



# A first unbiased global determination of polarized PDFs and their uncertainties

NNPDF Collaboration

Emanuele R. Nocera<sup>b</sup>, Richard D. Ball<sup>a</sup>, Stefano Forte<sup>b,\*</sup>,  
Giovanni Ridolfi<sup>c</sup>, Juan Rojo<sup>d,e</sup>

<sup>a</sup> *Higgs Centre, University of Edinburgh, JCMB, KB, Mayfield Rd, Edinburgh EH9 3JZ, Scotland, United Kingdom*

<sup>b</sup> *Dipartimento di Fisica, Università di Milano and INFN, Sezione di Milano, Via Celoria 16, I-20133 Milano, Italy*

<sup>c</sup> *Dipartimento di Fisica, Università di Genova and INFN, Sezione di Genova, Genova, Italy*

<sup>d</sup> *PH Department, TH Unit, CERN, CH-1211 Geneva 23, Switzerland*

<sup>e</sup> *Rudolf Peierls Centre for Theoretical Physics, 1 Keble Road, University of Oxford, OX1 3NP Oxford, United Kingdom*

Received 25 June 2014; received in revised form 17 August 2014; accepted 18 August 2014

Available online 22 August 2014

Editor: Tommy Ohlsson

---

## Abstract

We present a first global determination of spin-dependent parton distribution functions (PDFs) and their uncertainties using the NNPDF methodology: NNPDF<sub>p011.1</sub>. Longitudinally polarized deep-inelastic scattering data, already used for the previous NNPDF<sub>p011.0</sub> PDF set, are supplemented with the most recent polarized hadron collider data for inclusive jet and  $W$  boson production from the STAR and PHENIX experiments at RHIC, and with open-charm production data from the COMPASS experiment, thereby allowing for a separate determination of the polarized quark and antiquark PDFs, and an improved determination of the medium- and large- $x$  polarized gluon PDF. We study the phenomenological implications of the NNPDF<sub>p011.1</sub> set, and we provide predictions for the longitudinal double-spin asymmetry for semi-inclusive pion production at RHIC.

© 2014 CERN for the benefit of the NNPDF Collaboration. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/3.0/>). Funded by SCOAP<sup>3</sup>.

---

\* Corresponding author.

## 1. A global polarized PDF determination

In a recent paper, we presented the NNPDF<sub>p011.0</sub> parton set [1], a first unbiased determination of polarized parton distribution functions (PDFs) of the proton and their associated uncertainties based on the NNPDF methodology [2–5]. This methodology differs from that used in other recent next-to-leading order (NLO) analyses [6–10], in that it relies on a Monte Carlo sampling and representation of PDFs, and it uses a parametrization of PDFs based on neural networks with a very large number of free parameters.

The NNPDF<sub>p011.0</sub> parton set was determined from all available inclusive deep-inelastic scattering (DIS) data with longitudinally polarized beams. One important limitation of only using (neutral-current) inclusive DIS is that only the quark PDF combinations  $\Delta u^+ = \Delta u + \Delta \bar{u}$ ,  $\Delta d^+ = \Delta d + \Delta \bar{d}$ ,  $\Delta s^+ = \Delta s + \Delta \bar{s}$ , and the gluon  $\Delta g$  are accessible. Furthermore, in DIS the gluon is mostly determined by scaling violations, and thus subject to sizable uncertainties due to the restricted lever-arm in  $Q^2$  of polarized DIS measurements.

In recent years, the set of experimental data which may be used for the determination of longitudinally polarized PDFs has been extended impressively. They include now semi-inclusive DIS (SIDIS) in fixed-target experiments [11–15], one- or two-hadron and open-charm production in lepton–nucleon scattering [16–20], and semi-inclusive particle production [21–27], high- $p_T$  jet production [28–30] and parity-violating  $W^\pm$  boson production [31–33] in polarized proton–proton collisions. Many of these data probe individual quark flavors separately, or combinations of them. For instance, semi-inclusive DIS and  $W^\pm$  production data allow one to determine the light quark–antiquark separation of polarized PDFs. In addition, inclusive jet and pion production in polarized proton–proton collisions, as well as hadron or open-charm electroproduction provide a handle on the polarized gluon. Available measurements that provide information on polarized PDFs, the corresponding leading partonic subprocesses, the PDFs that are being probed, and the approximate range of  $x$  and  $Q^2$  covered by available data, are summarized in Table 1.

It is clear from Table 1 that currently available processes do not extend significantly the kinematic coverage of polarized DIS data, even though they do provide independent information, and, in some cases, first information on some PDF combinations. It follows that only moderate improvements are expected from these data on the first moments of polarized PDFs, which are limited by the extrapolation to the unconstrained small- $x$  region. Only future accelerators, such as a high-energy polarized Electron–Ion Collider (EIC) [34–36] or a neutrino factory [37], could extend significantly the coverage of the small- $x$  regime and improve knowledge of the first moments of polarized PDFs [38–40].

The goal of this paper is to include in the NNPDF determination of polarized PDFs the new experimental information provided by polarized hadron collider data. This will lead to our first global polarized PDF set: NNPDF<sub>p011.1</sub>. We will not include in our global fit processes which require knowledge of fragmentation functions for light quarks, such as for instance SIDIS or pion production. Fragmentation functions are on the same footing as PDFs: they can only be determined from a fit to experimental data [41–47], and as such they are subject to the same potential sources of bias. Because our methodology aims at reducing bias, and a determination of fragmentation functions based on our methodology is not yet available, we prefer not to use data which require their use for the time being. We do, however, include the open-charm lepto-production data, because the fragmentation function for heavy quarks is almost computable in perturbation theory and only introduces a very moderate uncertainty.

Our PDF determination thus includes, in addition to DIS data, open-charm production data from the COMPASS experiment at CERN and the most recent high- $p_T$  inclusive jet and  $W^\pm$

Table 1

Summary of available processes that allow for the determination of polarized PDFs. For each process, we show the leading partonic subprocesses, polarized PDFs, and the approximate ranges of  $x$  and  $Q^2$  that are accessible using the data from the references given. Processes listed in the upper part of the table do not depend on fragmentation functions, while those in the lower part of the table do.

Reaction	Partonic subprocess	PDF probed	$x$	$Q^2$ [GeV <sup>2</sup> ]	Ref.
$\ell^\pm\{p, d, n\} \rightarrow \ell^\pm X$	$\gamma^* q \rightarrow q$	$\Delta q + \Delta\bar{q}$	$0.003 \lesssim x \lesssim 0.8$	$1 \lesssim Q^2 \lesssim 70$	[51–60]
$\vec{p}\vec{p} \rightarrow \text{jet}(s)X$	$gg \rightarrow gq$ $qg \rightarrow qg$	$\Delta g$	$0.05 \lesssim x \lesssim 0.2$	$30 \lesssim p_T^2 \lesssim 800$	[28–30]
$\vec{p}p \rightarrow W^\pm X$	$u_L \bar{d}_R \rightarrow W^+$ $d_L \bar{u}_R \rightarrow W^-$	$\Delta u \Delta\bar{u}$ $\Delta d \Delta\bar{d}$	$0.05 \lesssim x \lesssim 0.4$	$\sim M_W^2$	[31–33]
$\ell^\pm\{p, d\} \rightarrow \ell^\pm DX$	$\gamma^* g \rightarrow c\bar{c}$	$\Delta g$	$0.06 \lesssim x \lesssim 0.2$	$0.04 \lesssim p_T^2 \lesssim 4$	[20]
$\ell^\pm\{p, d\} \rightarrow \ell^\pm hX$	$\gamma^* q \rightarrow q$	$\Delta u \Delta\bar{u}$ $\Delta d \Delta\bar{d}$	$0.005 \lesssim x \lesssim 0.5$	$1 \lesssim Q^2 \lesssim 60$	[11–15]
$\vec{p}\vec{p} \rightarrow \pi X$	$gg \rightarrow gg$ $qg \rightarrow qg$	$\Delta g$	$0.05 \lesssim x \lesssim 0.4$	$1 \lesssim p_T^2 \lesssim 200$	[21–27]

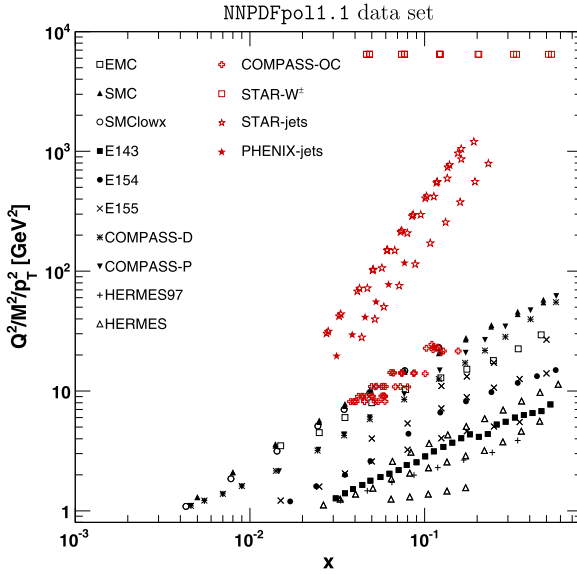


Fig. 1. Kinematic coverage in the  $(x, Q^2)$  plane of the new experimental data included in NNPDFpo11.1 (red points, listed in Table 1) together with that of the inclusive DIS data already in NNPDFpo11.0 (black points). The new experiments are listed in the second column of the legend. For hadronic data, LO kinematics are assumed. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

production data from the STAR and PHENIX experiments at RHIC. The kinematic coverage of these data is shown in Fig. 1, together with that of the fixed-target inclusive DIS data already included in [1].

Other available polarized PDF sets include some of the non-DIS data of Table 1: in particular, the fits from the DSSV family (DSSV08 [6] and DSSV+/DSSV++ [48]) include SIDIS

data, inclusive jet and identified hadron production measurements from polarized proton–proton collisions at RHIC, while the LSS10 fit [7] includes SIDIS data.

The new data sets will be added to those already included in the NNPDF<sub>p011.0</sub> polarized PDF determination [1] using the Bayesian reweighting method described in Refs. [49,50]. This methodology consists of updating the representation of the probability distribution in the space of PDFs provided by an available PDF set by means of Bayes' theorem in such a way that the information contained in the new data sets is included. The method has the dual advantage of not involving any further approximation once the starting Monte Carlo set is given – no parametrization or minimization is necessary – and also of being computationally rather light, in that the predictions to be compared to the new data can be made only once, rather than at each iteration of a minimization algorithm. On the other hand, the method becomes impractical if the new data bring in a large amount of new information. Indeed reweighting has the effect of *zooming in* on the part of the space of PDFs which is compatible with the new data, by giving small weights to replicas which have little compatibility with them. As a consequence, after reweighting the number of replicas in the Monte Carlo set is effectively smaller than the starting one, and thus only if the starting number of replicas was sufficiently large will the final representation of the probability density remain accurate. If too much new information is brought in by the new data, the method becomes impractical because a very large number of starting replicas would be required in order to obtain accurate results after reweighting.

The reweighting method is especially useful in our case because on the one hand, fast interfaces for the computation of hadronic observables, such as the FastNLO framework [61], the general-purpose interface APPLgrid [62], and the FastKernel method [5], used for the unpolarized PDF determinations in Ref. [1] are not yet available in the polarized case.<sup>1</sup> On the other hand, we wish to add to our data set only a few dozen hadronic data points, out of a total of approximately 300 data points. Nevertheless, as we shall see, the construction of the prior probability distribution will require some care, due to the fact that the new data affect some combinations of PDFs which were completely undetermined before reweighting.

As a consequence, we will start in Section 2 with a discussion on how to set up the reweighting method, and in particular how to construct a suitable prior. We will then analyze separately the impact of each of the new processes which are included in our determination on polarized PDFs: first, in Section 3, open-charm and jet production data, which affect the gluon PDF  $\Delta g$ , and then in Section 4,  $W^\pm$  production which allows for a determination of the light antiquark PDF  $\Delta \bar{u}$  and  $\Delta \bar{d}$ . Finally, in Section 5, we will discuss a simultaneous reweighting with all new data sets and present our final PDF set, NNPDF<sub>p011.1</sub>. Specifically, we will discuss the phenomenological implications of NNPDF<sub>p011.1</sub> compared to NNPDF<sub>p011.0</sub>: we will reassess the spin content of the proton by computing PDF first moments, and give an illustrative example of the predictive power of our new PDF set for the case of longitudinal double-spin asymmetry for single-inclusive particle production in proton–proton collisions recently measured at RHIC.

## 2. Reweighting: construction of the prior

We wish to include the polarized hadronic data into our polarized PDF determination by means of Bayesian reweighting. As a starting point, we construct a set of  $N_{\text{rep}} = 1000$  NNPDF<sub>p011.0</sub> PDF replicas according to the procedure of Ref. [1]: a large prior set is needed because

<sup>1</sup> Note however that recent progress on interfacing MadGraph5\_aMC@NLO [63] to APPLgrid will provide the required fast computations in the near future.

reweighting always entails some loss of efficiency, so the final reweighted replica set corresponds to a smaller effective number of unweighted replicas [49,50].

However, this is not sufficient, because the NNPDF polarized parton set NNPDFp011.0 [1] does not allow for a separation between quark and antiquark parton distributions, since it was determined from a fit to inclusive DIS data only. Because the new data that we wish to include are sensitive to such separation, we need to supplement the prior based on NNPDFp011.0 with some assumption on the  $\Delta q - \Delta \bar{q}$  PDF combinations which are left unspecified in it.

Clearly, choosing a completely unbiased flat prior for the PDF combinations which are undetermined in NNPDFp011.0 would be extremely inefficient, given that PDFs span a space of functions. We will thus choose a prior based on an existing PDF set in which these PDF combinations are determined, and then check that our reweighted results are independent of this choice by varying the prior. In practice, we construct the prior by supplementing the PDFs which are determined in the NNPDFp011.0 set, namely  $\Delta u^+$ ,  $\Delta d^+$ ,  $\Delta s^+$  and  $\Delta g$ , with  $\Delta \bar{u}$  and  $\Delta \bar{d}$  from the DSSV08 [6] set, but with the uncertainties inflated by a given factor. Of course, if the uncertainty was infinite this would be equivalent to an (unbiased) flat prior. Hence, we will verify that our results are independent of the choice of prior by inflating the uncertainty by an increasingly large factor, until the results stabilize.

