The strangest proton?

Ferran Faura¹, Shayan Iranipour², Emanuele R. Nocera¹, Juan Rojo¹,³,a, Maria Ubiali²

¹ Nikhef Theory Group, Science Park 105, 1098 XG Amsterdam, The Netherlands
² DAMTP, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, UK
³ Department of Physics and Astronomy, VU, 1081 HV Amsterdam, The Netherlands

Received: 25 September 2020 / Accepted: 10 December 2020 / Published online: 19 December 2020
© The Author(s) 2020

Abstract We present an improved determination of the strange quark and antiquark parton distribution functions of the proton by means of a global QCD analysis that takes into account a comprehensive set of strangeness-sensitive measurements: charm-tagged cross sections for fixed-target neutrino–nucleus deep-inelastic scattering, and cross sections for inclusive gauge-boson production and W-boson production in association with light jets or charm quarks at hadron colliders. Our analysis is accurate to next-to-next-to-leading order in perturbative QCD where available, and specifically includes charm-quark mass corrections to neutrino–nucleus structure functions. We find that a good overall description of the input dataset can be achieved and that a strangeness moderately suppressed in comparison to the rest of the light sea quarks is strongly favored by the global analysis.

Contents

1 Introduction .......................................................... 1
2 Analysis settings .................................................... 2
  2.1 Experimental data .................................... 2
  2.2 Theoretical calculations .............................. 3
    2.2.1 The NOMAD ratio ............................ 3
    2.2.2 Theoretical improvements ................. 3
  2.3 PDF sets .................................................. 5
3 Results ............................................................ 6
  3.1 Fit quality ............................................... 6
  3.2 Comparison with experimental data ............ 7
  3.3 Parton distributions .................................. 8
    3.3.1 Comparison with NNPDF3.1 .............. 8
    3.3.2 Total and valence strange distributions . 9
    3.3.3 Light-quark, charm, and gluon PDFs .... 11
  3.4 The strange content of the proton revisited .... 12
4 Summary .......................................................... 15

a e-mail: j.rojo@vu.nl (corresponding author)

1 Introduction

An accurate determination of the strange quark and antiquark parton distribution functions (PDFs) of the proton [1–3] is key to carrying out precision phenomenology at current and future colliders, specifically for measuring fundamental parameters of the standard model (SM) such as the mass of the W boson [4], the Weinberg angle [5], and electroweak parameters in general [6]. Because of the limited experimental information available, however, the strange quark and antiquark PDFs remain much more uncertain than the up and down sea quark PDFs.

The strange quark and antiquark PDFs have been determined from neutrino–nucleus deep-inelastic scattering (DIS) for a long time, specifically from measurements of dimuon cross sections, whereby the secondary muon originates from the decay of a charmed meson, νμ + N → μ + c + X with c → D → μ + X [7–10]. When interpreted in terms of the ratio between strange and non-strange sea quark PDFs, Rs ≡ (s + ¯s)/(¯u + ¯d), these measurements favor values around Rs ∼< 0.5 when PDFs are evaluated at values of the momentum fraction x = 0.023 and scale Q = 1.6 GeV. Therefore, it came as a surprise when a QCD analysis of the W- and Z-boson rapidity distributions measured by the ATLAS experiment in proton–proton collisions [11], later corroborated by an analysis based on an increased integrated luminosity [12], suggested instead a ratio closer to Rs ≃ 1. Complementary information on the strange quark and antiquark PDFs is provided by W-boson production in association with light jets [13] and charm quarks [14], the latter process being dominated by the partonic scattering g + s → W + c. Measurements of these processes were performed by the ATLAS [15,16] and CMS [17,18] experiments recently. Although ATLAS and CMS W+c measure-
ments turned out to be consistent at the parton level, different interpretations in terms of $R_s$ were claimed [16,17].

This state of affairs has motivated studies of the proton strangeness within the CT, MMHT, and NNPDF global fits, with overall consistent findings. The NNPDF3.1 analysis [19] found that, whereas the ATLAS $W$, $Z$ dataset [12] does indeed favour a larger total strangeness, its $\chi^2$ remains non-optimal when fitted together with the neutrino dimuon data. The recent CT18 global analysis [20] also presented fits with and without the ATLAS measurement of [12], with the resulting PDFs differing by more than one-sigma both for the gluon and for the total strangeness. An update of the global PDF analysis from the MMHT collaboration [21], which for the first time accounted for the NNLO massive corrections to the neutrino dimuon cross sections within a PDF fit, also revealed an enhanced strangeness driven by the ATLAS $W$, $Z$ dataset. The resulting PDFs were, however, consistent within uncertainties with the corresponding fit once this dataset was excluded. Additional dedicated studies of the strange quark and antiquark PDFs have been presented [22–26], however, these focused on a restricted set of processes or datasets, or were based on theoretical and methodological assumptions that can potentially bias the results.

Given its phenomenological relevance for precision physics at the LHC, a global reinterpretation of all of the strangeness-sensitive measurements within an accurate theoretical and methodological framework appears to be therefore timely and compelling. This paper fulfills this purpose: we present the results of these fits, we assess their quality, and we use them to understand how the datasets and the theoretical framework affect the PDFs, in particular in relationship with the strangeness content of the proton. We finally provide a summary of our work in Sect. 4.

2 Analysis settings

In this section we present the experimental datasets used as input to our analysis, we then discuss the details of the corresponding theoretical computations, and we finally explain which PDF fits we perform to study their impact on the proton strangeness.

2.1 Experimental data

The bulk of the dataset included in our analysis corresponds to the one used in [27], which is in turn a variant of the dataset used in the NNPDF3.1 NNLO analysis [19]. It contains in particular measurements of the dimuon neutrino–nucleus DIS cross sections from the NuTeV experiment [9], and of inclusive gauge-boson production in proton–(anti)proton collisions from several Tevatron and LHC experiments [12,28–31]. These measurements represented the most constraining source of experimental information on the strange quark and antiquark PDFs in the NNPDF3.1 analysis.

We supplement this dataset with a number of new measurements. Concerning neutrino–nucleus DIS, we include measurements of the ratio of dimuon to inclusive charged-current cross sections, $R_{\nu\mu}(\omega) = \sigma_{\nu\mu}(\omega)/\sigma_{CC}(\omega)$, from the NOMAD experiment [10], see Sect. 2.2 for details. The data is presented for three kinematic variables $\omega$: the neutrino beam energy $E_\nu$, the momentum fraction $x$, and the square root of the final-state invariant mass $\sqrt{s}$. Given that experimental correlations are not provided among measurements in different kinematic variables, only one measurement can be included in the fit at a time: we select the $n_{dat} = 19$ data points as a function of $E_\nu$, the only variable which is directly measured by the experiment among the three. We will nevertheless verify that similar results can be obtained for instance with the $\sqrt{s}$-dependent dataset. The kinematic sensitivity of the NOMAD measurements is roughly $0.03 < x < 0.7$, as illustrated by the coverage of the $x$-dependent dataset.

Concerning proton–proton collisions, we augment the inclusive gauge-boson production measurement from the ATLAS experiment at a center-of-mass energy (c.m.e.) of 7 TeV [12] with the off-peak and forward rapidity bins (not included in NNPDF3.1) for a total of $n_{dat} = 61$ data points. Furthermore, we include the $n_{dat} = 37$ data points corresponding to the ATLAS (at a c.m.e. of 7 TeV) [16] and CMS (at a c.m.e. of 7 TeV and 13 TeV) [17,18] $W+c$ measurements; for ATLAS, we consider the charm-jet dataset, which is amenable to fixed-order calculations (instead of the $D$-meson dataset). Finally we take into account the $n_{dat} = 32$ data points corresponding to the ATLAS $W+$jets...
measurement (at a c.m.e. of 8 TeV) differential in the transverse momentum of the $W$ boson [15]. Overall, these LHC datasets are sensitive to the proton strangeness in the region $10^{-3} \lesssim x \lesssim 0.1$. The present analysis contains a total of $n_{\text{dat}} = 4096$ data points; experimental correlations within each dataset are available for all of the new measurements considered here and are therefore included in our analysis.

