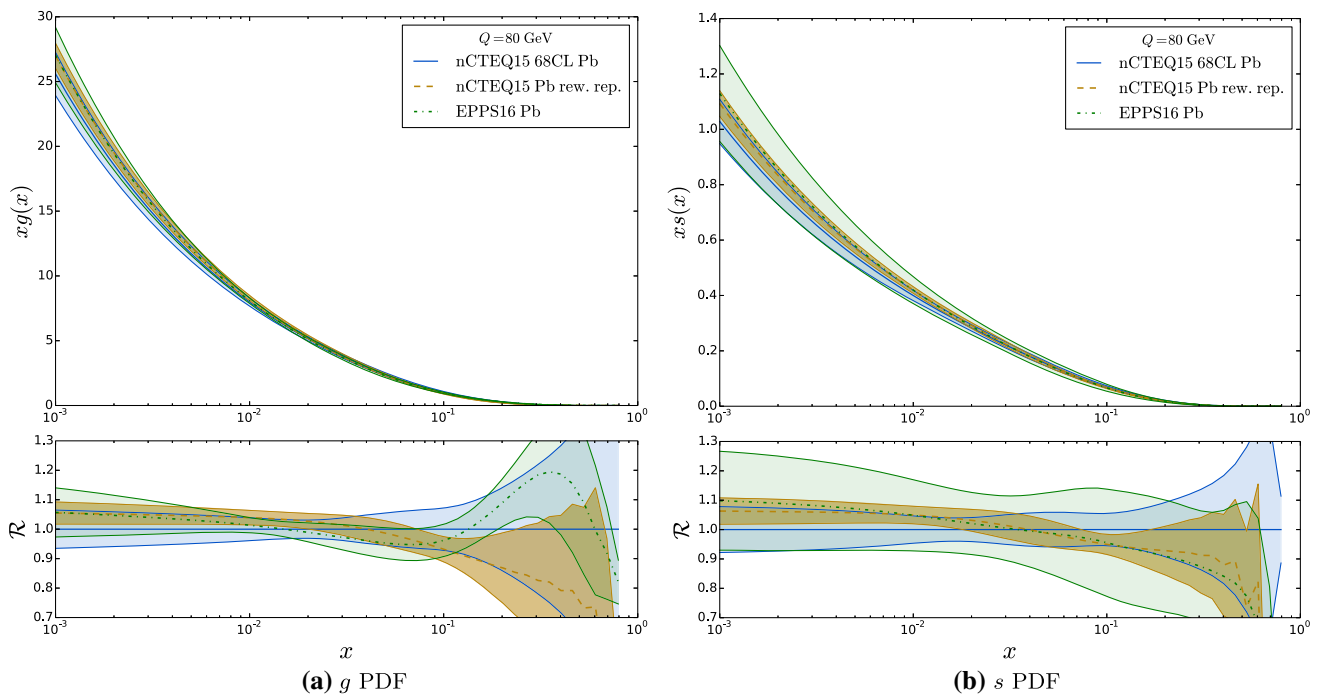


Fig. 31 Continuation of Fig. 30

4 Conclusions

We have presented a comprehensive study of vector boson production (W^\pm, Z) from lead collisions at the LHC. This LHC lead data is of particular interest for a number of reasons.

1. Comparisons with LHC proton data can determine nuclear corrections for large A values; this is a kinematic $\{x, Q^2\}$ range very different from nuclear corrections provided by fixed-target measurements.
2. The W^\pm, Z lead data are sensitive to the heavier quark flavors (especially the strange PDF), so this provides important information on the nuclear flavor decomposition.
3. Improved information on the nuclear corrections from the LHC lead data can also help reduce proton PDF uncertainties as fixed-target nuclear data is essential for distinguishing the individual flavors.

Predictions from the recent nCTEQ15 nPDFs are generally compatible with the LHC experimental data; however, this is partially due to the large uncertainties from both the nuclear corrections and the data. We do see suggestive trends (for example W^\pm production in pPb at large y_{ℓ^+}) which may impose influential constraints on the nPDF fits as the experimental uncertainties are reduced. Intriguingly, the large rapidity W/Z data seem to prefer nuclear PDFs with no or delayed shadowing at small x , similar to what has been observed in ν -Fe DIS. This observation was validated

by our reweighting study that demonstrated the impact of the W/Z pPb data on nPDFs.