In order to do this, we sample the DSSV08  $\Delta \bar{u}$  and  $\Delta \bar{d}$  distributions at a fixed reference scale  $Q_0^2 = 1 \text{ GeV}^2$ . We then select ten points, half logarithmically and half linearly spaced in the interval of momentum fraction  $10^{-3} \lesssim x \lesssim 0.4$ , which roughly corresponds to the range covered by SIDIS experimental data relevant for separating quark–antiquark contributions. These points with the corresponding PDF uncertainties are treated as sets of experimental pseudo-observables. Henceforth, they will be labeled as DSSV<sub>U</sub> and DSSV<sub>D</sub> respectively. In these pseudo-data prior fits, the experimental data are always taken as the central value from the DSSV08 best fit, while the experimental uncertainties are the corresponding nominal  $\Delta\chi^2 = 1$  Hessian uncertainties, multiplied by a factor one, two, three and four respectively in order to obtain the sets with inflated uncertainty as explained above.

We then generate  $N_{\text{rep}} = 1000$  replicas of the original pseudo-data, following the procedure described in Section 2 of Ref. [1], and we fit each replica with a set of neural networks. To this purpose, we supplement the input PDF basis given in Section 3 of Ref. [1], namely  $\Delta\Sigma$ ,  $\Delta T_3$ ,  $\Delta T_8$  and  $\Delta g$ , with two new linearly independent light quark combinations: the total valence,  $\Delta V$ , and the valence isotriplet,  $\Delta V_3$ ,

$$\Delta V(x, Q_0^2) = \Delta u^-(x, Q_0^2) + \Delta d^-(x, Q_0^2), \quad (1)$$

$$\Delta V_3(x, Q_0^2) = \Delta u^-(x, Q_0^2) - \Delta d^-(x, Q_0^2), \quad (2)$$

where  $\Delta q^- = \Delta q - \Delta \bar{q}$ ,  $q = u, d$ . Eq. (1) holds under the assumption that  $\Delta s = \Delta \bar{s}$ , i.e.  $\Delta V_8 = \Delta V$ . This assumption is not based on a theoretical motivation, but simply on the observation that present data are insufficient to determine  $\Delta s^-$ : hence this PDF combination should be simply viewed as undetermined in our fit.<sup>2</sup>

Each of the PDF combinations in Eqs. (1)–(2) is parametrized as usual by means of a neural network supplemented with a preprocessing function,

$$\Delta V(x, Q_0^2) = (1-x)^{m_{\Delta V}} x^{n_{\Delta V}} \text{NN}_{\Delta V}(x), \quad (3)$$

$$\Delta V_3(x, Q_0^2) = (1-x)^{m_{\Delta V_3}} x^{n_{\Delta V_3}} \text{NN}_{\Delta V_3}(x), \quad (4)$$

<sup>2</sup> In the corresponding LHAPDF grids, which require input of both  $\Delta s$  and  $\Delta \bar{s}$ , we will assume  $\Delta s^- = 0$ .

Table 2  
Ranges for the small- and large- $x$  preprocessing exponents in Eqs. (3)–(4).

PDF	$m$	$n$
$\Delta V(x, Q_0^2)$	[1.5, 3.0]	[0.05, 0.60]
$\Delta V_3(x, Q_0^2)$	[1.5, 3.0]	[0.01, 0.60]

Table 3  
The value of the  $\chi_{\text{tot}}^2$  per data point for both separate and combined  $\Delta\bar{u}$  and  $\Delta\bar{d}$  data sets after the neural network fit to  $\Delta\bar{u}$  and  $\Delta\bar{d}$  pseudo-data sampled from DSSV08 [6].

Experiment	Set	$N_{\text{dat}}$	$\chi_{\text{tot}}^2/N_{\text{dat}}$			
			$1\sigma$	$2\sigma$	$3\sigma$	$4\sigma$
DSSV		20	1.04	1.10	1.09	0.97
	DSSV <sub>U</sub>	10	1.13	1.09	1.08	0.97
	DSSV <sub>D</sub>	10	0.96	1.10	1.09	0.96

where  $\text{NN}_{\Delta\text{pdf, pdf}} = V, V_3$  is the output of the neural network, and the preprocessing exponents  $m, n$  are linearly randomized for each Monte Carlo replica within the ranges given in Table 2. We have checked that our choice of preprocessing exponents does not bias the fit, according to the procedure discussed in Section 4.1 of Ref. [1]. The neural network architecture is the same as in NNPDFp011.0, namely 2-5-3-1.

The DSSV08 PDFs provide us with pseudo-data for  $\Delta\bar{u}$  and  $\Delta\bar{d}$  which in terms of the PDF basis are given by

$$\Delta\bar{u}(x, Q_0^2) = \frac{1}{12}[2\Delta\Sigma + 3\Delta T_3 + \Delta T_8 - 3\Delta V - 3\Delta V_3](x, Q_0^2), \quad (5)$$

$$\Delta\bar{d}(x, Q_0^2) = \frac{1}{12}[2\Delta\Sigma - 3\Delta T_3 + \Delta T_8 - 3\Delta V + 3\Delta V_3](x, Q_0^2). \quad (6)$$

Each DSSV08 pseudo-data replica is then combined at random with an NNPDFp011.0 PDF replica, and the two missing basis combinations Eqs. (1)–(2) are determined by fitting with the standard NNPDF methodology, including the theoretical constraints which are relevant in the polarized case, as discussed in Ref. [1]. In particular, the positivity constraints Eqs. (61)–(62) of Ref. [1] have been enforced by letting  $f = u, \bar{u}, d, \bar{d}$  separately. Note that no additional sum rules affect  $\Delta V$  and  $\Delta V_3$ . As a consequence, the new PDF combinations  $\Delta V$  and  $\Delta V_3$  are completely uncorrelated to the PDF combinations of the NNPDFp011.0 set, as they are based on completely independent information and there is no further theoretically-induced cross-talk. The quality of the pseudo-data fits is quantitatively assessed by the  $\chi^2$  values per data point quoted in Table 3 which are close to one for both DSSV<sub>U</sub> and DSSV<sub>D</sub> data sets, and their combination.

We thus end up with four separate prior PDF ensembles, labeled as  $1\sigma, 2\sigma, 3\sigma$  and  $4\sigma$ , corresponding to the different factors by which the DSSV08 nominal PDF uncertainty has been enlarged. In Sections 3–4 we will explicitly show that this is sufficient to obtain reweighted results which are independent of the choice of prior, the  $3\sigma$  and  $4\sigma$  sets both being effectively unbiased priors.

In Fig. 2, we show the  $x\Delta\bar{u}(x, Q_0^2)$  and  $x\Delta\bar{d}(x, Q_0^2)$  PDFs at the initial energy scale  $Q_0^2 = 1 \text{ GeV}^2$  from the  $1\sigma$  and  $4\sigma$  sets. The other priors,  $2\sigma$  and  $3\sigma$ , consistently provide intermediate

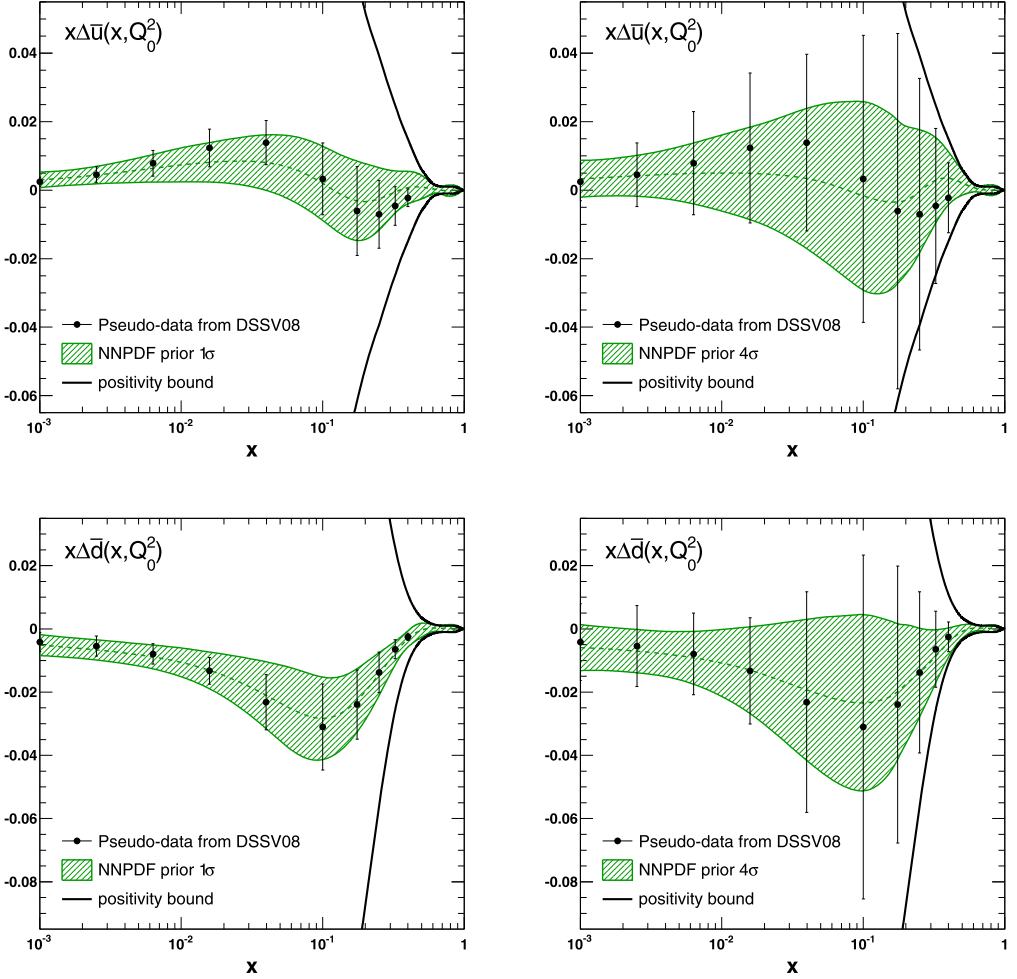


Fig. 2. The polarized sea quark densities,  $x\Delta\bar{u}(x, Q_0^2)$  (upper plots) and the  $x\Delta\bar{d}(x, Q_0^2)$  (lower plots) at the initial energy scale  $Q_0^2 = 1 \text{ GeV}^2$  from the neural network fit (hatched band) to the DSSV08 pseudo-data (points with uncertainties). Results are shown for the  $1\sigma$  (left plots) and  $4\sigma$  (right plots) prior ensembles (see text). The PDF positivity bounds from the corresponding unpolarized NNPDF2.3 counterpart are also shown.

results. In these plots, the positivity bound discussed in Ref. [1] and pseudo-data points sampled from DSSV08 are also shown.

These priors will be used in the next section for the inclusion of the new data sets, listed in the upper part of Table 1.

### 3. The polarized gluon: open charm and jet production

In this section, we include by reweighting the information coming from the data of Table 1 which mostly affect the gluon distributions: open-charm production data from the COMPASS experiment at CERN, and high- $p_T$  jet production measurements in proton–proton collisions from

the STAR and PHENIX experiments at RHIC. We discuss the inclusion of each of these two data sets in turn.

### 3.1. Open-charm muoproduction at COMPASS

Open-charm production in polarized DIS [64] directly probes the gluon distribution because at leading-order (LO) this process proceeds through photon–gluon fusion (PGF),  $\gamma^* g \rightarrow c\bar{c}$ , followed by the fragmentation of the charm quarks into charmed mesons, typically  $D^0$  mesons. The corresponding measurable photon–nucleon asymmetry  $A_{LL}^{\gamma N \rightarrow D^0 X}$  at LO is thus given by

$$A_{LL}^{\gamma N \rightarrow D^0 X} \equiv \frac{d\Delta\sigma_{\gamma N}}{d\sigma_{\gamma N}} = \frac{d\Delta\hat{\sigma}_{\gamma g} \otimes \Delta g \otimes D_c^{D^0}}{d\hat{\sigma}_{\gamma g} \otimes g \otimes D_c^{D^0}}, \quad (7)$$

where  $\Delta\hat{\sigma}_{\gamma g}$  ( $\hat{\sigma}_{\gamma g}$ ) is the spin-dependent (spin-averaged) partonic cross-section,  $\Delta g$  ( $g$ ) is the polarized (unpolarized) gluon PDF, and  $D_c^{D^0}$  is the fragmentation function for a charm quark into a  $D^0$  meson, assumed to be spin independent.

The COMPASS Collaboration has measured the photon–nucleon asymmetry  $A_{LL}^{\gamma N \rightarrow D^0 X}$ , Eq. (7), obtained by the scattering of polarized muons of energy  $E_\mu = 160$  GeV off a fixed target of longitudinally polarized protons or deuterons [20]. Three different data sets are available, depending on the  $D^0$  decay mode used to reconstruct the charmed hadron in the final state:  $D^0 \rightarrow K^-\pi^+$ ,  $D^0 \rightarrow K^-\pi^+\pi^0$  or  $D^0 \rightarrow K^-\pi^+\pi^+\pi^-$ . In the following, these will be referred to as COMPASS  $K1\pi$ , COMPASS  $K2\pi$  and COMPASS  $K3\pi$  respectively. Assuming LO kinematics, the polarized gluon PDF is being probed at intermediate momentum fraction values,  $0.06 \lesssim x \lesssim 0.22$ , and at energy scale  $Q^2 = 4(m_c^2 + p_T^2) \sim 13$  GeV<sup>2</sup>, where  $m_c$  is the charm quark mass and  $p_T$  is the transverse momentum of the produced charmed hadron, as shown in Fig. 1 and Table 1.

In order to include COMPASS open-charm muoproduction data in our polarized PDF determination [1] via reweighting, we need to compute the predictions for the virtual photon–nucleon asymmetry  $A_{LL}^{\gamma N \rightarrow D^0 X}$  given in Eq. (7). We perform this computation, separately for the numerator and the denominator of Eq. (7), using the LO expressions in Ref. [65], to which we refer the reader for more details. The NLO prediction is also available [66]; however, as we shall see shortly, these data have a negligible impact, so much so that results would be essentially unchanged by the use of NLO theory.