2.2 Theoretical calculations

The measurements outlined in the previous section correspond to hadronic observables already considered in [19], except for the ratio $R_{\mu\mu}$ measured by the NOMAD experiment, and for the production of $W$ bosons in association with light jets measured by the ATLAS experiment. Likewise, the theoretical settings adopted in the present analysis closely follow those described in the NNLO analysis of [19,27] (whereby, in particular, the charm PDF is fitted), except for some improvements. In this section we discuss in turn the new NOMAD observable and the theoretical details unique to the present analysis.

2.2.1 The NOMAD ratio

As mentioned above, the NOMAD experiment measured the ratio of dimuon to inclusive charged-current cross sections, $R_{\mu\mu}$. Both the numerator and the denominator of $R_{\mu\mu}$ are evaluated as two-dimensional integrals of the differential cross sections over the fiducial phase space. For the $E_v$-dependent dataset, which we include by default, we have

$$
\sigma_i(E_v) = \int_{x_0}^1 \frac{dx}{x} \int_{Q^2_{\min}}^{Q^2_{\max}(x)} dQ^2 \frac{d^2\sigma_i}{dx dQ^2}(x, Q^2, E_v), \quad (2.1)
$$

where $Q^2_{\max}(x) = 2m_p E_v x$ and $x_0 = Q^2_{\min}/(2m_p E_v)$, with $m_p$ the proton mass. While the NOMAD measurements are reconstituted for $Q^2 \geq 1$ GeV$^2$, we assume $Q^2_{\min} = Q_0^2$, where $Q_0 = 1.65$ GeV is the initial parametrization scale adopted in our analysis [19]. We explicitly verified that results are unaffected if $Q^2_{\min} = 1$ GeV$^2$ is chosen instead. The integrand in Eq. (2.1) is either the dimuon ($i = \mu\mu$, entering the numerator or $R_{\mu\mu}$) or the inclusive ($i = CC$, entering the denominator of $R_{\mu\mu}$) charged-current cross section,

$$
\frac{d^2\sigma_i}{dx dQ^2}(x, Q^2, E_v) = \frac{G_F M_W^4}{4\pi} \frac{1}{(Q^2 + M_W^2)^2} \left[ Y_+ - \frac{2m_p^2 y^2}{Q^2} \right] F_2^i(x, Q^2) - y^2 F_L^i(x, Q^2) + Y_- x F_3^i,
$$

The kinematic factors $Y_\pm = 1 \pm (1 - y)^2$ are related to the inelasticity $y = Q^2/(2m_p E_v x)$; $G_F$ and $M_W$ are respectively the Fermi constant and the mass of the $W$ boson. The factor $K^i$ is either the identity, for $i = CC$, or the charm semileptonic branching ratio $B_\mu$, for $i = \mu\mu$. In the latter case we use the $E_v$-dependent parametrization

$$
B_\mu(E_v) = a(1 + b/E_v)^{-1},
$$

with the values of the parameters $a$ and $b$ determined in [10], $a = 0.097 \pm 0.003$ and $b = 6.7 \pm 1.8$. The corresponding uncertainty is included in the experimental covariance matrix of the measurement.

Both the charm (for $i = \mu\mu$) and the total (for $i = CC$) structure functions $F_j^i (\rho = 2, L, 3)$ entering Eq. (2.2) are evaluated with APFEL [32]. We benchmarked our results against those obtained from an independent computation based on [33]. After the correction of a bug in APFEL, which affected the computation of the large-$x$ DIS coefficient functions at next-to-leading (NLO) order, the relative difference between the two is found to be of the order of permille, apart from the lowest $E_v$ bins, in which it reaches the percent level, as displayed in the left panel of Fig. 1.

2.2.2 Theoretical improvements

In comparison to the earlier NNPDF analyses [19,27], here we introduce several theoretical improvements, which are summarized in turn below.

**NNLO massive corrections in neutrino DIS** We incorporate the recently computed NNLO charm-quark massive corrections [33,34] in the description of the NuTeV and NOMAD measurements. We do so by multiplying the NLO theoretical prediction in the FONLL general-mass variable flavor number scheme [35,36] by a $K$-factor defined as the ratio between the NNLO result in the fixed-flavor number (FFN) scheme with and without the charm-mass correction in the matrix elements (ME); NNLO PDFs are used in both cases. The resulting $K$-factor,

$$
K_{\text{NNLO}} \equiv \frac{\sigma_{\text{FFN}}(\text{NNLO PDFs, NNLO ME})}{\sigma_{\text{FFN}}(\text{NNLO PDFs, NLO ME})},
$$

is such that the prediction for the NuTeV dimuon cross sections becomes

$$
\frac{d^2\sigma_{\mu\mu}}{dx dQ^2} \bigg|_{\text{FONLL (NNLO ME)}} = K_{\text{NNLO}} \times \frac{d^2\sigma_{\mu\mu}}{dx dQ^2} \bigg|_{\text{FONLL (NLO ME)}},
$$

and an analogous expression holds for the NOMAD observables.

This approach provides a good approximation of the exact result, because theoretical predictions in the FFN and FONLL schemes are very close for the NuTeV and NOMAD kinematics. This fact was demonstrated in [36] in the case of
NuTeV data; we nevertheless checked that it remains true with the independent computation of [33,34], and that it also applies to the NOMAD measurements. To this purpose, we computed the relative difference between the FONLL-A and FFN scheme predictions for the NuTeV and NOMAD datasets based on structure functions accurate to $\mathcal{O}(\alpha_s)$. We found that differences were less than 1% in the entire kinematic range for NuTeV, and of about 1.5% irrespective of the value of $E_\nu$ for NOMAD. These differences are well below the experimental and the PDF uncertainties.\footnote{A $\simeq 1.5\%$ difference in $K_{\text{NNLO}}$ corresponds to at most a $\simeq 0.1\%$ effect in the absolute cross sections.} We therefore conclude that using a NNLO $K$-factor determined in the FFN scheme in fits that otherwise use the FONLL scheme throughout is unlikely to affect the final results. Given the current implementation of the FONLL scheme in the DIS observables [35] the matching between the NNLO massless and massive calculations would require non-trivial modifications of the code of [33], e.g. to extract the collinear logarithms, with little practical advantage.

The $K$-factors are in general smaller than unity, and thus enhance the (anti-)strange quark PDF when accounted for in the fit. This fact is consistent with what was already observed in [21], and is further illustrated in the right panel of Fig. 1, where we display the charm production cross section, Eq. (2.1), with $i = \mu\mu$, as a function of $E_\nu$ in the kinematic range measured by the NOMAD experiment. The cross section is obtained in the FFN scheme (with $n_f = 3$) at different perturbative orders using the NNPDF3.1 NNLO PDF set (consistently with $n_f = 3$). The inset displays the ratio to the leading order (LO) calculation. Higher-order corrections clearly suppress the cross section, in particular as $E_\nu$ increases. For instance, in the highest energy bin the NNLO cross section is about 10% smaller than the LO one. The size of the NNLO correction is comparable to or larger than the size of the NLO one, therefore its inclusion is mandatory to achieve a good description of the data. While the comparison of Fig. 1 is presented in the FFN scheme, all the fits discussed below are based on the FONLL scheme.