The uncertainties of the currently available data are relatively large, and correlated errors are not yet available. Fortunately, we can look forward to more data (with improved statistics) in the near future as additional heavy ion runs are scheduled.

While the above reweighting technique provides a powerful method to quickly assess the impact of new data, there are limitations. For example, the reweighting method cannot introduce or explore new degrees of freedom. Thus, if the original fit imposes artificial constraints (such as linking the strange PDF to the up and down sea distributions), this limitation persists for the reweighted PDF [41].

Most importantly, our correlation study (Sect. 2.4) demonstrated the importance of the strange distribution for the vector boson (W/Z) production at the LHC, possibly even pointing to a nuclear strangeness asymmetry ($s(x) > \bar{s}(x)$). The comparison of the two flavor and five flavor results illustrates how flavor decomposition and nuclear corrections can become entangled. Therefore, it is imperative to separately control the strange PDF and the nuclear correction factor if we are to obtain unambiguous results. The investigations performed in this paper provide a foundation for improving our determination of the PDFs in lead, especially the strange quark component. Combining this information in a new nCTEQ fit across the full A range can produce improved nPDFs, and thus yield improved nuclear correction factors. These improved nuclear correction factors, together with the

LHC W/Z production data for pp , can refine our knowledge of the strange PDF in the proton.

Acknowledgements The authors would like to thank C. Bertulani, M. Botje, C. Keppel, S. Kumano, J.G. Morfín, P. Nadolsky, J.F. Owens, F. Petriello, R. Plačákytė, and V. Radescu for valuable discussions. We acknowledge the hospitality of CERN, DESY, and Fermilab where a portion of this work was performed. This work was also partially supported by the U.S. Department of Energy under Grant No. DE-SC0010129 and by the National Science Foundation under Grant No. NSF PHY11-25915.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. Funded by SCOAP³.