Notice that here we do not need the prior Monte Carlo samples constructed in Section 2, since the asymmetry (7) only depends on the polarized gluon. We use the Peterson parametrization of the fragmentation function  $D_c^{D^0}$  [67] with  $\epsilon = 0.06$ ; we checked that results are unaffected by reasonable variations of the fragmentation function [68], and indeed it was pointed out in Ref. [66] that the dependence on the fragmentation function is weaker than scale uncertainties of the NLO computation.

To get a feeling for the potential impact of these data, in Fig. 3 we compare the LO predictions obtained using various PDF sets to the COMPASS data, separated into individual decay channels and three bins for the energy  $E_{D^0}$  of the charmed hadron. Specifically, we show results obtained using DSSV08, AAC08 [8], BB10 [9] and NNPDFpol1.0 polarized sets supplemented with the following unpolarized sets: CTEQ6 [69] for DSSV08; MRST2004 [70] for AAC08 and BB10; and NNPDF2.3 [71] (for NNPDFpol1.0). In all cases, the PDF uncertainties shown are obtained neglecting the uncertainties due to the unpolarized sets. It is clear from Fig. 3 that the



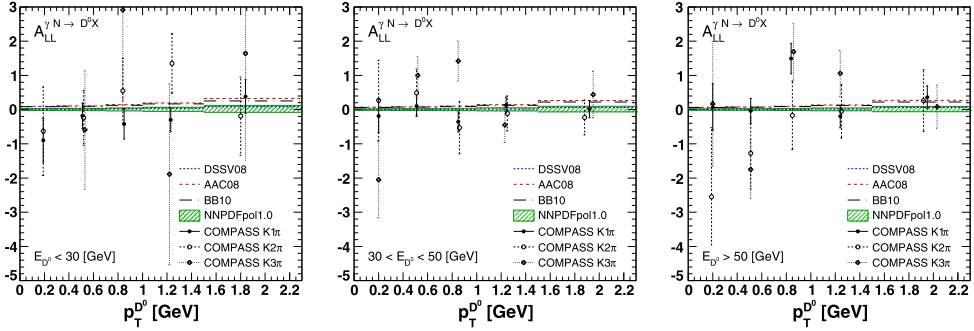


Fig. 3. Double-spin asymmetry for  $D^0$  meson photoproduction  $A_{LL}^{\gamma N \rightarrow D^0 X}$ , Eq. (7), as measured by COMPASS [20] from the three different decay channels, compared to the corresponding LO theoretical prediction obtained using NNPDFpol1.0 DSSV08, AAC08 and BB10 polarized parton sets, supplemented in each case by a suitable unpolarized set (see text). Results are presented for three bins of the  $D^0$  meson energy,  $E_{D^0}$ , and in five bins of its transverse momentum,  $p_T^{D^0}$ .

Table 4

Quality of the fit to combined and individual COMPASS open-charm data sets obtained using the polarized PDF sets (and unpolarized counterparts) of Fig. 3, as well as a set (denoted as reweighted) obtained starting with NNPDFpol1.0 and including the COMPASS data by reweighting. In each case, we show the number of data points, the effective number of replicas after reweighting for the reweighted set, and the  $\chi^2$  per data point obtained using each set.

Experiment	Set	$N_{\text{dat}}$	$N_{\text{eff}}$	$\chi^2/N_{\text{dat}}$				
				NNPDFpol1.0	Rewighted	DSSV08	AAC08	BB10
COMPASS		45	980	1.23	1.23	1.23	1.27	1.25
	COMPASS $K 1\pi$	15	990	1.27	1.27	1.27	1.43	1.38
	COMPASS $K 2\pi$	15	990	0.51	0.51	0.51	0.56	0.55
	COMPASS $K 3\pi$	15	970	1.90	1.89	1.90	1.81	1.82

COMPASS data have very large uncertainties on the scale of the gluon PDF uncertainty, despite the fact that the PDF uncertainty is very large, especially for the NNPDFpol1.0 set [1].

The COMPASS data have been included into the NNPDFpol1.0 set by Bayesian reweighting. The values of  $\chi^2$  per data point before and after reweighting are shown in Table 4, along with those obtained using other PDF sets in which the COMPASS data are not included; in each case the unpolarized set is the same as that used in Fig. 3. We also list the number of data points, and the effective number of replicas after reweighting a prior of  $N_{\text{rep}} = 1000$  replicas. Note that, because information on the correlation of systematics is unavailable, statistical and systematic uncertainties are added in quadrature when computing the  $\chi^2$ .

It is clear from Table 4 that the inclusion of the COMPASS data has a negligible impact on the fit, as shown by the fact that the  $\chi^2$  values before and after reweighting are either the same or extremely close, and the effective number of replicas after reweighting is always very close to the number of replicas in the prior set. Also, despite very significant differences in the shape of the central gluon distribution for the various PDF sets considered here, in each case the  $\chi^2$  values are essentially the same for all sets. This means that the  $\chi^2$  is mostly determined by the mutual consistency or inconsistency of the data themselves, rather than by the actual shape of the gluon.

The observable  $A_{LL}^{\gamma N \rightarrow D^0 X}$ , Eq. (7), and the polarized gluon PDF  $x\Delta g(x, Q_0^2)$  at  $Q_0^2 = 1 \text{ GeV}^2$  before and after reweighting, are shown in Fig. 4 and Fig. 5 respectively: as expected,

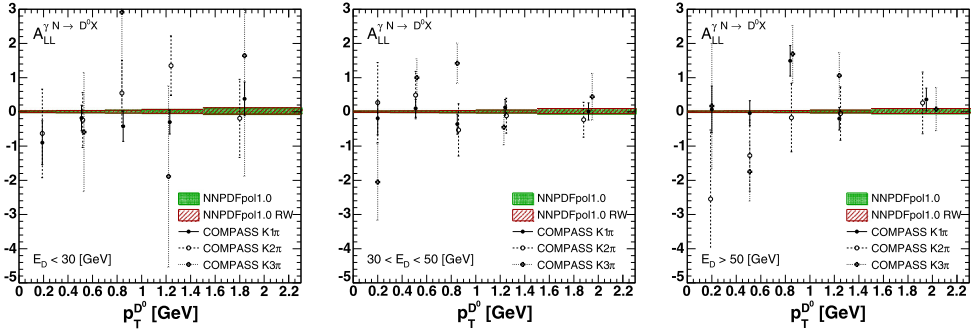


Fig. 4. The double-spin asymmetry  $A_{LL}^{\gamma N \rightarrow D^0 X}$  determined before and after reweighting compared to the COMPASS data.

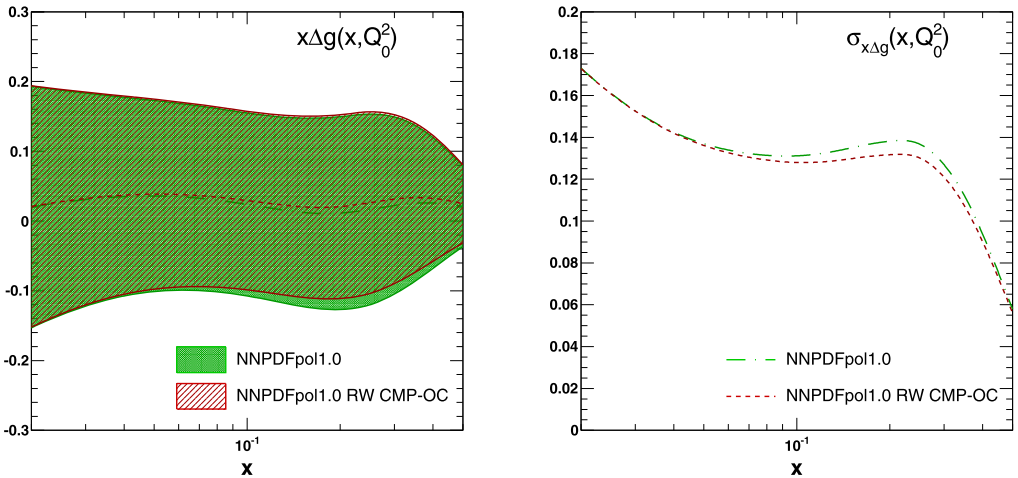


Fig. 5. The polarized gluon  $x\Delta g(x, Q_0^2)$  at  $Q_0^2 = 1 \text{ GeV}^2$  before and after reweighting with the COMPASS data (left); the absolute uncertainty on the gluon is also shown (right).

they are essentially unaffected by the inclusion of the COMPASS data. Note in particular that the uncertainty on the gluon, also shown in Fig. 5, is essentially unchanged. We conclude that the COMPASS data have little or no effect on polarized PDFs, and specifically the polarized gluon PDF.

### 3.2. High- $p_T$ jet production at STAR and PHENIX

We now turn to inclusive jet production in longitudinally polarized proton–proton collisions for which RHIC data are available (see e.g. [72,73]). These data are expected to have a significant impact on the gluon PDF because of the dominance of  $gg$  and  $qg$  initiated subprocesses in the accessible kinematic range (see e.g. [74,75]). Semi-inclusive hadron production in polarized collisions [21–26,76] is also sensitive to the gluon PDF, but it requires knowledge of fragmentation functions, which should be consistently determined along with parton distributions. Predictions for some semi-inclusive processes will be provided in Section 5.3 below with

the goal of assessing their potential relevance, but we do not include them in our PDF determination in order not to have to rely on poorly known fragmentation functions.

We consider specifically data for the longitudinal double-spin asymmetry,

$$A_{LL}^{1\text{-jet}} = \frac{\sigma^{++} - \sigma^{+-}}{\sigma^{++} + \sigma^{+-}}, \quad (8)$$

defined as the ratio of the difference to the sum of inclusive jet cross-sections with equal ( $\sigma^{++}$ ) or opposite ( $\sigma^{+-}$ ) proton beam polarizations. For dijet production, the leading-order parton kinematics is

$$x_1 = \frac{p_T}{\sqrt{s}}(e^{\eta_3} + e^{\eta_4}), \quad x_2 = \frac{p_T}{\sqrt{s}}(e^{-\eta_3} + e^{-\eta_4}), \quad (9)$$

where  $p_T$  is the transverse jet momentum,  $\eta_{3,4}$  are the rapidities of the two jets and  $\sqrt{s}$  is the center-of-mass energy. In single-inclusive jet production, the underlying Born kinematics is not fixed uniquely because the second jet is being integrated over (in Fig. 1 we have conventionally assumed equal longitudinal momentum of the incoming partons, so  $\eta_3 = -\eta_4 \equiv \eta$  and  $x_{1,2} = \frac{2p_T}{\sqrt{s}}e^{\pm\eta}$ ).

The NLO QCD computation for inclusive high- $p_T$  jet production in polarized hadron collisions was first presented in Ref. [77], based on the subtraction method of Refs. [78,79], along with a code for parton-level event generation. More recently, spin-dependent and spin-averaged cross-sections for single-inclusive high- $p_T$  jet production have been determined in Ref. [80] using the so-called *narrow-cone* approximation, which holds in the limit of not too large jet radius  $R$ . In this approximation, analytical results for the corresponding NLO partonic cross-sections can be derived, leading to a faster and more efficient computer code, as all singularities arising in the intermediate steps of the calculation explicitly cancel. The narrow-cone approximation was shown to be close to the result of Ref. [77] in the particular case of RHIC kinematics [80]. We will therefore use the code of Ref. [80], rather than that of Ref. [77]. The approach of Ref. [80] has been recently extended in Ref. [81] to  $k_T$ -type jet algorithms, used for the latest STAR data.

The two general purpose experiments at RHIC, STAR and PHENIX, have presented measurements of the longitudinal double-spin asymmetry for inclusive jet production, Eq. (8). Results from STAR are available for the 2005 and 2006 runs and, since very recently, also for the 2009 run. On the other hand, a single data set is available from PHENIX, corresponding to data taken in 2005, while further jet measurements from this experiment are not foreseen due to detector limitations in angular coverage. In the following, these data sets will be referred to as STAR 1j-05, STAR 1j-06, STAR 1j-09A, STAR 1j-09B and PHENIX 1j respectively.

The features of these data sets are summarized in Table 5 and the corresponding Born-level kinematic coverage is shown in Fig. 1. The experimental covariance matrix is available only for the STAR 1j-09A and STAR 1j-09B data sets: we included it in our analysis through a routine provided by the STAR Collaboration [82], which takes into account also additional fully correlated systematics arising from relative luminosity and jet energy scale uncertainties. For the other data sets, systematic and statistical uncertainties are added in quadrature. For the STAR 1j-05 and STAR 1j-06 data sets, we have to account for the fact that the data are taken in bins of  $p_T$ , whereas the corresponding theoretical predictions are computed for the center of each bin. We estimate the corresponding uncertainty as the maximal variation of the observable within each bin and take that value as a further uncorrelated systematic uncertainty. Furthermore, although these

Table 5

Features of the RHIC inclusive jet data included in the present analysis: the number of data points,  $N_{\text{dat}}$ , the algorithm used for jet reconstruction and the value of the jet radius,  $R$ , the range over which the rapidity  $\eta$  is integrated, the center-of-mass energy of the collisions,  $\sqrt{s}$ , and the integrated luminosity for each run,  $\mathcal{L}$ .

Data set	$N_{\text{dat}}$	jet-algorithm	$R$	$[\eta_{\text{min}}, \eta_{\text{max}}]$	$\sqrt{s}$ [GeV]	$\mathcal{L}$ [pb $^{-1}$ ]	Ref.
STAR 1j-05	10	midpoint-cone	0.4	[+0.20, +0.80]	200	2.1	[29]
STAR 1j-06	9	midpoint-cone	0.7	[−0.70, +0.90]	200	5.5	[29]
STAR 1j-09A	11	anti- $k_T$	0.6	[−0.50, +0.50]	200	25	[30]
STAR 1j-09B	11	anti- $k_T$	0.6	[−1.00, −0.50] [+0.50, +1.00]	200	25	[30]
PHENIX 1j	6	seeded-cone	0.3	[−0.35, +0.35]	200	2.1	[28]

Table 6

The value of the  $\chi^2$  per data point  $\chi^2/N_{\text{dat}}$  ( $\chi_{\text{rw}}^2/N_{\text{dat}}$ ) before (after) reweighting the prior samples with the RHIC inclusive jet data. Results are presented for both separate and total data sets and for each prior discussed in Section 2.