\section*{NNLO corrections in collider gauge-boson production}

Theoretical predictions for inclusive $W$- and $Z$-boson production and for $W$-boson production in association with charm quarks or light jets are evaluated at NLO using MCFM+APPLGRID [37,38], and are supplemented with NNLO QCD $K$-factors. These are evaluated with FEWZ [39] for inclusive gauge-boson production, and with the $N_{\text{jetty}}$ program [40,41] for $W$-boson production with light jets. In the first case, a fixed factorization and renormalization scale is used, equal to the mass of the gauge boson; in the second case, a dynamic scale is used, where the factorization and renormalization scales are defined as $\mu_F = \mu_R = \sqrt{m_W^2 + p_{T,j}^2}$, with $m_W$ the mass of the $W$ boson and $p_{T,j}$ the transverse momentum of the hadronic jet. Because NNLO QCD corrections for $W + c$ production have been presented only very recently [42], in this case we accompany the data with an additional correlated uncertainty, estimated from the 9-point scale variations of the NLO calculation [43,44].

\section*{Nuclear corrections in neutrino DIS

Neutrino-DIS measurements from the NuTeV and NOMAD experiments are subject to nuclear corrections, because they both utilize an iron (Fe) target. In this analysis, however, we do not include such corrections because they are expected to be subdominant in comparison to other sources of uncertainties. For NuTeV, they were found to be moderate in a global fit based on the same methodology used here [45]; for NOMAD, they are known to approximately cancel out in the ratio $R_{\mu\mu}$. To verify this last statement, we recomputed the NOMAD ratio $R_{\mu\mu}$ with the recently presented nNNPDF2.0 NLO Fe nuclear PDF set.
and compared the result with the predictions obtained with the NLO free proton PDF set consistently determined in [46]. The full set of correlations between Fe and proton PDFs were therefore appropriately taken into account. The relative uncertainties.

The exclusion of the ATLAS $W$, $Z$ measurements of [11,12] from the $str_{base}$ fit allows us to quantify the effect of the full LHC strangeness-sensitive dataset by comparing this fit with the $str_{prior}$ one. This second fit, $str_{prior}$, is finally further supplemented with the NOMAD data, specifically the set that depends on $E_{w}$, to determine the third fit ($str$). Bayesian reweighting and unweighting [48,49] are used in this last step, because they allow one to evaluate the two-dimensional integral in Eq. (2.1) only once, a task that would otherwise be computationally very intensive in a fit. After reweighting, one ends up with $N_{eff} = 105$ effective replicas, from which we construct a set of $N_{rep} = 100$ replicas.

We also produced variants of these three fits. First of all, in order to assess the impact of the choice of the specific NOMAD dataset, we performed the $str_{s\_hat}$ fit, except for the fact that the NOMAD $E_{w}$-dependent dataset is replaced by its $\sqrt{s}$-dependent counterpart. In this case, after reweighting one ends up with $N_{eff} = 135$ effective replicas (out of $N_{rep} = 850$ initial replicas), from which we construct an ensemble of $N_{rep} = 100$ replicas. Second, in order to assess the impact of parametrizing the charm PDF, we performed the $str_{prior\_pch}$ and $str_{pch}$ fits. These fits are equivalent to the $str_{prior}$ and $str$ fits, except for the fact

| Fit ID      | Dataset                          | Theory                        |
|------------|----------------------------------|-------------------------------|
| str_base   | Same as [27] –                   | Same as [27] +                |
|            | ATLAS $W$, $Z$ [11,12]           | NNLO $K$-fact NuTeV           |
| str_prior  | Same as str_base +               | Same as str_base +            |
|            | ATLAS $W$, $Z$ (full) [11,12]    | Correlated theory             |
|            | ATLAS $W + c$ [16]               | Uncertainty for               |
|            | CMS $W + c$ [17,18]              | Unknown NNLO QCD              |
|            | ATLAS $W$+jets [15]              | Corrections                   |
| str        | Same as str_prior +              | Same as str_prior +           |
|            | NOMAD $E_{w}$ Set [10]           | NNLO $K$-fact NOMAD           |
| str_s_hat  | Same as str_prior +              | Same as str                    |
|            | NOMAD $\sqrt{s}$ set [10]        | Perturbative charm            |
| str_prior_pch | Same as str_prior               | Same as str_prior_pch +      |
|            | NOMAD $E_{w}$ Set [10]           | NNLO $K$-fact NOMAD           |

Table 1 A list of the PDF fits presented in this work; see the text for details

[46], and compared the result with the predictions obtained with the NLO free proton PDF set consistently determined in [46]. The full set of correlations between Fe and proton PDFs were therefore appropriately taken into account. The relative difference between the two computations (with and without nuclear PDF corrections) turned out to range between 3%, in the lowest $E_{w}$ bin, and a fraction of percent, in the bins at the highest $E_{w}$. These differences are smaller than both the data and the PDF uncertainties, therefore ignoring nuclear PDF uncertainties is a well-justified approximation. We note that nuclear non-isoscalarity effects are treated as in [45]. In the future, it might be interesting to extend the present analysis in a way that systematically accounts for nuclear PDF uncertainties.

Positivity of cross sections We enforce the positivity of the charm structure function $F_{c}^{x}$ with a procedure similar to that described in [47] for light quarks. This additional positivity constraint is required to prevent the fitted charm PDF becoming unphysically negative once the new datasets are included in the fit.

2.3 PDF sets

We assess the impact of the datasets and of the theoretical choices outlined in Sects. 2.1–2.2 on PDFs by performing the series of fits summarized in Table 1. All of them are accurate to NNLO in perturbative QCD (where available), and are based on the NNPDF methodology; see [47] and the references therein for a comprehensive description.

The first fit ($str_{base}$) is our baseline, and corresponds to the fit of [27] with the addition of the NNLO charm-mass $K$-factors for the NuTeV data and of the positivity constraint on $F_{c}^{x}$, and with the removal of the 2010 and 2011 ATLAS $W$, $Z$ inclusive measurements of [11,12]. We will present a comparison of this baseline fit with the NNPDF3.1 PDF set of [19] in Sect. 3.3. This fit is then supplemented with all the new LHC data, including the ATLAS $W$, $Z$ measurements from [11,12], to obtain the second fit ($str_{prior}$), for which we generate $N_{rep} = 850$ Monte Carlo replicas. The exclusion of the ATLAS $W$, $Z$ measurements of [11,12] from the $str_{base}$ fit allows us to quantify the effect of the full LHC strangeness-sensitive dataset by comparing this fit with the $str_{prior}$ one. This second fit, $str_{prior}$, is finally further supplemented with the NOMAD data, specifically the set that depends on $E_{w}$, to determine the third fit ($str$). Bayesian reweighting and unweighting [48,49] are used in this last step, because they allow one to evaluate the two-dimensional integral in Eq. (2.1) only once, a task that would otherwise be computationally very intensive in a fit. After reweighting, one ends up with $N_{eff} = 105$ effective replicas, from which we construct a set of $N_{rep} = 100$ replicas.

We also produced variants of these three fits. First of all, in order to assess the impact of the choice of the specific NOMAD dataset, we performed the $str_{s\_hat}$ fit, except for the fact that the NOMAD $E_{w}$-dependent dataset is replaced by its $\sqrt{s}$-dependent counterpart. In this case, after reweighting one ends up with $N_{eff} = 135$ effective replicas (out of $N_{rep} = 850$ initial replicas), from which we construct an ensemble of $N_{rep} = 100$ replicas. Second, in order to assess the impact of parametrizing the charm PDF, we performed the $str_{prior\_pch}$ and $str_{pch}$ fits. These fits are equivalent to the $str_{prior}$ and $str$ fits, except for the fact
that the charm PDF is generated perturbatively off the gluon and the light-quark PDFs. In this case, we produced only $N_{\text{rep}} = 500$ replicas in the str_prior_pch fit; after reweighting we are left with $N_{\text{eff}} = 157$ effective replicas, from which we constructed an ensemble of $N_{\text{rep}} = 100$ replicas for the str_pch fit.