References

- G. Aad et al., Phys. Rev. C **92**(4), 044915 (2015). doi:[10.1103/PhysRevC.92.044915](https://doi.org/10.1103/PhysRevC.92.044915)
- ATLAS Collaboration, Measurement of $W \rightarrow \mu\nu$ production in p +Pb collision at $\sqrt{s_{NN}} = 5.02$ TeV with ATLAS detector at the LHC. ATLAS-CONF-2015-056 (2015). <http://cds.cern.ch/record/2055677>
- V. Khachatryan et al., Phys. Lett. B **759**, 36 (2016). doi:[10.1016/j.physletb.2016.05.044](https://doi.org/10.1016/j.physletb.2016.05.044)
- V. Khachatryan et al., Phys. Lett. B **750**, 565 (2015). doi:[10.1016/j.physletb.2015.09.057](https://doi.org/10.1016/j.physletb.2015.09.057)
- R. Aaij et al., JHEP **09**, 030 (2014). doi:[10.1007/JHEP09\(2014\)030](https://doi.org/10.1007/JHEP09(2014)030)
- K.J. Senosi, PoS Bormio **2015**, 042 (2015)
- G. Aad et al., Phys. Rev. Lett. **110**(2), 022301 (2013). doi:[10.1103/PhysRevLett.110.022301](https://doi.org/10.1103/PhysRevLett.110.022301)
- G. Aad et al., Eur. Phys. J. C **75**(1), 23 (2015). doi:[10.1140/epjc/s10052-014-3231-6](https://doi.org/10.1140/epjc/s10052-014-3231-6)
- S. Chatrchyan et al., JHEP **03**, 022 (2015). doi:[10.1007/JHEP03\(2015\)022](https://doi.org/10.1007/JHEP03(2015)022)
- S. Chatrchyan et al., Phys. Lett. B **715**, 66 (2012). doi:[10.1016/j.physletb.2012.07.025](https://doi.org/10.1016/j.physletb.2012.07.025)
- H. Paukkunen, C.A. Salgado, JHEP **03**, 071 (2011). doi:[10.1007/JHEP03\(2011\)071](https://doi.org/10.1007/JHEP03(2011)071)
- N. Armesto, H. Paukkunen, J.M. Penín, C.A. Salgado, P. Zurita, Eur. Phys. J. C **76**(4), 218 (2016). doi:[10.1140/epjc/s10052-016-4078-9](https://doi.org/10.1140/epjc/s10052-016-4078-9)
- P. Ru, S.A. Kulagin, R. Petti, B.W. Zhang, Phys. Rev. D **94**(11), 113013 (2016). doi:[10.1103/PhysRevD.94.113013](https://doi.org/10.1103/PhysRevD.94.113013)
- K.J. Eskola, H. Paukkunen, C.A. Salgado, JHEP **04**, 065 (2009). doi:[10.1088/1126-6708/2009/04/065](https://doi.org/10.1088/1126-6708/2009/04/065)
- D. de Florian, R. Sassot, P. Zurita, M. Stratmann, Phys. Rev. D **85**, 074028 (2012). doi:[10.1103/PhysRevD.85.074028](https://doi.org/10.1103/PhysRevD.85.074028)
- K.J. Eskola, P. Paakkinen, H. Paukkunen, C.A. Salgado, Eur. Phys. J. C **77**(3), 163 (2017). doi:[10.1140/epjc/s10052-017-4725-9](https://doi.org/10.1140/epjc/s10052-017-4725-9)
- R.D. Ball et al., JHEP **04**, 040 (2015). doi:[10.1007/JHEP04\(2015\)040](https://doi.org/10.1007/JHEP04(2015)040)
- L.A. Harland-Lang, A.D. Martin, P. Motylinski, R.S. Thorne, Eur. Phys. J. C **75**(5), 204 (2015). doi:[10.1140/epjc/s10052-015-3397-6](https://doi.org/10.1140/epjc/s10052-015-3397-6)
- S. Dulat, T.J. Hou, J. Gao, M. Guzzi, J. Huston, P. Nadolsky, J. Pumplin, C. Schmidt, D. Stump, C.P. Yuan, Phys. Rev. D **93**(3), 033006 (2016). doi:[10.1103/PhysRevD.93.033006](https://doi.org/10.1103/PhysRevD.93.033006)
- H. Khanpour, S. Atashbar Tehrani, Phys. Rev. D **93**(1), 014026 (2016). doi:[10.1103/PhysRevD.93.014026](https://doi.org/10.1103/PhysRevD.93.014026)
- K. Kovarik et al., Phys. Rev. D **93**(8), 085037 (2016). doi:[10.1103/PhysRevD.93.085037](https://doi.org/10.1103/PhysRevD.93.085037)
- R. Gavin, Y. Li, F. Petriello, S. Quackenbush, Comput. Phys. Commun. **182**, 2388 (2011). doi:[10.1016/j.cpc.2011.06.008](https://doi.org/10.1016/j.cpc.2011.06.008)
- R. Gavin, Y. Li, F. Petriello, S. Quackenbush, Comput. Phys. Commun. **184**, 208 (2013). doi:[10.1016/j.cpc.2012.09.005](https://doi.org/10.1016/j.cpc.2012.09.005)