Experiment	Set	$N_{\text{dat}}$	$\chi^2/N_{\text{dat}}$				$\chi_{\text{rw}}^2/N_{\text{dat}}$			
			$1\sigma$	$2\sigma$	$3\sigma$	$4\sigma$	$1\sigma$	$2\sigma$	$3\sigma$	$4\sigma$
STAR		41	1.50	1.49	1.50	1.50	1.05	1.04	1.04	1.04
	STAR 1j-05	10	1.04	1.05	1.04	1.04	1.01	1.02	1.02	1.02
	STAR 1j-06	9	0.75	0.76	0.76	0.76	0.59	0.58	0.59	0.59
	STAR 1j-09A	11	1.40	1.39	1.39	1.40	0.98	0.99	0.98	0.98
	STAR 1j-09B	11	3.04	3.05	3.03	3.05	1.18	1.17	1.17	1.18
PHENIX	PHENIX 1j	6	0.24	0.24	0.24	0.24	0.24	0.24	0.24	0.24
		47	1.35	1.35	1.35	1.36	1.00	1.01	1.01	1.00

data are provided with asymmetric systematic uncertainties, we symmetrize them, according to Eqs. (7)–(8) of Ref. [2].

For each data set in Table 5, we have computed the longitudinal double-spin asymmetry Eq. (8), at NLO, using the narrow-cone approximation code of Refs. [80,81], suitably modified in order to use the NNPDF polarized parton sets via the LHAPDF interface [83,84]. In each case we use the jet algorithm and cone radius which are appropriate for the given data set, as listed in Table 5. Polarized and unpolarized PDFs are taken respectively from the prior ensembles constructed in Section 2 and from the NNPDF2.3 NLO reference parton set. As for open-charm muoproduction, the numerator in Eq. (8) is computed for each replica in the prior ensembles ( $N_{\text{rep}} = 1000$ ), while the denominator is evaluated only once for the central unpolarized replica. This is justified because uncertainties of the polarized PDF completely dominate over those of the unpolarized ones.

We have included the jet data of Table 5 by reweighting the prior sets discussed in Section 2. The  $\chi^2$  per data point before and after reweighting are listed in Table 6, for each data set and a combined set including all these data, for each of the four prior sets. Various measures of the effectiveness of the reweighting process are listed in Table 7: the effective number of replicas  $N_{\text{eff}}$ , and the modal value of the  $\mathcal{P}(\alpha)$  distribution,  $\langle \alpha \rangle$ , defined by Eq. (12) of Ref. [49]. The parameter  $\alpha$  measures the consistency of the data which are used for reweighting with those included in the prior set:  $\alpha$  is the factor by which the uncertainty on the new data must be rescaled in order for both sets to be consistent with each other. A value of  $\alpha$  close to one means that the uncertainties have been correctly estimated.

Table 7

The effective number of replicas after reweighting,  $N_{\text{eff}}$  (the starting sample has  $N_{\text{rep}} = 1000$ ), and the modal value of the  $\mathcal{P}(\alpha)$  distribution,  $\langle\alpha\rangle$  (note that here and henceforth  $\langle\alpha\rangle$  denotes the mode, not the mean of the  $P(\alpha)$  distribution). Results are quoted for separate and combined data sets and for each of the different prior PDF ensembles.

Experiment	Set	$N_{\text{dat}}$	$N_{\text{eff}}$				$\langle\alpha\rangle$			
			$1\sigma$	$2\sigma$	$3\sigma$	$4\sigma$	$1\sigma$	$2\sigma$	$3\sigma$	$4\sigma$
STAR		41	256	256	256	260	1.01	1.10	1.02	1.01
	STAR 1j-05	10	931	931	931	931	1.03	1.02	1.05	1.06
	STAR 1j-06	9	621	623	622	621	0.97	0.98	0.95	0.96
	STAR 1j-09A	11	244	244	244	244	1.20	1.20	1.15	1.17
	STAR 1j-09B	11	299	299	299	300	1.10	1.05	1.05	1.10
PHENIX	PHENIX 1j	6	740	740	740	741	0.55	0.50	0.50	0.55
		47	334	338	334	333	1.00	1.02	0.99	0.99

Inspection of Tables 6–7 allows us to draw the following conclusions:

- Results are essentially independent of the choice of prior. This was to be expected, given the very mild sensitivity of this observable to the polarized flavor–antiflavor decomposition.
- The effective number of replicas after reweighting is always significantly lower than the size of the initial sample  $N_{\text{rep}} = 1000$ , suggesting that the RHIC data do have a significant impact on the fit. The most constraining data sets, for which  $N_{\text{eff}}$  is smallest, are STAR 1j-09A and STAR 1j-09B: this is to be expected since these are the measurements with smallest statistical and systematic uncertainties.
- The effective number of replicas after reweighting is nevertheless always rather larger than a hundred, which is a typical size needed for the final replica sample to provide accurate results. This means that the size of the prior sample is large enough for final results to be reliable.
- The modal value of the  $\mathcal{P}(\alpha)$  distribution for all the STAR data as well as for the global data set is always close to one, suggesting correct uncertainty estimation (even for the earlier data sets for which no information on correlated systematics is available). However, the modal value of  $\alpha$  is significantly below one for the PHENIX data, suggesting that for this data set uncertainties are overestimated, possibly due to the missing information on correlated systematics.
- The  $\chi^2$  after reweighting is of order one for the data sets for which information on correlated systematics is available, and it shows a significant improvement, with an especially remarkable agreement for the STAR 1j-09B which, as mentioned, has the smallest uncertainties. This suggests that these data are bringing in significant new information.

Predictions for the asymmetry  $A_{LL}^{1\text{-jet}}$  Eq. (8) obtained using the NNPDFp011.0 PDF set before and after reweighting are compared to the RHIC data in Fig. 6. The curves shown correspond to the  $1\sigma$  prior PDF ensemble, reweighted with the complete data set of Table 5. The improvement in experimental uncertainties in STAR 1j-09A and STAR 1j-09B as compared to all other data sets is clearly visible, and it leads to an equally visible reduction of the uncertainty on the theoretical prediction, as well as a significant change of its central value.

This is mostly due to a corresponding reduction in uncertainty and change of shape of the polarized gluon upon reweighting: this can be seen in Fig. 7, where the polarized gluon distributions

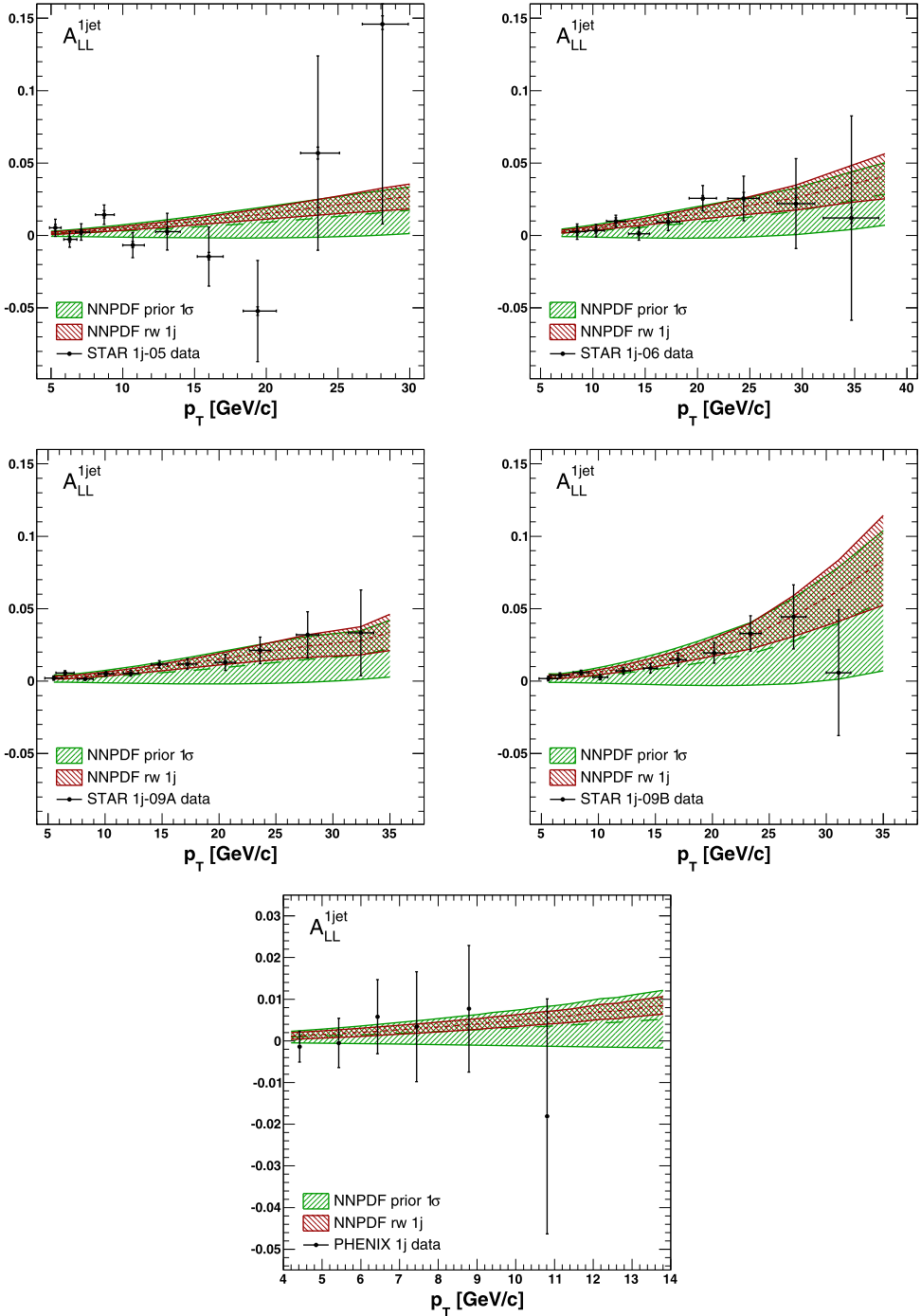


Fig. 6. Predictions for the double longitudinal spin asymmetry for single-inclusive jet production  $A_{LL}^{1jet}$ , Eq. (8), as a function of the jet  $p_T$ , before and after reweighting, compared to the corresponding RHIC data. Note the different scale of vertical axis for the PHENIX data as compared to the STAR data.

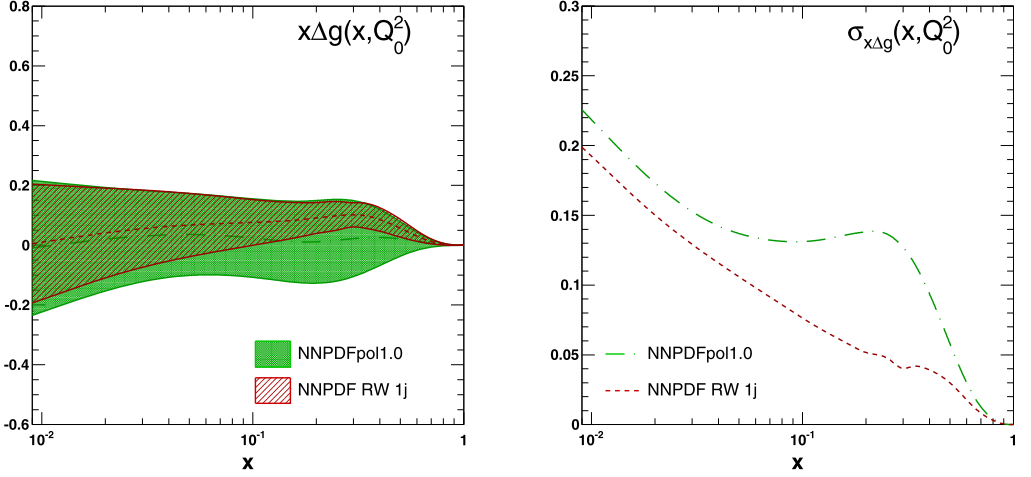


Fig. 7. The NNPDF<sub>pol1.0</sub> polarized gluon distribution  $x\Delta g(x, Q_0^2)$  at  $Q_0^2 = 1 \text{ GeV}^2$  before and after reweighting with the RHIC jet data, labeled NNPDF RW 1j (left); the absolute uncertainty on the gluon is also shown (right).

before and after reweighting are compared. Here too the result corresponds to the  $1\sigma$  prior PDF ensemble, reweighted with the complete data set of Table 5. We have explicitly checked that the reweighted gluon is independent of the choice of prior. In the kinematic range probed at RHIC,  $x \in [0.04, 0.2]$  (see Fig. 1), the polarized gluon PDF tends to become positive and its uncertainty is reduced by more than a factor two for  $x \sim 0.3$ . This feature is qualitatively consistent with that reported by the DSSV group in Refs. [85,86], based on the analysis of the same data. The other PDFs are essentially unaffected by the inclusion of the RHIC jet data.

#### 4. Polarized quark–antiquark separation: the $W$ asymmetry

We now turn to polarized hadron collider data that constrain the flavor separation of polarized quarks and antiquarks. We consider in particular  $W$  production. Because of the chiral nature of the weak interactions, the polarized parton content may be accessed both through (parity violating) single-spin and (parity conserving) double-spin asymmetries. The single-spin asymmetry is defined as

$$A_L \equiv \frac{\sigma^+ - \sigma^-}{\sigma^+ + \sigma^-} = \frac{\Delta\sigma}{\sigma}, \quad (10)$$

where  $\sigma^{+(-)}$  denotes the cross-section for  $W$  production when colliding positive (negative) longitudinally polarized protons off unpolarized protons, and the double-spin asymmetry

$$A_{LL} \equiv \frac{\sigma^{++} - \sigma^{+-}}{\sigma^{++} + \sigma^{+-}}, \quad (11)$$

where  $\sigma^{++}$  ( $\sigma^{+-}$ ) is the cross-section for  $W$  production with equal (opposite) proton beam polarizations.