### 3 Results

In this section we present the main results of our analysis. First, we discuss the quality of the fits that we performed; then, we compare the data to our theoretical predictions; next, we present the PDFs that we determine; and finally, we revisit the strangeness content of the proton in the light of these. We conclude by focusing on the impact of the NOMAD dataset and of the implications that the treatment of the charm PDF has on our results.

#### 3.1 Fit quality

In Table 2 we summarize the values of $\chi^2$ per data point obtained from five of the six fits discussed in Sect. 2.3; see also Table 1: $\chi^2_{\text{base}}$ for str_base; $\chi^2_{\text{str}}$ for str; $\chi^2_{\text{str}_\text{s_hat}}$ for str_s_hat; and $\chi^2_{\text{str}_{\text{pr}}}$ for str_pch. The value of $\chi^2$ of the str_prior_pch fit is not reported, because it is not particularly more informative than the one of the str_pch fit, which includes the complete dataset. In all cases, the value of $\chi^2$ per data point correspond to the definition given in Eqs. (3.2)–(3.3) in [50].

The values in square brackets are for datasets not included in the corresponding fit.

| Process | Dataset | $n_{\text{dat}}$ | $\chi^2_{\text{base}}$ | $\chi^2_{\text{str}}$ | $\chi^2_{\text{str}_\text{s_hat}}$ | $\chi^2_{\text{str}_{\text{pr}}}$ |
|---------|---------|-----------------|------------------------|------------------------|-------------------------------|------------------------|
| vDIS ($\mu\mu$) | NuTeV [9] | 76/76/95/91/95 | 0.70 | 0.71 | 0.53 | 0.52 | 0.63 |
| | NOMAD [10] | $-19/19/19$ | [9.0] | [8.8] | 0.55 | 0.35 | 0.69 |
| | W, Z (incl.) | 327/418/418/418/418 | 3.22 | 1.65 | 1.67 | 1.64 | 1.80 |
| | ATLAS [12] | $-22/22/22/22/22$ | [0.76] | 0.68 | 0.60 | 0.66 | 0.68 |
| | CMS [17,18] | $-15/15/15/15/15$ | [1.10] | 0.98 | 0.96 | 1.00 | 1.00 |
| | ATLAS [16] | $-22/22/22/22$ | [0.53] | 0.48 | 0.42 | 0.43 | 0.46 |
| Total | ATLAS [15] | $-32/32/32/32$ | [1.58] | 1.18 | 1.18 | 1.18 | 1.18 |

The values in square brackets are for datasets not included in the corresponding fit.

In this section we present the main results of our analysis. First, we discuss the quality of the fits that we performed; then, we compare the data to our theoretical predictions; next, we present the PDFs that we determine; and finally, we revisit the strangeness content of the proton in the light of these. We conclude by focusing on the impact of the NOMAD dataset and of the implications that the treatment of the charm PDF has on our results.

We finally assess the effect of reweighting the str_prior fit with each dataset; see the discussion in Sect. 2.3.
Fig. 2 Comparison between the theoretical predictions, obtained from the str_prior and str (left) or str_s_hat (right) fits, and the experimental data for the $E_\nu$- and $\sqrt{s}$-dependent NOMAD measurements. The insets display the ratio to the central value of each data point. The error bands on the theory predictions indicate the one-sigma PDF uncertainties.

Fig. 3 Comparison between theoretical predictions and experimental data for the neutrino (left) and antineutrino (right) charm dimuon cross sections measured by the NuTeV experiment [9]. Data and theory are normalized to the central value of the former; data points are sorted by their ID values, roughly corresponding to increasing $x$ and $Q^2$ values (for fixed pseudo-rapidity bins $y$) when the plots are read from left to right.

1.20 when comparing the str and the str_pch fits. We therefore confirm previous studies indicating that fitting the charm PDFs improves the description of the experimental data within a global PDF analysis.

3.2 Comparison with experimental data

We now compare the strangeness-sensitive datasets included in our analysis with the corresponding theoretical predictions. Our aim is to assess the impact of the various datasets. To this purpose, we compare the fits obtained without and with a specific dataset included.

We first focus on the neutrino-DIS datasets. In Fig. 2 we display the comparison for the $E_\nu$-dependent and the $\sqrt{s}$-dependent NOMAD measurements. We compare the theoretical predictions obtained from the str_prior fit and, respectively, either from the str or the str_s_hat fits. The insets display the ratio to the central value of each measured data point. In the two cases, we observe a consistent picture: the theoretical prediction obtained from the str_prior fit overshoots the data points by about 20% (10%) for the $E_\nu$-dependent ($\sqrt{s}$-dependent) dataset; after reweighting, the theoretical prediction nicely describes the data points with an uncertainty consistently reduced by up to a factor of 4. We explicitly checked that the same reduction of the uncertainty occurs also in the case of perturbative charm fits without (str_prior_pch) and with (str_pch) the $E_\nu$-dependent NOMAD dataset included. In this case, however, the underlying PDFs and the strangeness ratio $R_s$ vary in comparison to fitted charm fits, as discussed in Sect. 3.4.

In Fig. 3 we display the data/theory comparison for the charm dimuon cross sections from the NuTeV measurement of [9] (for both neutrino and antineutrino beams). In this case, predictions are determined from the str_base and str PDF input sets, and are normalized to the central value of the data points. These are sorted by their ID value, roughly cor-
responding to increasing $x$ and $Q^2$ values (for fixed pseudo-rapidity bins $y$) when the plot is read from left to right. A fair agreement between data and theory is observed, as expected from the pattern of the $\chi^2$ values reported in Table 2. The inclusion of the NOMAD data in the str fit suppresses the theoretical expectation for the NuTeV neutrino cross sections (but not for the antineutrino ones, for which no analogue observable is measured by NOMAD); uncertainties are reduced by up to a factor of 2 (again, more markedly for the neutrino data points than for the antineutrino ones). Both the shift in the central value and the reduction of the uncertainty remain smaller than the comparatively large experimental uncertainty.

We now turn to the hadron collider data. In Fig. 4 we display: the $W+c$ lepton rapidity distributions corresponding to the ATLAS measurement of [16] (for both $W^+$ and $W^−$) and to the CMS measurements (sum of $W^+$ and $W^−$) of [17, 18] (respectively at a c.m.e. of 7 TeV and 13 TeV); and the $Z$ dilepton rapidity distributions from the ATLAS measurement of [12] at a c.m.e. of 7 TeV (for both the central and the forward selection cuts). The insets display the ratio of the theory to the central value of the experimental measurement. As in Fig. 3, theoretical predictions are evaluated with the str_base and str PDF sets.