- H.L. Lai, M. Guzzi, J. Huston, Z. Li, P.M. Nadolsky, J. Pumplin, C.P. Yuan, Phys. Rev. D **82**, 074024 (2010). doi:[10.1103/PhysRevD.82.074024](https://doi.org/10.1103/PhysRevD.82.074024)
- A.D. Martin, W.J. Stirling, R.S. Thorne, G. Watt, Eur. Phys. J. C **63**, 189 (2009). doi:[10.1140/epjc/s10052-009-1072-5](https://doi.org/10.1140/epjc/s10052-009-1072-5)
- A. Kusina, T. Stavreva, S. Berge, F.I. Olness, I. Schienbein, K. Kovarik, T. Jezo, J.Y. Yu, K. Park, Phys. Rev. D **85**, 094028 (2012). doi:[10.1103/PhysRevD.85.094028](https://doi.org/10.1103/PhysRevD.85.094028)
- K. Kovarik, I. Schienbein, F.I. Olness, J.Y. Yu, C. Keppel, J.G. Morfin, J.F. Owens, T. Stavreva, Phys. Rev. Lett. **106**, 122301 (2011). doi:[10.1103/PhysRevLett.106.122301](https://doi.org/10.1103/PhysRevLett.106.122301)
- I. Schienbein, J.Y. Yu, K. Kovarik, C. Keppel, J.G. Morfin, F. Olness, J.F. Owens, Phys. Rev. D **80**, 094004 (2009). doi:[10.1103/PhysRevD.80.094004](https://doi.org/10.1103/PhysRevD.80.094004)
- S.X. Nakamura et al., Rep. Prog. Phys. **80**(5), 056301 (2017). doi:[10.1088/1361-6633/aa5e6c](https://doi.org/10.1088/1361-6633/aa5e6c)
- P.M. Nadolsky, H.L. Lai, Q.H. Cao, J. Huston, J. Pumplin, D. Stump, W.K. Tung, C.P. Yuan, Phys. Rev. D **78**, 013004 (2008). doi:[10.1103/PhysRevD.78.013004](https://doi.org/10.1103/PhysRevD.78.013004)
- D. Mason et al., Phys. Rev. Lett. **99**, 192001 (2007). doi:[10.1103/PhysRevLett.99.192001](https://doi.org/10.1103/PhysRevLett.99.192001)
- J. Qiu, Nucl. Phys. A **715**, 309–318 (2003). doi:[10.1016/S0375-9474\(02\)01441-0](https://doi.org/10.1016/S0375-9474(02)01441-0)
- W.T. Giele, S. Keller, Phys. Rev. D **58**, 094023 (1998). doi:[10.1103/PhysRevD.58.094023](https://doi.org/10.1103/PhysRevD.58.094023)
- R.D. Ball, V. Bertone, F. Cerutti, L. Del Debbio, S. Forte, A. Guffanti, J.I. Latorre, J. Rojo, M. Ubiali, Nucl. Phys. B **849**, 112 (2011). doi:[10.1016/j.nuclphysb.2011.03.017](https://doi.org/10.1016/j.nuclphysb.2011.03.017). doi:[10.1016/j.nuclphysb.2011.09.011](https://doi.org/10.1016/j.nuclphysb.2011.09.011). [Erratum: Nucl. Phys. B **855**, 927 (2012)]
- R.D. Ball, V. Bertone, F. Cerutti, L. Del Debbio, S. Forte, A. Guffanti, N.P. Hartland, J.I. Latorre, J. Rojo, M. Ubiali, Nucl. Phys. B **855**, 608 (2012). doi:[10.1016/j.nuclphysb.2011.10.018](https://doi.org/10.1016/j.nuclphysb.2011.10.018)
- N. Sato, J.F. Owens, H. Prosper, Phys. Rev. D **89**(11), 114020 (2014). doi:[10.1103/PhysRevD.89.114020](https://doi.org/10.1103/PhysRevD.89.114020)
- H. Paukkunen, P. Zurita, JHEP **12**, 100 (2014). doi:[10.1007/JHEP12\(2014\)100](https://doi.org/10.1007/JHEP12(2014)100)
- G. Watt, R.S. Thorne, JHEP **08**, 052 (2012). doi:[10.1007/JHEP08\(2012\)052](https://doi.org/10.1007/JHEP08(2012)052)
- T.J. Hou et al., JHEP **03**, 099 (2017). doi:[10.1007/JHEP03\(2017\)099](https://doi.org/10.1007/JHEP03(2017)099)
- F. Arleo, E. Chapon, H. Paukkunen, Eur. Phys. J. C **76**(4), 214 (2016). doi:[10.1140/epjc/s10052-016-4049-1](https://doi.org/10.1140/epjc/s10052-016-4049-1)
- A. Kusina, F. Lyonnet, D.B. Clark, E. Godat, T. Jezo, K. Kovarik, F.I. Olness, I. Schienbein, J.Y. Yu, Acta Phys. Polon. B **48**, 1035 (2017). doi:[10.5506/APhysPolB.48.1035](https://doi.org/10.5506/APhysPolB.48.1035)
- G. Onengut et al., Phys. Lett. B **632**, 65 (2006). doi:[10.1016/j.physletb.2005.10.062](https://doi.org/10.1016/j.physletb.2005.10.062)