At leading order, neglecting Cabibbo-suppressed channels, the former is given by

$$A_L^{W^+} \approx \frac{\Delta u(x_1)\bar{d}(x_2) - \Delta\bar{d}(x_1)u(x_2)}{u(x_1)\bar{d}(x_2) + \bar{d}(x_1)u(x_2)}, \quad (12)$$

and the latter is

$$A_{LL}^{W^+} \approx -\frac{\Delta u(x_1)\Delta\bar{d}(x_2) + \Delta\bar{d}(x_1)\Delta u(x_2)}{u(x_1)\bar{d}(x_2) + \bar{d}(x_1)u(x_2)}, \quad (13)$$

where for fixed  $W$  rapidity  $y_W$ , the momentum fractions  $x_{1,2}$  carried by the colliding partons are given by

$$x_{1,2} = \frac{M_W}{\sqrt{s}} e^{\pm y_W}. \quad (14)$$

It is thus clear that first, each of these asymmetries is sensitive to the flavor decomposition of polarized quark and antiquark distributions and second, that their simultaneous measurement provides an especially strong constraint, as both linear and quadratic combinations of polarized PDFs are being measured. This last fact has the interesting implication that the single- and double-spin asymmetries (10)–(11) must satisfy a nontrivial positivity bound, derived in Ref. [87]

$$1 \pm A_{LL}(y_W) > |A_L(y_W) \pm A_L(-y_W)|, \quad (15)$$

where  $y_W$  is the  $W$  boson rapidity (not to be confused with  $\eta_l$ , the pseudo-rapidity of the lepton from the  $W$  decay, which is used for the experimental measurements).

Both the STAR and PHENIX Collaborations have presented measurements of the parity-violating spin asymmetry  $A_L^{W^\pm}$ , Eq. (10), based on the RHIC 2009 run at  $\sqrt{s} = 500$  GeV [31,32]. These measurements are based on data sets with low integrated luminosities:  $\mathcal{L} = 12$  pb $^{-1}$  and  $\mathcal{L} = 8.6$  pb $^{-1}$  respectively for STAR and PHENIX. Each of these two experiments provides a determination of the asymmetry for a single value of the rapidity and for outgoing  $W^\pm$ , but as discussed in Ref. [88] they are affected by very large uncertainties, do not provide any significant constraint, and we need not discuss them further.

Also, the STAR Collaboration has recently presented [33] results for both the  $A_L^{W^\pm}$  and  $A_{LL}^{W^\pm}$  asymmetries, based on  $\mathcal{L} = 9$  pb $^{-1}$  of  $\sqrt{s} = 500$  GeV data from the 2010 and  $\mathcal{L} = 77$  pb $^{-1}$  of data at  $\sqrt{s} = 510$  GeV from the 2011 run. These data have been combined into a single data set at the nominal energy of  $\sqrt{s} = 510$  GeV, and have greatly reduced uncertainties as compared to previous measurements.

The STAR data are provided for both  $W^+$  and  $W^-$  final states, which we will refer to as STAR- $A_L^{W^\pm}$  and STAR- $A_{LL}^{W^\pm}$ , presented in bins of  $\eta_l$ , the rapidity of the lepton from the  $W$  decay. In particular the single-spin asymmetry data are given in six rapidity bins, and the double-spin asymmetry in three rapidity bins, integrated over the lepton transverse momentum in the range  $25 < p_T^l < 50$  GeV. Correlated beam polarization systematics (3.4% and 6.5% respectively for single- and double-spin asymmetries) are provided, while an additional uncorrelated systematics, due to the relative luminosity, affects  $A_L^{W^\pm}$  [33]. Using LO kinematics we see that these STAR  $W$  production data constrain light quark and antiquark PDFs with  $0.05 \lesssim x \lesssim 0.4$  and  $Q^2 \sim M_W^2$ , see Fig. 1.

We have included the STAR single- and double-asymmetry data by reweighting. The asymmetries have been computed using the CHE code [89], suitably modified to handle NNPDF parton sets via the LHAPDF interface. The values of  $\chi^2$  per data point before and after reweighting (and the number of data points) are collected in Table 8, while the measures of the reweighting process (defined as for Table 7 of Section 3) are given in Table 9. In each case, results are given both for separate and combined STAR data sets, and for all four priors constructed in Section 2.



Table 8

The value of the  $\chi^2$  per data point  $\chi^2/N_{\text{dat}}$  ( $\chi_{\text{rw}}^2/N_{\text{dat}}$ ) before (after) reweighting, for the different priors. The number of data points  $N_{\text{dat}}$  in each of the two STAR data sets is also given.

Experiment	Set	$N_{\text{dat}}$	$\chi^2/N_{\text{dat}}$				$\chi_{\text{rw}}^2/N_{\text{dat}}$			
			$1\sigma$	$2\sigma$	$3\sigma$	$4\sigma$	$1\sigma$	$2\sigma$	$3\sigma$	$4\sigma$
STAR- $A_L$		12	1.38	1.44	1.39	1.33	1.08	0.88	0.74	0.74
	STAR- $A_L^{W^+}$	6	0.75	0.75	0.86	0.90	0.75	0.75	0.68	0.70
	STAR- $A_L^{W^-}$	6	1.92	2.03	1.82	1.67	1.32	1.08	0.83	0.82
STAR- $A_{LL}$		6	0.82	0.81	0.78	0.78	0.82	0.80	0.76	0.76
	STAR- $A_{LL}^{W^+}$	3	0.92	0.88	0.81	0.80	0.90	0.85	0.77	0.76
	STAR- $A_{LL}^{W^-}$	3	0.73	0.74	0.75	0.76	0.73	0.74	0.75	0.76
		18	1.19	1.20	1.15	1.15	1.00	0.87	0.78	0.77

Table 9

The effective number of replicas after reweighting,  $N_{\text{eff}}$ , and the modal value of the  $\mathcal{P}(\alpha)$  distribution, for the different priors.

Experiment	Set	$N_{\text{dat}}$	$N_{\text{eff}}$				$\langle\alpha\rangle$			
			$1\sigma$	$2\sigma$	$3\sigma$	$4\sigma$	$1\sigma$	$2\sigma$	$3\sigma$	$4\sigma$
STAR- $A_L$		12	697	475	441	497	1.27	1.25	1.10	1.10
	STAR- $A_L^{W^+}$	6	905	852	723	709	0.92	0.95	0.98	0.98
	STAR- $A_L^{W^-}$	6	529	444	492	516	1.45	1.43	1.20	1.20
STAR- $A_{LL}$		6	993	983	968	963	0.95	0.95	0.95	0.95
	STAR- $A_{LL}^{W^+}$	3	972	953	933	924	0.98	0.97	0.95	0.97
	STAR- $A_{LL}^{W^-}$	3	999	995	994	990	0.85	0.93	0.95	0.90
		18	717	569	508	501	1.11	1.10	1.15	1.11

Inspection of Tables 8–9 allows us to draw the following conclusions:

- Independence of the choice of prior is achieved between the  $3\sigma$  and  $4\sigma$  priors, both in terms of fit quality (Table 8) and of reweighting parameters (Table 9). We will thus discuss henceforth results obtained using the  $4\sigma$  prior.
- The effective number of replicas after reweighting is always significantly lower than the size of the prior sample for the single-spin asymmetry, but quite close to it for the double-spin asymmetry, thus suggesting that the former data have a significant impact, while the latter do not. This is consistent with the fact that statistical uncertainties are smaller for single-spin asymmetries.
- The effective number of replicas, however, even for the double-asymmetry data remains rather larger than that found in Section 3.2 after reweighting with the STAR jet data: hence, as the impact of the reweighting is more moderate, the number of replicas after reweighting remains adequate for accurate phenomenology.
- The modal value of the  $\mathcal{P}(\alpha)$  distribution are generally close to one, though a bit higher for  $A_L^{W^-}$  ( $\langle\alpha\rangle \sim 1.2$ ), suggesting a mild underestimation of uncertainties, while the effective number of replicas after reweighting with these data is smaller than that found when reweighting with the  $A_L^{W^+}$  data, which should have a very similar impact.

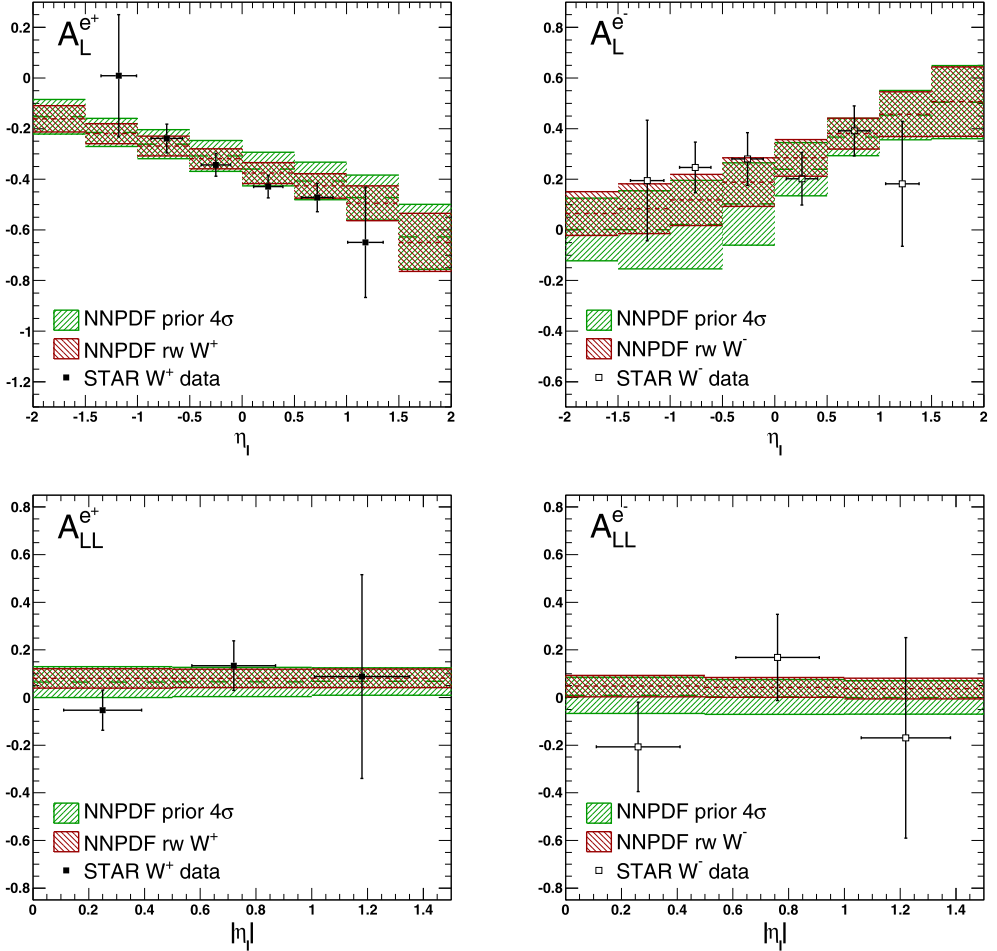


Fig. 8. Predictions for the longitudinal positron (left plots) and electron (right plots) single-spin asymmetries  $A_L^{e^+}$  and  $A_L^{e^-}$  (top row) and double-spin asymmetries  $A_{LL}^{e^+}$  and  $A_{LL}^{e^-}$  (bottom row) before and after reweighting, compared to the STAR data from the 2012 run [33]. Results obtained with the  $4\sigma$  prior are shown. The uncertainties shown are statistical only.

- After reweighting, the  $\chi^2$  per data point is below one (in fact, a little more than one  $\sigma$  below one), thus showing perfect agreement of the reweighted prediction with the data; interestingly the  $\chi^2$  before reweighting for the  $A_L^{e^-}$  set was much greater than one, thus showing that the prior (based on DSSV, but with inflated uncertainties) does not agree well with these data.

Predictions for the longitudinal positron (electron) single- and double-spin asymmetries  $A_L^{e^+}$ ,  $A_{LL}^{e^+}$  ( $A_L^{e^-}$ ,  $A_{LL}^{e^-}$ ) before and after reweighting are shown in Fig. 8 as a function of the lepton rapidity  $\eta_l$ , and compared to the STAR data [33]. As usual, the numerator in Eqs. (10)–(11) is computed for each replica in the PDF ensemble ( $N_{\text{rep}} = 1000$ ), while the denominator is evaluated only for the central unpolarized replica. The reduction in uncertainty after reweighting

is clearly visible in all plots, and so is the good agreement between the data and the prediction after reweighting, the improvement being especially clear in the single-spin asymmetry plots.

The impact of the STAR data on the PDFs is seen in Fig. 9, where we compare the up and down antiquark polarized PDFs before and after reweighting; the corresponding absolute uncertainties are explicitly shown in the right plots. The change in shape in comparison to the prior (whose shape was determined by that of the DSSV best-fit PDFs) is especially remarkable for the  $\Delta\bar{u}$  distribution, and indicates that the STAR  $W$  data pull in a different direction for  $\Delta\bar{u}$  than the SIDIS data used in DSSV08. The reduction in uncertainty is very visible in the region of the peak, where it amounts to almost 20%.

## 5. The NNPDFp011.1 polarized parton set

In this section, we combine the different pieces of information obtained in Sections 3 and 4 and produce a global polarized PDF set based on the NNPDF methodology, NNPDFp011.1, which is the main result of this paper. This set is constructed by simultaneous reweighting of the prior PDF samples with all the new data from the COMPASS, STAR and PHENIX experiments listed in Table 1 (upper part), displayed in Fig. 1, and discussed in the previous sections.

In summary, the NNPDFp011.1 set represents the state-of-the-art in our understanding of the proton's spin content from polarized observables which do not entail parton-to-hadron fragmentation. Further constraints will be provided by a variety of semi-inclusive measurements, which in turn will require the development of a set of parton fragmentation functions using the NNPDF methodology. In the long term, the final word on the spin content of the proton will require brand new facilities such as an Electron–Ion Collider, which could bring polarized PDF determinations to a similar level of accuracy as their unpolarized counterparts.

Based on our results, we will also reassess the status of polarized quark and gluon first moments, and, as an example of application, we will compute the longitudinal double-spin asymmetry for single-inclusive particle production in proton–proton collisions, and compare results to recent RHIC data.