A fair agreement between data and theory is found in all cases, as expected from the pattern of the $\chi^2$ values reported in Table 2. However, we clearly see that the size of the PDF uncertainty relative to the size of the data uncertainty depend on the dataset. Concerning the ATLAS and CMS $W+c$ measurements, experimental uncertainties span the range between 10 and 20%, and are consistently larger than PDF uncertainties. We note that the PDF uncertainties in the theory predictions in the str fit in comparison to str_base fit, as highlighted by the ratios in the insets. Concerning the ATLAS $Z$ distribution, the total experimental uncertainty is much smaller than the $W$ counterpart, around 2% for the central rapidity bin, and in the central region it is comparable to the PDF uncertainty. We therefore expect this measurement to be one of the most constraining among all of the LHC measurements considered in this work. Interestingly, once the NOMAD dataset is included in the fit, the central value of the theoretical prediction approaches the central value of the ATLAS data, and PDF uncertainties are slightly reduced. A similar trend can be observed for the forward selection data. This behavior is a further sign of the good overall compatibility of all of the datasets, and in particular of neutrino DIS and LHC gauge-boson production measurements.

3.3 Parton distributions

We now turn to study the impact of the theoretical assumptions and of the new datasets considered in this analysis on the PDFs. We first present a comparison between the str_base and the NNPDF3.1 parton sets, and then a comparison among the str_base, str_prior and str PDF sets. In the latter case, because the new datasets are expected to mainly affect the strange quark and antiquark distributions, we will focus on the total and valence strange distributions, first, and on the other PDFs, then.

3.3.1 Comparison with NNPDF3.1

Our baseline fit str_base differs from NNPDF3.1 [19] in several respects. As explained in Sect. 2.2, these include: the treatment of inclusive jet production from ATLAS and CMS with NNLO $K$-factors, see [27]; an updated treatment of non-isoscalarity effects in neutrino-DIS data, see [45]; the inclusion of the NNLO massive corrections to the NuTeV structure functions; the new $F_2^s$ positivity constraint; and the correction of the APFEL bug found in the benchmark reported in Fig. 1, which affected the large-$x$ implementation of the NLO coefficient functions. Furthermore, following the motivation presented in Sect. 2.3, the ATLAS $W, Z$ rapidity distributions from [11, 12] that were part of NNPDF3.1 are excluded from str_base.

In order to gauge the impact of all these differences, in Fig. 5 we compare the NNPDF3.1 and str_base parton sets. We display the up (valence and sea), down (valence and sea), strange (valence and total), gluon, and charm distributions at a scale $Q = 100$ GeV; PDFs are normalized to the central value of the NNPDF3.1 parton set, except for the strange valence distribution, for which the absolute PDFs are shown. In addition, we also display in this comparison the results of a variant of str_base obtained without the NNLO $K$-factors, Eq. (2.3), for the NuTeV cross sections, in order to isolate their impact in the resulting PDFs.

In comparison to NNPDF3.1, in the str_base fit we observe: a rearrangement of the quark flavor separation at medium and large $x$; an increase in the central value of the strange PDF for $x \gtrsim 10^{-3}$; a similar effect in the case of the charm PDF for $x \gtrsim 10^{-2}$ (mostly due to the new $F_2^s$ positivity constraint); and a harder gluon at large $x$ (mostly due to the improved NNLO treatment of jet data). All in all, while the two fits agree within uncertainties, the improvements introduced in the str_base fit lead to PDF differences that are sufficiently large to adopt it as baseline in the current study. Therefore the NNPDF3.1 set will not be discussed further in the sequel.

From the comparison of str_base with its variant without the $K$-factors, the most noticeable effect is the moderate suppression of $s^+$ in the region $x \gtrsim 0.05$, which can represent a shift of up to half a sigma in units of the PDF uncertainty. One can also observe a small correlated enhancement of the up and down quark sea distributions in the same region of $x$. 
Fig. 4 Comparison between theoretical predictions and experimental data for some of strangeness-sensitive proton collider measurements used in this work. Top: the $W^+ + c$ lepton rapidity distributions (separately for $W^+$ and $W^-$) corresponding to the ATLAS measurement at 7 TeV [16]. Middle: the $W + c$ lepton rapidity distributions (sum of $W^+$ and $W^-$) corresponding to the CMS measurements at 7 TeV [17] and 13 TeV [18]. Bottom: the $Z$ dilepton rapidity distributions from the ATLAS measurement of [12]. The insets display the theory to data ratio. Theoretical predictions are evaluated with the str_base and str fits.

The impact of the $K$-factors turns out to be negligible for the gluon and other flavour combinations.

3.3.2 Total and valence strange distributions

In Fig. 6 we display the total and valence strange distributions at $Q = 10$ GeV. We compare in turn the PDFs obtained from the str_base, str_prior and str fits, and those obtained from the str fit with other recent parton sets. Specifically, we consider CT18, CT18A [20] (CT18A is a variant of CT18 that includes the ATLAS $W, Z$ data), MMHT14 [51], and ABMP16 [52]. They all include only a subset of the strangeness-sensitive data included in our analysis (see Table 2), in particular: the NuTeV dataset is part of all PDF sets; the NOMAD dataset is only part of ABMP16; and the off-peak and forward ATLAS $W, Z$ bins, the $W + c$ and the $W + jets$ datasets are not part of any of these PDF sets. We also emphasize that, apart from the more extensive dataset, our analysis differs from all of the other PDF determinations shown in Fig. 6 in that the charm-quark PDF is fitted on the same footing as the other light-quark PDFs [53]. This feature was demonstrated to improve the description of
Fig. 5 Comparison between the NNPDF3.1 NNLO fit [19] and the baseline fit used in this work, str_base. From top to bottom and left to right we show the up (valence and sea), down (valence and sea), strange (valence and total), gluon, and charm distributions at a scale $Q = 100$ GeV; PDFs are normalized to the central value of the NNPDF3.1 parton set (ref), except for the strange valence distribution, for which the absolute PDFs are shown.
Fig. 6 A comparison of the total (top) and valence (bottom) strange distributions, \( s^+ \) and \( s_V \), at \( Q = 10 \) GeV. PDFs are from the \texttt{str_base}, \texttt{str_prior} and \texttt{str} fits (left panels) and from the \texttt{str} and other recent PDF fits (right panels); see the text for details. The total strangeness \( s^+ \) is normalized to the central value of the \texttt{str_base} fit. The insets display the corresponding relative (\( \delta s^+/s^+ \)) and absolute (\( \delta s_V \)) PDF uncertainties.

DIS and LHC datasets, and in particular to partially relieve tensions between the NuTeV and the ATLAS \( W, Z \) datasets [19]. The insets in Fig. 6 display the relative and absolute PDF uncertainties for the total (\( \delta s^+/s^+ \)) and valence (\( \delta s_V \)) strange distributions, respectively. In the case of \( s^+ \), the curves are normalized to the central value of the \texttt{str_base} fit.

A comparison among the \texttt{str_base}, \texttt{str_prior} and \texttt{str} fits reveals that the impact of the data is consistent for the total and valence strange distributions. The inclusion of the LHC datasets in the \texttt{str_prior} fit does not alter the central value of the PDFs in a significant way, while it narrows the PDF uncertainty across most of the \( x \) range. The inclusion of the NOMAD dataset in the \texttt{str} fit is associated to a larger effect: the central value of both \( s^+ \) and \( s_V \) is suppressed by about 20% (or more) for \( x \gtrsim 0.1 \); the uncertainty is further reduced by up to a third in the same \( x \) region.

A comparison among the \texttt{str} fit and other recent parton sets reveals differences in the shape of the central value of the \( s^+ \) and \( s_V \) distributions. In the second case, in particular, only MMHT14 [51] allows for a non-zero parametrization. Within the larger uncertainties of the CT18, CT18A and MMHT14 PDF sets, however, results are overall consistent. In this respect, note that the very small uncertainty of the ABM16 set is an artifact of the lack of a tolerance criterion in their analysis. In this case, uncertainties should be rescaled by a factor which is, however, not determined in their analysis.