### 5.1. Simultaneous reweighting of RHIC and COMPASS data

The NNPDFp011.1 parton set is obtained by performing a global reweighting of the prior polarized PDF ensembles (constructed as described in Section 2) with all the relevant data from the COMPASS, STAR and PHENIX experiments. The reweighting parameters and the  $\chi^2$  per data point before and after reweighting are collected in Table 10, to be compared to Tables 6–7 of Section 3 and Tables 8–9 of Section 4, where the results of reweighting with individual data sets were shown.

Inspection of Table 10 allows us to draw the following conclusions:

- Independence of the prior is achieved already with the  $3\sigma$  prior, possibly even between the  $2\sigma$  and the  $3\sigma$  prior. The fact that independence of the prior is achieved somewhat earlier when performing the global reweighting than when reweighting with the most constraining data set (see Tables 8–9) is to be expected, as a consequence of the fact that the global data set carries more information.
- The effective number of replicas after reweighting is by almost one order of magnitude smaller than the size of the starting data set, thereby showing that the data have a significant impact on the fit.

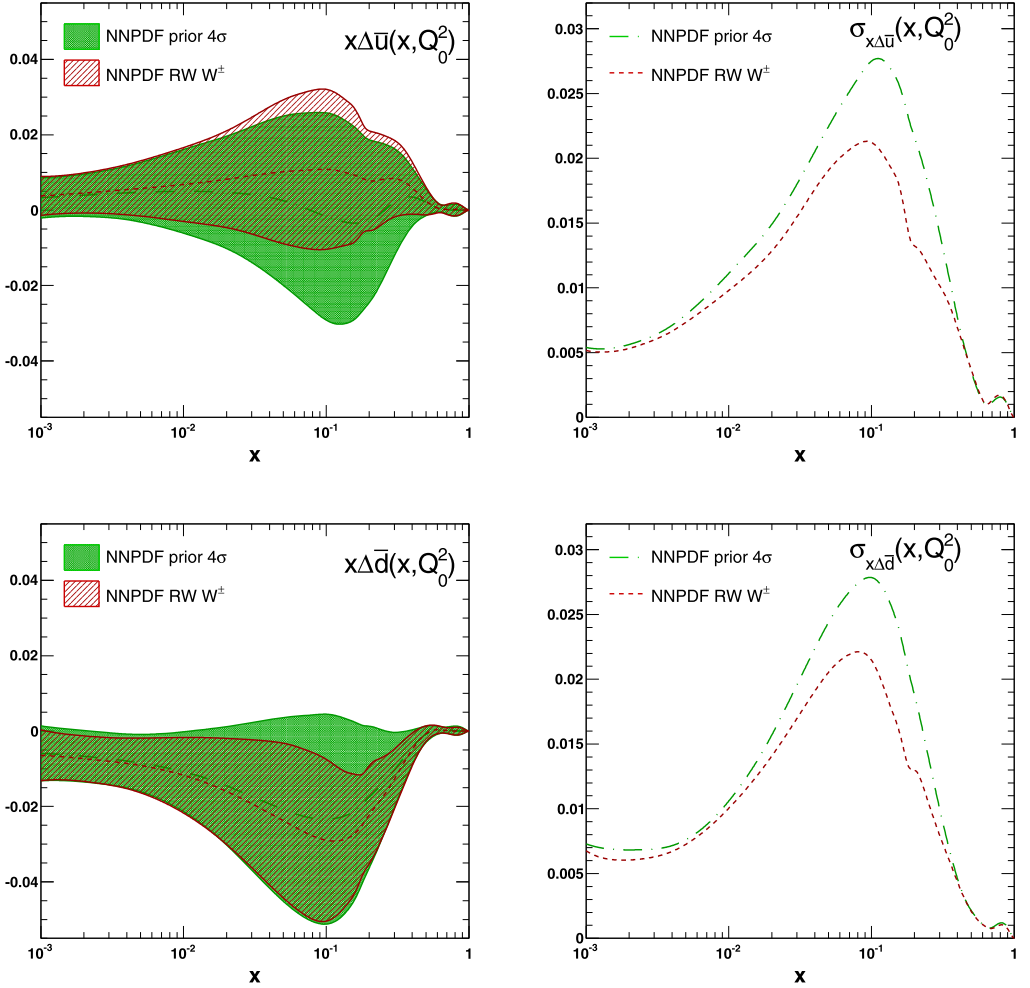


Fig. 9. Comparison between the polarized antiquark sea densities  $x\Delta\bar{u}$  (upper plots) and  $x\Delta\bar{d}$  (lower plots) before and after reweighting the  $4\sigma$  ensemble with complete STAR  $W^\pm$  data set, at  $Q_0^2 = 1 \text{ GeV}^2$ . The absolute PDF uncertainty is also shown (right plots).

- The effective number of replicas after reweighting remains nevertheless larger than the conventional threshold  $N_{\text{rep}} = 100$  which is necessary to ensure reliability of the results obtained using the reweighted set.
- The modal value of the  $\mathcal{P}(\alpha)$  distribution is just above one, suggesting good compatibility of the new global data set with the previous fit, even though, as discussed in Section 4, Table 9, there is some evidence for moderate uncertainty underestimation for the  $A_L^{W^-}$  STAR data, which may explain the value slightly above one.
- The  $\chi^2$  per data point improves substantially upon reweighting, and is of order one after reweighting. Note however that these values should be treated with care because information on correlated systematics is not available for all experiments, and thus for all experiments for which statistical and systematic uncertainties must be added in quadrature the  $\chi^2$  per data point might be expected to be less than one.

Table 10

The value of the  $\chi^2$  per data point  $\chi^2/N_{\text{dat}}$  ( $\chi_{\text{rw}}^2/N_{\text{dat}}$ ) before (after) reweighting, the effective number of replicas after reweighting  $N_{\text{eff}}$ , and the modal value of the  $\mathcal{P}(\alpha)$  distribution. All results refer to reweighting with all the  $N_{\text{dat}} = 110$  data points corresponding to the data sets shown in Table 1 (upper part), with different choices for the prior.

	$1\sigma$	$2\sigma$	$3\sigma$	$4\sigma$
$\chi^2/N_{\text{dat}}$	1.30	1.31	1.30	1.30
$\chi_{\text{rw}}^2/N_{\text{dat}}$	1.09	1.07	1.05	1.05
$N_{\text{eff}}$	181	164	142	135
$\langle\alpha\rangle$	1.22	1.21	1.23	1.22

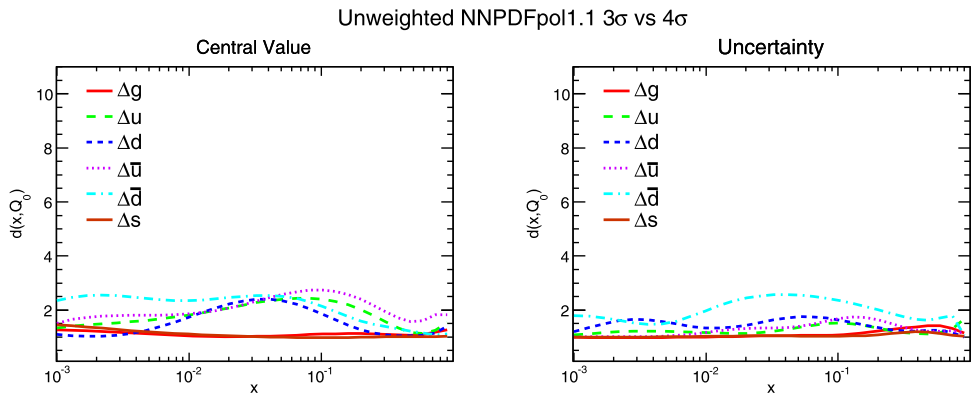


Fig. 10. Distances between parton sets obtained via global reweighting of  $3\sigma$  and  $4\sigma$  prior PDF ensembles at  $Q^2 = 10 \text{ GeV}^2$ .

In the sequel, we will choose the PDF set obtained by reweighting the  $4\sigma$  prior as our default NNPDFpol1.1 PDF set: this guarantees maximal independence of the choice of prior, without significant loss of accuracy, given that the values of  $N_{\text{eff}}$  for the  $3\sigma$  and  $4\sigma$  prior are roughly equal. As a further check of independence of the prior, in Fig. 10 we display the distance  $d(x, Q^2)$ , as defined in Appendix A of Ref. [5], between PDFs obtained with the  $3\sigma$  and  $4\sigma$  priors, at  $Q^2 = 10 \text{ GeV}^2$ . This statistical estimator is expected to take the value  $d \sim 1$  when two samples of  $N_{\text{rep}}$  replicas are extracted from the same underlying probability distribution, while it is  $d = \sqrt{N_{\text{rep}}}$  when the two samples are extracted from two distributions which differ on average by one standard deviation. As we see in Fig. 10, the two ensembles are indeed statistically equivalent.

In order to construct the NNPDFpol1.1 set, made of  $N_{\text{rep}} = 100$  replicas, the results from the global reweighting discussed above are unweighted following the procedure in Ref. [50]. We show in Table 11 the  $\chi^2$  of each of the experiments included in the NNPDFpol1.1 analysis computed with the NNPDFpol1.1 PDF set; for experiments which were already included in NNPDFpol1.0 the value obtained using that set is also shown. It is clear that the description of inclusive DIS data in NNPDFpol1.1 is as good as in the original set, and that the description of each individual set used for reweighting is as good in the combined reweighted set as it was when reweighting with each individual set (as shown in Tables 6 and 8 of Sections 3 and 4 respectively).

Table 11

The  $\chi^2$  per data point of all the experiments included in the NNPDF<sub>po11.1</sub> analysis computed with the corresponding PDF set; for experiments which were already included in NNPDF<sub>po11.0</sub> the value obtained using that set is also shown.

Experiment	NNPDF <sub>po11.0</sub>	NNPDF <sub>po11.1</sub>
EMC	0.44	0.43
SMC	0.93	0.90
SMC <sub>lowx</sub>	0.97	0.97
E143	0.64	0.67
E154	0.40	0.45
E155	0.89	0.85
COMPASS-D	0.65	0.70
COMPASS-P	1.31	1.38
HERMES97	0.34	0.34
HERMES	0.79	0.82
COMPASS (OC)	–	1.22
STAR (jets)	–	1.05
PHENIX (jets)	–	0.24
STAR- $A_L$	–	0.72
STAR- $ALL$	–	0.75

The comparison of predictions obtained using NNPDF<sub>po11.1</sub> to the data is accordingly very similar to that shown in Figs. 4, 6, 8, and is therefore not shown.

We now turn to the NNPDF<sub>po11.1</sub> PDFs which we first compare in Fig. 11 with those of the previous set, NNPDF<sub>po11.0</sub>. Because in NNPDF<sub>po11.0</sub> only the PDF combinations  $\Delta u^+$ ,  $\Delta d^+$ ,  $\Delta s^+$ , and the gluon PDF,  $\Delta g$ , are determined, only these can be compared directly. The comparison is shown at  $Q^2 = 10 \text{ GeV}^2$ , where the positivity bound coming from the unpolarized NNPDF2.3 set is also displayed. In order to quantitatively assess the impact of the new data for these PDF combinations, in Fig. 12 we also plot the distance  $d(x, Q)$  between NNPDF<sub>po11.0</sub> and NNPDF<sub>po11.1</sub> at  $Q^2 = 10 \text{ GeV}^2$ .

The main differences are found for  $\Delta g$  in the region  $0.02 \lesssim x \lesssim 0.5$  covered by the STAR inclusive jet production data. In this region, the gluon from NNPDF<sub>po11.1</sub> parton set is positive at the one-sigma level, and it has a much reduced uncertainty in comparison to its NNPDF<sub>po11.0</sub> counterpart. At small  $x$ , outside the kinematical coverage of the STAR jet data, the NNPDF<sub>po11.0</sub> and NNPDF<sub>po11.1</sub> are again statistically equivalent. On the other hand, the total quark–antiquark combinations  $\Delta u^+$ ,  $\Delta d^+$  and  $\Delta s^+$  are only moderately affected by the new data, and they are still mostly constrained by the polarized inclusive DIS data, with the RHIC data only leading to a minor reduction in uncertainty mostly for the  $\Delta d^+$  distribution and in the small  $x$  region.

We now turn to the individual NNPDF<sub>po11.1</sub> PDFs, which are compared in Fig. 13 to those from the global DSSV08 fit [6] at  $Q^2 = 10 \text{ GeV}^2$ ; PDF uncertainties are nominal one-sigma error bands for NNPDF<sub>po11.1</sub>, and Hessian uncertainties ( $\Delta\chi^2 = 1$ ) for DSSV08. The main conclusions from the comparison in Fig. 13 are the following:

- The  $\Delta u$  and  $\Delta d$  PDFs of the NNPDF<sub>po11.1</sub> and DSSV08 are qualitatively similar, though for NNPDF<sub>po11.1</sub> uncertainties are typically larger. Note however that the default  $\Delta\chi^2 = 1$  adopted by DSSV08 may lead to uncertainty underestimation: it is well known that in Hessian global unpolarized fits a tolerance  $\Delta\chi^2 = T$  with  $T > 1$  is needed for faithful

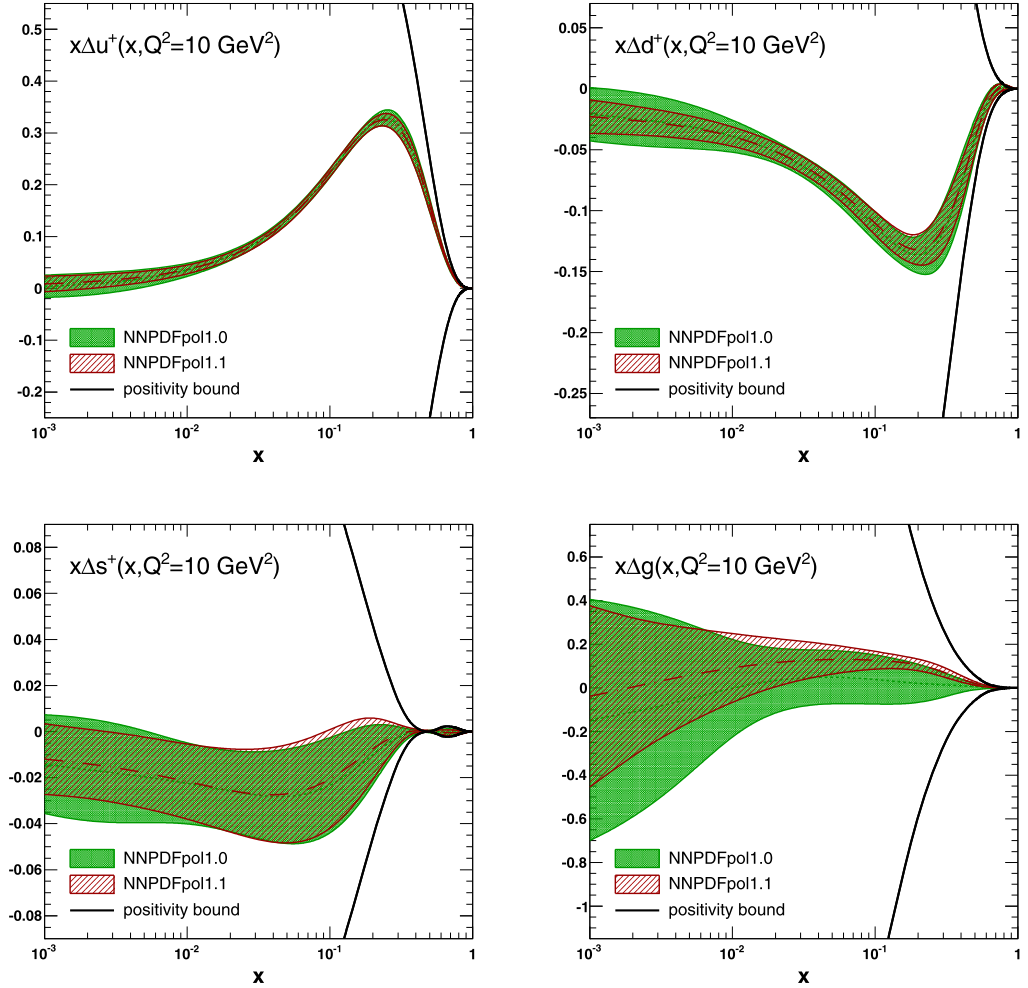


Fig. 11. Comparison between NNPDF<sub>pol1.0</sub> and NNPDF<sub>pol1.1</sub> parton sets at  $Q^2 = 10 \text{ GeV}^2$  for the  $x\Delta u^+$ ,  $x\Delta d^+$ ,  $x\Delta s^+$  and  $x\Delta g$  polarized PDFs. The positivity bound from the unpolarized NNPDF2.3 set is also shown.

uncertainty estimation, as also recognized in Ref. [6] (see also the discussion of first moments in Section 5.2 below).

- The NNPDF<sub>pol1.1</sub> polarized gluon PDF is consistent at the one-sigma level with its DSSV08 counterpart in the large- $x$  region  $x \gtrsim 0.2$ , where they have similar uncertainties. However, for  $x < 0.2$ ,  $\Delta g$  has a node in the DSSV08 determination, while it is clearly positive in NNPDF<sub>pol1.1</sub>. This result is mostly driven by the recent precise inclusive jet production data from STAR (STAR 1j-09A and STAR 1j-09B of Section 3), which were not available at the time of the original DSSV08 analysis shown in Fig. 13 (only STAR 1j-05 and STAR 1j-06 were included). Recent updates of the DSSV08 fit including also STAR 1j-09A and STAR 1j-09B data sets [85,86] suggest a positive  $\Delta g$  consistent with the NNPDF<sub>pol1.1</sub> result.
- As already noticed in Section 4, inclusion of the  $W^\pm$  data (not included in the DSSV08 fit) visibly affects the shape of the  $\Delta \bar{u}$  distribution, especially above  $x \sim 3 \cdot 10^{-2}$ , which

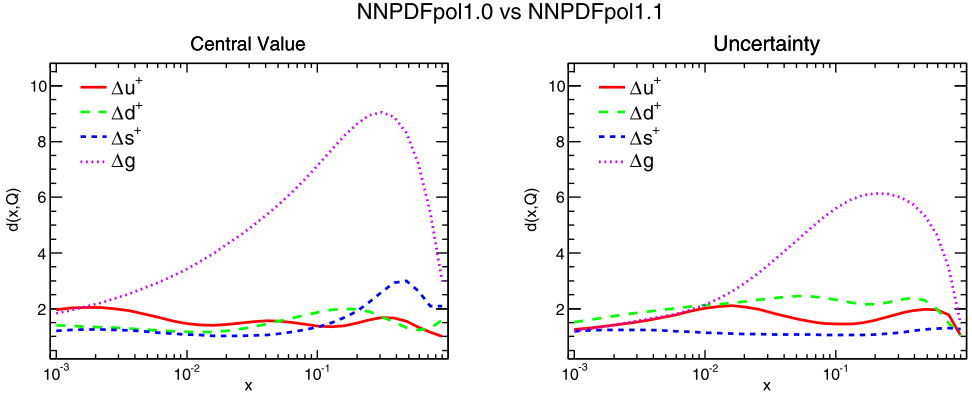


Fig. 12. Distances  $d(x, Q^2)$  between NNPDF<sub>pol1.0</sub> and NNPDF<sub>pol1.1</sub> at  $Q^2 = 10 \text{ GeV}^2$ .

thus differs from that in the DSSV08 set. This might be a signal of tension between  $W^\pm$  and semi-inclusive DIS data, possibly due to limited knowledge of fragmentation functions. A similar change in the shape of  $\Delta\bar{u}$ , making its peak less negative, was also found in a preliminary global fit including STAR data in the DSSV framework [90].

- Since  $W$  boson production data in the kinematic regime probed by STAR are not sensitive to strangeness, the discrepancy between NNPDF and DSSV determinations of  $\Delta s$ , already found in Ref. [1], is still present. As discussed in our previous work [1], in the NNPDF analysis the polarized strange PDF is obtained from inclusive DIS data through its  $Q^2$  evolution and assumptions about flavor symmetry of the proton sea enforced by experimentally measured baryon octet decay constants. On the other hand, the DSSV determination of polarized PDFs also includes semi-inclusive data with identified kaons in final states, which are directly sensitive to strangeness, but subject to the uncertainty in the kaon fragmentation functions (which is difficult to quantify).

In view of the comparison with other PDF sets, it is interesting to examine positivity bounds which, as discussed in Section 2, our PDFs satisfy by construction for each individual flavor, as is also clear from Fig. 13. Note that positivity also implies that single- and double-spin asymmetries must satisfy the bounds Eq. (15) [87]. Results for the combinations of asymmetries which are bounded to be non-negative are shown in Fig. 14, determined using NNPDF<sub>pol1.1</sub> and DSSV08 PDFs (with, as in Section 4 the NNPDF2.3 [71] and MRST02 [91] respectively as corresponding unpolarized sets), for the RHIC center-of-mass energy  $\sqrt{s} = 510 \text{ GeV}$ , integrated over the lepton transverse momentum in the range  $25 < p_T < 50$ , for  $W^-$  and  $W^+$  production.

Interestingly, while NNPDF<sub>pol1.1</sub> satisfies the bounds (as it must to by construction) DSSV08 PDFs appear to violate both the bounds for  $W^-$  production when  $|y_W| \gtrsim 1.6$ , which roughly corresponds to momentum fractions  $x \gtrsim 0.8$ . This ties in neatly with the observation that while NNPDF<sub>pol1.1</sub> and DSSV08 results are in good agreement for  $W^+$  production, they differ for  $W^-$ , reflecting the different shape of the  $\Delta\bar{u}$  distribution, also seen in Fig. 13.

### 5.2. The spin content of the proton revisited

The first moments of the polarized PDFs can be directly related to the fraction of the proton spin carried by individual partons. In Ref. [1] we presented a detailed analysis of the first moments of various PDF combinations from the fit to polarized inclusive DIS data only. In this



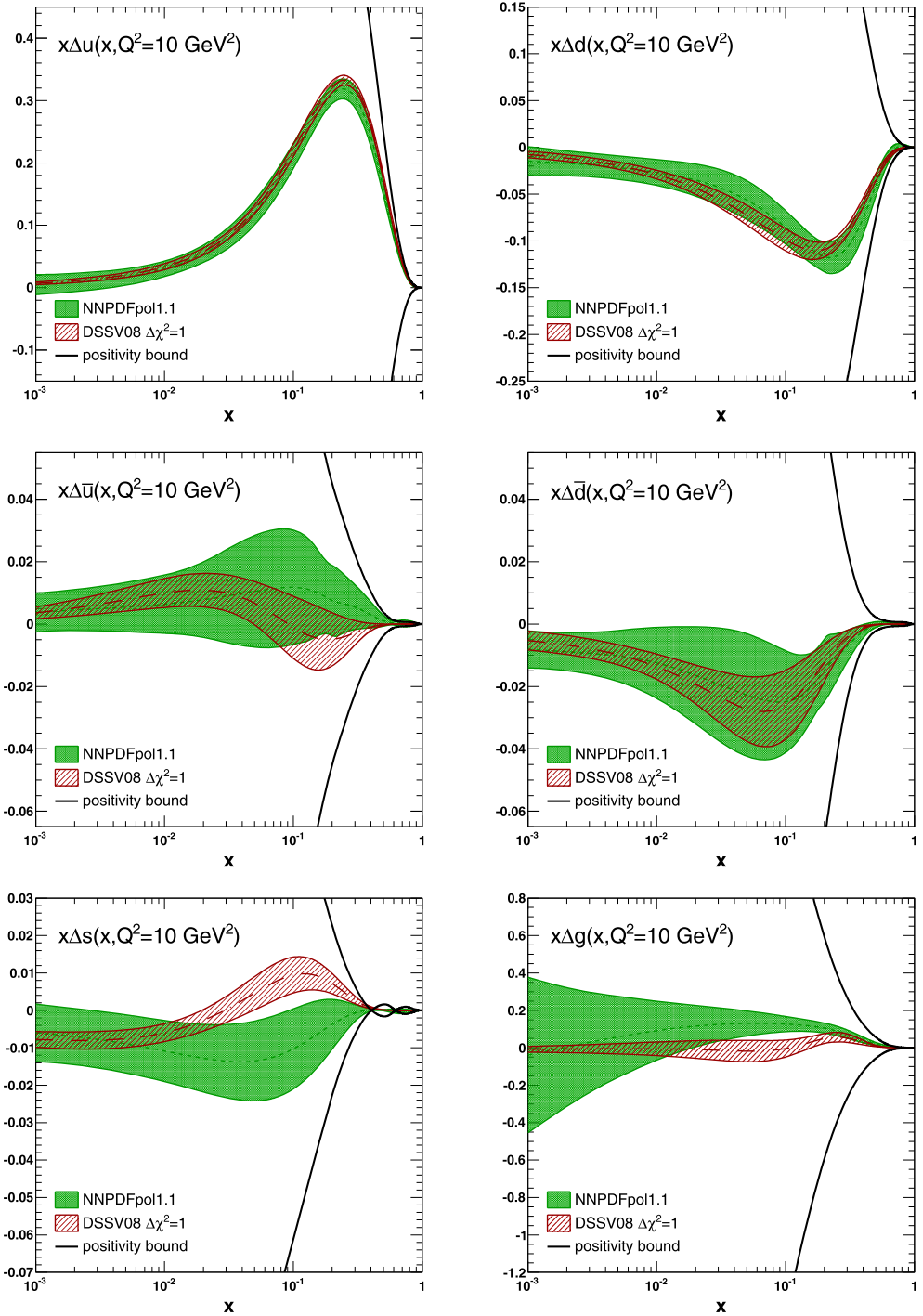


Fig. 13. The NNPDFpol1.1 parton set compared to DSSV08 [6] at  $Q^2 = 10 \text{ GeV}^2$ .

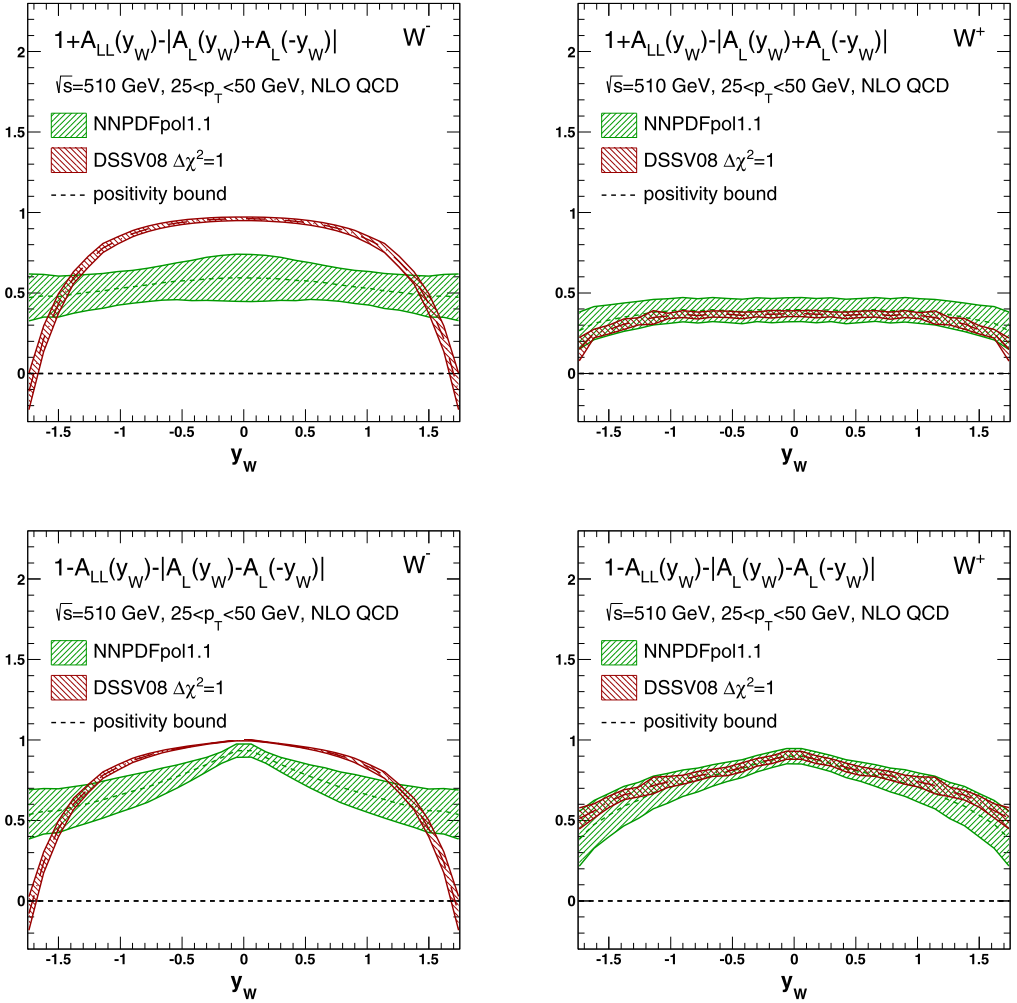


Fig. 14. The combinations  $1 \pm A_{LL}(y_W) - |A_L(y_W) \pm A_L(-y_W)|$  of longitudinal single- and double-spin asymmetries,  $A_{LL}$  and  $A_L$ , for  $W$  production in polarized proton–proton collisions plotted as a function of the  $W$  boson rapidity  $y_W$ . Results are shown for both the NNPDFpol1.1 and DSSV08 parton sets. The positivity bounds (15) are satisfied whenever the curves are positive.

section we revisit this analysis with the NNPDFpol1.1 parton set, and quantify the impact of the RHIC data on the proton spin content. We define the (truncated) first moments of the polarized PDFs  $\Delta f(Q^2, x)$  in the region  $[x_{\min}, x_{\max}]$

$$\langle \Delta f(Q^2) \rangle^{[x_{\min}, x_{\max}]} \equiv \int_{x_{\min}}^{x_{\max}} dx \Delta f(x, Q^2). \quad (16)$$

We will provide results for both full moments,  $\langle \Delta f(Q^2) \rangle^{[0,1]}$ , and truncated moments, restricted to the  $x$  region which roughly corresponds to the kinematic coverage of experimental data, i.e.  $\langle \Delta f(Q^2) \rangle^{[10^{-3}, 1]}$ , see Fig. 1.