3.3.3 Light-quark, charm, and gluon PDFs

In Fig. 7 we compare the up (valence and sea), down (valence and sea), gluon and charm distributions resulting from the \texttt{str_base}, \texttt{str_prior} and \texttt{str} fits at \( Q = 100 \) GeV. Results are normalized to the \texttt{str_base} fit.

From these comparisons, we observe that the new datasets have a little impact on the gluon PDF, both on central values and on uncertainties, as expected. A bigger effect is observed instead on the quark PDFs. For light quarks and antiquarks, the electroweak LHC datasets constrain the distributions at low to mid values of \( x, x \lesssim 0.1 \), while the NOMAD datasets do so at larger values of \( x, x \gtrsim 0.1 \). The two datasets are therefore complementary, and concur together to enhance the central value of the down distributions in the region \( 0.01 \lesssim x \sim 0.1 \) by a few percent, and to make all the light valence and sea quark PDFs more precise: overall, uncertainties are reduced by up to a factor 2 in the same region for the \texttt{str} fit.
3.4 The strange content of the proton revisited

We finally revisit the strange content of the proton in the light of our results. To this purpose, we consider the strange fraction of proton quark sea and the corresponding ratio of momentum fraction, respectively, defined as

\begin{align}
R_s(x, Q^2) &= \frac{s(x, Q^2) + \bar{s}(x, Q^2)}{\bar{u}(x, Q^2) + \bar{d}(x, Q^2)} , \\
K_s(Q^2) &= \frac{\int_{0}^{1} dx x [s(x, Q^2) + \bar{s}(x, Q^2)]}{\int_{0}^{1} dx x [\bar{u}(x, Q^2) + \bar{d}(x, Q^2)]} .
\end{align}

(3.1)

We first consider the ratio \(R_s\). In the left panel of Fig. 8 we display it for the \(\text{str}_{\text{base}}, \text{str}_{\text{prior}}\) and \(\text{str}\) fits at a scale \(Q = 10\) GeV as a function of \(x\). The inset displays the associated relative PDF uncertainty \(\delta R_s / R_s\). The impact of the new datasets is clearly visible. Concerning the central value, collider datasets do not alter its expectation (the results obtained from the \(\text{str}_{\text{base}}\) and \(\text{str}_{\text{prior}}\) fits are almost identical); the NOMAD dataset, instead, prefers a

For the charm PDF, the central value is suppressed in the \(\text{str}\) fit; uncertainties are reduced by up to a factor 2 for \(x \simeq 0.05\). This effect is almost entirely due to the NOMAD data, which is indirectly sensitive to the charm PDF through its interplay with the \(s\bar{c}\) and \(\bar{s}c\) contributions to \(W\)-boson production.

Fig. 7 Comparison between the fits presented in this work. From top to bottom and left to right we show the up (valence and sea), down (valence and sea), gluon and charm distributions resulting from the \(\text{str}_{\text{base}}, \text{str}_{\text{prior}}\) and \(\text{str}\) fits at \(Q = 100\) GeV. Results are normalized to the \(\text{str}_{\text{base}}\) fit (ref)
The ratio $R_s$, Eq. (3.1), as a function of $x$ at $Q = 10$ GeV. The PDF used are from the str_base, str_prior and str fits (left), and from the str fit and from recent parton sets (right) see text for details. The insets display the corresponding relative uncertainty $\delta R_s / R_s$.

Fig. 9  Same as Fig. 8, comparing fits to the $E_\nu$-dependent (str) or to the $\sqrt{s}$-dependent (str_s_hat) NOMAD dataset (left), and fits with fitted (str) or perturbative (str_pch) charm (right).

more suppressed strange sea for $x \gtrsim 0.1$. Concerning uncertainties, collider datasets lead to a reduction of the relative uncertainty on $R_s$ of about 4% for $x \lesssim 0.1$; the NOMAD dataset, instead, reduces it by about a factor of 2 for $x \gtrsim 0.1$. Overall, the impact of the new datasets depends on $x$, and is mostly significant for $x = 0.2$, where the uncertainty on $R_s$ is reduced from 20 to 8%. For $x \gtrsim 0.3$ no experimental constraints are available, hence the PDF uncertainty blows up.

The right panel of Fig. 8 compares the ratio $R_s$, computed at a scale $Q = 10$ GeV as a function of $x$, as obtained from the str fit and from the CT18/CT18A [20], MMHT14 [51], and ABMP16 [52] fits. The inset displays the relative PDF uncertainty $\delta R_s / R_s$. Our str determination agrees with the CT18A and ABMP16 results within uncertainties in the data region. However, it overshoots the CT18 and MMHT14 results. Note that the very small PDF uncertainties of the ABMP16 result should be realistically rescaled by a tolerance factor $T = \chi^2 > 1$ [1], which is, however, not accounted for in their analysis. With this caveat, our results for $s^+$ and $R_s$ are also the most precise, in particular around $x \sim 0.1$, thanks to the wider dataset (and specifically of NOMAD) employed to constrain the strange quark and antiquark PDFs.

We explicitly assessed how the results for $R_s$ obtained with our optimal fit str depend on the specific choice of the NOMAD dataset and on the fact that the charm PDF is parametrized on the same footing as light-quark PDFs. In Fig. 9 we display the ratio $R_s$, as a function of $x$ at $Q = 10$ GeV: in the left panel we compare results obtained with the str and str_s_hat fits; in the right panel, we compare results obtained with the str and str_pch fits. In the first case, both the central value and the PDF uncertainties of $R_s$ are very similar. This fact confirms the independence of our results upon the choice of the NOMAD dataset included in the fit. In the second case, while PDF uncertainties turn out to be very similar in both the perturbative and the fitted charm fits, the former prefers a central value which is systematically larger than the one obtained from the latter. The size of the shift, however, is at most as large as one-sigma in units of the PDF uncertainties, in line with previous studies [19].
Fig. 10 The ratio $R_s$, Eq. (3.1), computed at $x = 0.023$, $Q = 1.6$ GeV (left) and $x = 0.13$, $Q = 100$ GeV (right). The PDF sets are from the str_base, str_prior, str fits and from other recent PDF analyses; see the text for details.

Fig. 11 The ratio $K_s$, Eq. (3.1), for $Q = 1.6$ GeV (left) and $Q = 100$ GeV (right). The PDF sets are the same as in Fig. 10.

In order to further investigate how our results compare to those reported in the ATLAS studies [11,12], which claimed a symmetric strange quark sea, in Fig. 10 we display the values of $R_s$ for the base, prior and str fits and for the fits shown in Fig. 8. Here $R_s$ is evaluated for the two kinematic choices outlined in [11,12], first $x = 0.023$ and $Q = 1.6$ GeV, and second for $x = 0.13$ and $Q = 100$ GeV. Figure 10 makes it clear the consistent effect of the new datasets included in our analysis. Considering the results for $R_s$ at $x = 0.023$ and $Q = 1.6$ GeV, the value of $R_s = 0.69 \pm 0.22$ in the str_base fit is made more precise by the LHC datasets, which reduce its uncertainty by about a factor 2, while also increasing its central value, $R_s = 0.76 \pm 0.12$; then the neutrino-DIS NOMAD dataset shifts this number towards a lower value by a half-sigma bringing in also a further moderate reduction of the uncertainty, $R_s = 0.71 \pm 0.10$. We therefore conclude that the result $R_s = 1.13 \pm 0.11$, reported in [12] from an analysis of HERA and ATLAS $W, Z$ data within the xFITTER framework [54], is not compatible with ours, possibly because it is affected by a restricted dataset and/or methodological limitations. Similar remarks apply to the results for $R_s$ at $x = 0.13$ and $Q = 100$ GeV, for which we observe a consistent slight reduction of the central value of $R_s$, and a larger reduction of the uncertainty due to the stronger effect of NOMAD data at larger $x$; see Fig. 8. Our results are compatible, within uncertainties, with those of the other PDF determinations, but they are generally more precise.

We finally consider the ratio of momentum fraction $K_s$, defined in Eq. (3.1). Figure 11 displays $K_s$ for $Q = 1.6$ GeV and $Q = 100$ GeV, respectively. The qualitative interpretation of this quantity is consistent with that of $R_s$, in particular, PDF uncertainties are reduced by a factor of 2 in the str fit with respect to the str_base fit. The values of $K_s$ grow with the scale $Q$, as expected due to DGLAP evolution effects: for instance, using the str fit, one finds $K_s = 0.64 \pm 0.07$ at $Q = 1.6$ GeV, and $K_s = 0.81 \pm 0.04$ at $Q = 100$ GeV. Overall, our final str result indicates that the strange sea is mildly suppressed with respect to the rest of the light sea quarks. The value of $K_s$ lies halfway a highly suppressed ($K_s \sim 0.5$) and a purely symmetric ($K_s = 1$) scenarios. As for $R_s$, our final str result for $K_s$ is in agreement with the determinations obtained by other recent PDF analyses within uncertainties, although our results are generally more precise.
4 Summary

By means of a state-of-the-art global analysis, which combines all the relevant experimental and theoretical inputs, we have achieved a precise determination of the strangeness content of the proton. We have demonstrated the compatibility of a wide range of strangeness-sensitive datasets; quantified their relative impact on the fit; compared our results to other recent global analyses; and assessed the robustness of our results with respect to various methodological choices. Our analysis demonstrates that the strange PDF can be precisely determined and that, after all, the proton is not too strange: the momentum fraction carried by strange quark and antiquark PDFs ranges between about 65% and 80% of the momentum fraction carried by the other light sea quarks in a wide energy range (1.6 GeV < Q ≤ 100 GeV). The present determination of the strangeness content of the proton is found to agree, within uncertainties, with the results of other recent global PDF analyses.

Pivotal to this result is the complementary between the LHC gauge-boson production data and of the charmed-tagged neutrino-DIS data, in particular from the NOMAD experiment. Our PDF set, which combines all this information, is available in the LHAPDF format [55] together with its perturbative charm counterpart from http://nnpdf.mi.infn.it/nnpdf3-1strangeness/

This analysis represents an important input for phenomenology, for instance to carry out improved determinations of fundamental parameters of the SM or to be used as baseline in the determination of nuclear PDFs, where strange distributions are not well known [46, 56, 57]. Our determination of the strange and antistrange quark PDFs could be further stress-tested with more exclusive processes, e.g., measurements of kaon production in semi-inclusive DIS (SIDIS). Studies of the strange PDFs based on SIDIS [58–60] notoriously prefer a suppressed strangeness, but are also subject to the potential bias coming from their sensitivity to the fragmentation of the strange quarks into kaons.

Acknowledgements We are grateful to Jun Gao for providing us with the NNLO K-factors for the NuTeV and the NOMAD measurements, and for help into the benchmark of the corresponding NLO calculations. We thank Valerio Bertone for assistance with the usage of APFEL, Lucian Harland-Lang for discussions on the massive corrections to dimuon production, and Rabah Abdul-Khalek, Amanda Cooper-Sarkar, Stefano Forte, Katerina Lipka, Gavin Pownall and Cameron Voiysey for comments on the manuscript. E.R.N. is supported by the European Commission through the Marie Sklodowska-Curie Action ParDHonS FFS.TMDs (Grant number 752748). J.R. is partially supported by the Netherlands Organization for Scientific Research (NWO), M.U. and S.I. are partially supported by the STFC grant ST/L000385/1 and by the Royal Society grant RG1/EA/180148. The work of M.U. is also funded by the Royal Society grant DH150088 and supported by the European Research Council under the European Union’s Horizon 2020 research and innovation Programme (Grant agreement n.950246).

Data Availability Statement This manuscript has associated data in a data repository. [Authors’ comment: This manuscript has its associated data available from the following url: http://nnpdf.mi.infn.it/nnpdf3-1strangeness/]

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Funded by SCOAP³.

References

1. J. Gao, L. Harland-Lang, J. Rojo, The structure of the proton in the LHC precision era. Phys. Rep. 742, 1–121 (2018). arXiv:1709.04922
2. K. Kovařík, P.M. Nadolsky, D.E. Soper, Hadron structure in high-energy collisions. Rev. Mod. Phys. 92(4), 045003 (2020). arXiv:1905.06957
3. J.J. Eibier, E.R. Nocera, Parton distributions in nucleons and nuclei. Ann. Rev. Nucl. Part. Sci. 70, 1–34 (2020). arXiv:2001.07722
4. ATLAS Collaboration, M. Aaboud et al., Measurement of the W-boson mass in pp collisions at √s = 7 TeV with the ATLAS detector. Eur. Phys. J. C 78(2), 110 (2018). arXiv:1701.07240
5. CMS Collaboration, A.M. Sirunyan et al., Measurement of the weak mixing angle using the forward–backward asymmetry of Drell–Yan events in pp collisions at 8 TeV. Eur. Phys. J. 78(9), 701 (2018). arXiv:1806.00863
6. E. Bagnaschi, A. Vicini, A new look at the estimation of the PDF uncertainties in the determination of electroweak parameters at hadron colliders. arXiv:1910.04726
7. CCFR Collaboration, A.O. Bazarko et al., Determination of the strange quark content of the nucleon from a next-to-leading order QCD analysis of neutrino charm production. Z. Phys. C 65, 189–198 (1995). arXiv:hep-ex/9406007
8. CHORUS Collaboration, A. Kayis-Topaksu et al., Leading order analysis of neutrino induced dimuon events in the CHORUS experiment. Nucl. Phys. B 798, 1–16 (2008). arXiv:0804.1869
9. D. Mason et al., Measurement of the nucleon Strange–Antistrange asymmetry at next-to-leading order in QCD from NuTeV Dimuon data. Phys. Rev. Lett. 99, 192001 (2007)
10. NOMAD Collaboration, O. Samoylov et al., A precision measurement of charm dimuon production in neutrino interactions from the NOMAD experiment. Nucl. Phys. B 876, 339 (2013). arXiv:1308.4750
11. ATLAS Collaboration, G. Aad et al., Determination of the strange quark density of the proton from ATLAS measurements of the W, Z cross sections. Phys. Rev. Lett. 112, arXiv:1203.4051
12. ATLAS Collaboration, M. Aaboud et al., Precision measurement and interpretation of inclusive W⁺, W⁻ and Z/γ∗ production cross sections with the ATLAS detector. Eur. Phys. J. C 77(6), 367. (2017). arXiv:1612.03016
13. ATLAS Collaboration, M. Sutton, The PDF interpretation of the measurement of a vector boson produced in association with jets at the ATLAS detector. PoS DIS2019, 034 (2019)
14. W. Stirling, E. Vryonidou, Charm production in association with an electroweak gauge boson at the LHC. Phys. Rev. Lett. 109, 082002 (2012). arXiv:1203.6781

15. ATLAS Collaboration, M. Aaboud et al., Measurement of differential cross sections and $W^+/W^-$ cross-section ratios for $W$ boson production in association with jets at $\sqrt{s} = 8$ TeV with the ATLAS detector. JHEP 05, 077 (2018). arXiv:1711.03296

16. ATLAS Collaboration, G. Aad et al., Measurement of the production of a $W$ boson in association with a charm quark in $pp$ collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. JHEP 1045, 068 (2014). arXiv:1402.6263

17. CMS Collaboration, S. Chatrchyan et al., Measurement of associated $W +$ charm production in $pp$ collisions at $\sqrt{s} = 7$ TeV. JHEP 02, 013 (2014). arXiv:1310.1138

18. CMS Collaboration, A.M. Sirunyan et al., Measurement of associated production of a $W$ boson and a charm quark in proton–proton collisions at $\sqrt{s} = 13$ TeV. Eur. Phys. J. C 79(3), 260 (2019). arXiv:1811.10021

19. NNPDF Collaboration, R.D. Ball et al., Parton distributions for the strong coupling constant within a global PDF analysis of a collider: the APPLGRID Project. Eur. Phys. J. C 66, 503 (2010). arXiv:0911.2985

20. T.-J. Hou et al., New CTEQ global analysis of quantum chromodynamics with high-precision data from the LHC. arXiv:1410.8849

21. A.G. De Ridder, T. Gehrmann, E.W.N. Glover, A. Huss, T.A. Morgan, Precise QCD predictions for the production of a $Z$ boson in association with a hadronic jet. Phys. Rev. Lett. 117(2), 022001 (2016).

22. R. Boughezal, J.M. Campbell, R.K. Ellis, C. Focke, W. Giele, X. Liu, F. Petriello, C. Williams, Color singlet production at NNLO in MCFM. Eur. Phys. J. C 77(1), 7 (2017). arXiv:1605.08011

23. T. Carli et al., A posteriori inclusion of parton density functions in NLO QCD final-state calculations at hadron colliders: the APPLGRID Project. Eur. Phys. J. C 66, 503 (2010). arXiv:0911.2985

24. R. Gavin, Y. Li, F. Petriello, S. Quackenbush, FEWZ 2.0: a code for hadronic $Z$ production at next-to-next-to-leading order. Comput. Phys. Commun. 182, 2388–2403 (2011). arXiv:1011.3540

25. S. Alekhin, J.M. Campbell, R.K. Ellis, C. Focke, W. Giele, X. Liu, F. Petriello, C. Williams, Color singlet production at NNLO in MCFM. Eur. Phys. J. C 77(1), 7 (2017). arXiv:1605.08011

26. R. Boughezal, C. Focke, X. Liu, F. Petriello, W-boson production in association with a jet at next-to-next-to-leading order in perturbative QCD. Phys. Rev. Lett. 115(6), 062002 (2015). arXiv:1504.02131

27. R. Boughezal, J.M. Campbell, R.K. Ellis, C. Focke, W. Giele, X. Liu, F. Petriello, W-boson production in association with a jet at next-to-next-to-leading order in perturbative QCD. Phys. Rev. Lett. 115(6), 062002 (2015). arXiv:1504.02131

28. R.A. Khalek, J.J. Ethier, J. Rojo, G. van Weelden, nNNPDF2.0: next-to-next-to-leading order parton density functions. JHEP 05, 014 (2017). arXiv:1610.08011

29. M. Czakon, A. Mitov, M. Pellen, R. Poncelet, NNLO QCD predictions for $W+c$-jet production at the LHC. arXiv:2011.01011

30. NNPDF Collaboration, R. Abdul Khalek et al., Parton distributions with theory uncertainties: general formalism and first phenomenological studies. Eur. Phys. J. C 79(11), 931 (2019). arXiv:1906.10698

31. NNPDF Collaboration, R. Abdul Khalek et al., A first determination of parton distributions with theoretical uncertainties. Eur. Phys. J. C 79, 838 (2019). arXiv:1905.04311

32. NNPDF Collaboration, R.D. Ball, E.R. Nocera, R.L. Pearson, Nuclear Uncertainties in the Determination of Proton PDFs. Eur. Phys. J. C 79(3), 282 (2019). arXiv:1812.09074

33. A.G. De Ridder, T. Gehrmann, E.W.N. Glover, A. Huss, T.A. Morgan, Precise QCD predictions for the production of a $Z$ boson in association with a hadronic jet. Phys. Rev. Lett. 117(2), 022001 (2016).

34. R.A. Khalek, J.J. Ethier, J. Rojo, G. van Weelden, nNNPDF2.0: quark flavor separation in nuclei from LHC data. JHEP 09(1), 014 (2017). arXiv:1610.08011

35. M. Czakon, A. Mitov, M. Pellen, R. Poncelet, NNLO QCD predictions for $W+c$-jet production at the LHC. arXiv:2011.01011

36. NNPDF Collaboration, R. Abdul Khalek et al., Parton distributions with theory uncertainties: general formalism and first phenomenological studies. Eur. Phys. J. C 79(11), 931 (2019). arXiv:1906.10698

37. NNPDF Collaboration, R. Abdul Khalek et al., A first determination of parton distributions with theoretical uncertainties. Eur. Phys. J. C 79, 838 (2019). arXiv:1905.04311

38. NNPDF Collaboration, R.D. Ball, E.R. Nocera, R.L. Pearson, Nuclear Uncertainties in the Determination of Proton PDFs. Eur. Phys. J. C 79(3), 282 (2019). arXiv:1812.09074

39. R.A. Khalek, J.J. Ethier, J. Rojo, G. van Weelden, nNNPDF2.0: quark flavor separation in nuclei from LHC data. JHEP 09, 183 (2015). arXiv:1504.02131

40. NNPDF Collaboration, R.D. Ball et al., Parton distributions for the LHC Run II. JHEP 04, 040 (2015). arXiv:1410.8849

41. R.D. Ball, V. Bertone, F. Cerutti, L. Del Debbio, S. Forte et al., Reweighting and unweighting of parton distributions and the LHC W lepton asymmetry data. Nucl. Phys. B 855, 608–638 (2012). arXiv:1108.1758

42. The NNPDF Collaboration, R.D. Ball et al., Reweighting NNPDFs: the W lepton asymmetry. Nucl. Phys. B 849, 112–143 (2011). arXiv:1012.0836

43. E.R. Nocera, M. Ubiali, C. Vaisey, Single top production in PDF fits. JHEP 05, 067 (2020). arXiv:1912.05443
determination of the charm content of the proton. Eur. Phys. J. C 76(11), 647 (2016). arXiv:1605.06515
54. S. Alekhin et al., HERAFitter. Eur. Phys. J. C 75(7), 304 (2015). arXiv:1410.4412
55. A. Buckley, J. Ferrando, S. Lloyd, K. Nordström, B. Page et al., LHAPDF6: parton density access in the LHC precision era. Eur. Phys. J. C 75, 132 (2015). arXiv:1412.7420
56. K.J. Eskola, P. Paakkinen, H. Paukkunen, C.A. Salgado, EPPS16: nuclear parton distributions with LHC data. Eur. Phys. J. C 77(3), 163 (2017). arXiv:1612.05741
57. A. Kusina et al., Impact of LHC vector boson production in heavy ion collisions on strange PDFs. arXiv:2007.09100
58. HERMES Collaboration, A. Airapetian et al., Reevaluation of the parton distribution of strange quarks in the nucleon. Phys. Rev. D 89(9), 097101 (2014). arXiv:1312.7028
59. I. Borsa, R. Sassot, M. Stratmann, Probing the sea quark content of the proton with one-particle-inclusive processes. Phys. Rev. D 96(9), 094020 (2017). arXiv:1708.01630
60. JAM Collaboration, N. Sato, C. Andres, J. Ethier, W. Melnitchouk, Strange quark suppression from a simultaneous Monte Carlo analysis of parton distributions and fragmentation functions. Phys. Rev. D 101(7), 074020 (2020). arXiv:1905.03788