Table 12

Full and truncated first moments of the polarized quark distributions, Eq. (16), at  $Q^2 = 10 \text{ GeV}^2$ , for NNPDFp011.1, NNPDFp011.0 (when available) and DSSV08. The uncertainties shown are one-sigma for NNPDF and Lagrange multiplier with  $\Delta\chi^2/\chi^2 = 2\%$  for DSSV. The number in parenthesis for DSSV08 is the contribution that should be added to the truncated moment in order to obtain the full moment.

$\Delta f$	$\langle \Delta f(Q^2) \rangle^{[0,1]}$		$\langle \Delta f(Q^2) \rangle^{[10^{-3},1]}$		
	NNPDFp011.0	NNPDFp011.1	NNPDFp011.0	NNPDFp011.1	DSSV08
$\Delta u^+$	$+0.77 \pm 0.10$	$+0.79 \pm 0.07$	$+0.76 \pm 0.06$	$+0.76 \pm 0.04$	$+0.793^{+0.028}_{-0.034} (+0.020)$
$\Delta d^+$	$-0.46 \pm 0.10$	$-0.47 \pm 0.07$	$-0.41 \pm 0.06$	$-0.41 \pm 0.04$	$-0.416^{+0.035}_{-0.025} (-0.042)$
$\Delta \bar{u}$	–	$+0.06 \pm 0.06$	–	$+0.04 \pm 0.05$	$+0.028^{+0.059}_{-0.059} (+0.008)$
$\Delta \bar{d}$	–	$-0.11 \pm 0.06$	–	$-0.09 \pm 0.05$	$-0.089^{+0.090}_{-0.080} (-0.026)$
$\Delta s$	$-0.07 \pm 0.06$	$-0.07 \pm 0.05$	$-0.06 \pm 0.04$	$-0.05 \pm 0.04$	$-0.006^{+0.028}_{-0.031} (-0.051)$
$\Delta \Sigma$	$+0.16 \pm 0.30$	$+0.18 \pm 0.21$	$+0.23 \pm 0.15$	$+0.25 \pm 0.10$	$+0.366^{+0.042}_{-0.062} (+0.124)$

We consider first polarized quark and antiquark PDFs. In contrast to the previous NNPDFp011.0 determination, we can now determine the first moment of individual light flavor and antiflavours. In Table 12 we show results for the  $C$ -even combinations  $\Delta u^+$  and  $\Delta d^+$ , for the light antiquarks  $\Delta \bar{u}$  and  $\Delta \bar{d}$ , for the polarized strangeness  $\Delta s$  (recall that we assume  $\Delta s = \Delta \bar{s}$ ), and for the singlet PDF combination  $\Delta \Sigma = \sum_{q=u,d,s} \Delta q^+$ . The corresponding central values and one-sigma PDF uncertainties obtained from the  $N_{\text{rep}} = 100$  replicas of the NNPDFp011.1 parton set at  $Q^2 = 10 \text{ GeV}^2$  are collected in Table 12. We compare our results to both NNPDFp011.0 and DSSV08. In the latter case, we quote the PDF uncertainty obtained from the Lagrange multiplier method with  $\Delta\chi^2/\chi^2 = 2\%$ , which is recommended in Ref. [6] as more reliable. This corresponds to a value of the tolerance parameter, defined above in Section 5.1, of order  $T \sim 8$ , hence to uncertainties rather larger than those shown for DSSV PDFs in Fig. 13. For DSSV08 we also show in parenthesis the contribution that must be added to the truncated moment given in the table in order to obtain the first moment: this contribution comes from extrapolation in the unmeasured region, and it should be assigned 100% uncertainty [6].

The first moments obtained from NNPDFp011.1 and NNPDFp011.0 are perfectly consistent with each other, as expected based on the agreement of the PDFs seen in Fig. 11. The new constraints on polarized quark PDFs from the RHIC  $W$  data lead to a substantial reduction, by almost a factor two, of the PDF uncertainties on the first moments of  $\Delta u^+$  and  $\Delta d^+$ . The contribution to the uncertainty coming from the data and extrapolation regions are of comparable size in NNPDFp011.1, just as in NNPDFp011.0. This means that the uncertainty due to the extrapolation has also decreased by almost a factor two in NNPDFp011.1 in comparison to NNPDFp011.0, despite the fact that the kinematic coverage of the data at small  $x$  is not significantly extended by the hadronic data (recall Fig. 1): this is because the lower uncertainty in the data region also limits the spread of acceptable small- $x$  extrapolations. Interestingly, all quark truncated first moments are also compatible between NNPDFp011.1 and DSSV08, despite the differences in shape in the  $\Delta \bar{u}$  PDF.

We now turn to the gluon. Results for full and truncated moments at  $Q^2 = 10 \text{ GeV}^2$  are presented in Table 13, where, on top of NNPDFp011.1, NNPDFp011.0, and DSSV08 results, we also show the result found with a recent update of the DSSV family, DSSV++ [85]. The interest of this comparison lies in the fact that DSSV++ also includes jet production data from RHIC, so the data which constrain the gluon are essentially the same as in NNPDFp011.1 (the exception

Table 13

Same as Table 12 but for the gluon. Results are also shown for the truncated moment in the range of the RHIC jet data and for DSSV++.

	$\langle \Delta g(Q^2) \rangle^{[0,1]}$	$\langle \Delta g(Q^2) \rangle^{[10^{-3},1]}$	$\langle \Delta g(Q^2) \rangle^{[0.05,0.2]}$
NNPDF <sub>pol1.0</sub>	$-0.95 \pm 3.87$	$-0.06 \pm 1.12$	$+0.05 \pm 0.15$
NNPDF <sub>pol1.1</sub>	$+0.03 \pm 3.24$	$+0.49 \pm 0.75$	$+0.17 \pm 0.06$
DSSV08	–	$0.013^{+0.702}_{-0.314} (+0.097)$	$0.005^{+0.129}_{-0.164}$
DSSV++	–	–	$0.10^{+0.06}_{-0.07}$

being pion production data from RHIC, included in DSSV++ but not in NNPDF<sub>pol1.1</sub>). We show full and truncated moments in the measured region  $[10^{-3}, 1]$ , and also truncated moments in the region  $x \in [0.05, 0.2]$ , which corresponds to the range covered by the RHIC inclusive jet data (see Fig. 1).

The significant improvement in the uncertainties on the first moment of  $\Delta g$  in the data region when going from NNPDF<sub>pol1.0</sub> to NNPDF<sub>pol1.1</sub> is apparent; clearly, the improvement is concentrated in the region of the RHIC jet data. The uncertainty from the extrapolation, however, does not improve significantly and it dominates the result for the full first moment, which remains essentially undetermined. This strongly suggests that only with a wider kinematic coverage, such as would be obtained at a polarized Electron–Ion Collider, could a significantly more accurate determination of the polarized gluon first moment be achieved [39,40]. A moderate improvement might also be possible from RHIC jet data at higher center-of-mass energy, up to  $\sqrt{s} = 500$  GeV. Interestingly, the contribution to the first moment from the region of the RHIC jet data is clearly positive. The result found in this region is in good agreement (both in terms of central value and uncertainty) with that of DSSV++ [85]. This provides some evidence for a positive polarized gluon, though unfortunately the large uncertainty due to the extrapolation does not allow one to draw firm conclusions about the full first moment.

### 5.3. Single-particle inclusive production asymmetries at RHIC

As an illustration of the predictive power of the NNPDF<sub>pol1.1</sub> set, in this section we compute the longitudinal double-spin asymmetry for single-hadron production in polarized proton–proton collisions, which has also been measured at RHIC recently. As for inclusive jets, Eq. (8), this asymmetry is defined as

$$A_{LL}^h = \frac{\sigma^{++} - \sigma^{+-}}{\sigma^{++} + \sigma^{+-}} \equiv \frac{d\Delta\sigma}{d\sigma}, \quad (17)$$

where  $\sigma^{++}$  ( $\sigma^{+-}$ ) is the cross-section for the process with equal (opposite) proton polarizations. Theoretical predictions for the polarized and unpolarized differential cross-sections  $d\Delta\sigma$  and  $d\sigma$  may be obtained using factorized expression with the general structure

$$\begin{aligned} d\sigma &= \sum_{a,b,c=q,\bar{q},g} f_a \otimes f_b \otimes D_c^H \otimes d\hat{\sigma}_{ab}^c, \\ d\Delta\sigma &= \sum_{a,b,c=q,\bar{q},g} \Delta f_a \otimes \Delta f_b \otimes D_c^H \otimes d\Delta\hat{\sigma}_{ab}^c, \end{aligned} \quad (18)$$

where the sum runs over all initial- and final-state partonic channels,  $d\Delta\hat{\sigma}_{ab}^c$ ,  $d\hat{\sigma}_{ab}^c$  are respectively the polarized and unpolarized cross-sections for the partonic subprocess

$$a + b \rightarrow (c \rightarrow H) + X, \quad (19)$$

and  $D_c^H$  is the fragmentation function for parton  $c$  into hadron  $H$ .

Predictions for these processes are subject to a significant uncertainty due to lack of knowledge of the relevant fragmentation functions: therefore, we did not use them for PDF determination in order not to introduce a theoretical uncertainty over which we have poor control. On the other hand, the relevant partonic cross-sections are known up to NLO, and may be obtained using the code of Ref. [92].

In Figs. 15–16 we show predictions for several of these double-spin asymmetries, compared to RHIC data, obtained using the NNPDF<sub>pol1.1</sub> parton set. We use the fragmentation functions from the DSS07 set [45]. All uncertainties shown are PDF uncertainties from the  $N_{\text{rep}} = 100$  polarized replica set (with, as usual, the unpolarized denominator computed using the central PDF, as its uncertainty is negligible). Of course, an extra unknown uncertainty from the fragmentation function should also be included, but it cannot be reliably estimated at present.

We specifically compare to PHENIX data for neutral pion production at  $\sqrt{s} = 200$  GeV [25, 27] and  $\sqrt{s} = 62.4$  GeV [24], and mid-rapidity ( $|\eta| < 0.35$ ) charged hadron production at  $\sqrt{s} = 62.4$  GeV [76], and to STAR data for neutral-pion production with forward rapidity ( $0.8 < \eta < 2.0$ ) at  $\sqrt{s} = 200$  GeV [26]. Earlier PHENIX data for neutral pion production [21–23], with significantly larger uncertainties, are not considered.

Our predictions are always in good agreement with the data within experimental uncertainties; they suggest that double-spin asymmetries for single-hadron production remain quite small in all the available  $p_T$  range, typically below the 1% level. Our predictions for negatively charged pion asymmetry is also small for all transverse momenta, see Fig. 16. In contrast,  $A_{LL}^{\pi^+}$  is larger than  $A_{LL}^{\pi^0}$ . High- $p_T$  data (both polarized and unpolarized) are potentially sensitive to the gluon distribution, hence these data might eventually provide a further handle on the polarized gluon, if sufficiently accurate fragmentation functions become available.

## 6. Conclusions and outlook

We have presented a first global polarized PDF determination based on NNPDF methodology, which includes, on top of the deep-inelastic scattering data already used in our previous NNPDF<sub>pol1.0</sub> polarized PDF set, COMPASS charm production data and all relevant inclusive hadronic data from polarized collisions at RHIC, i.e. essentially all available data which do not require knowledge of light-quark fragmentation functions. We have thus achieved a significant improvement in accuracy in the determination of the gluon distribution in the medium and small- $x$  region (from jet data), with evidence for a positive gluon polarization in this region, and a determination of individual light quark and antiquark PDFs (from  $W^\pm$  productions data). Together with the available NNPDF unpolarized PDF sets (currently NNPDF2.3 [71]) this provides a first global set of polarized and unpolarized PDFs determined with a consistent methodology, including mutually consistent constraints from cross-section positivity. This provides a reliable framework for phenomenological applications, also including possible searches for new physics with polarized beams [93].

Future inclusive RHIC data (specifically from the PHENIX and STAR Collaborations) with improved accuracy and kinematic coverage will lead to even more precise polarized PDF determinations. In addition, a significant potential for improvement lies in the use of semi-inclusive data, whose availability and accuracy is constantly increasing both from fixed-target [11–15]

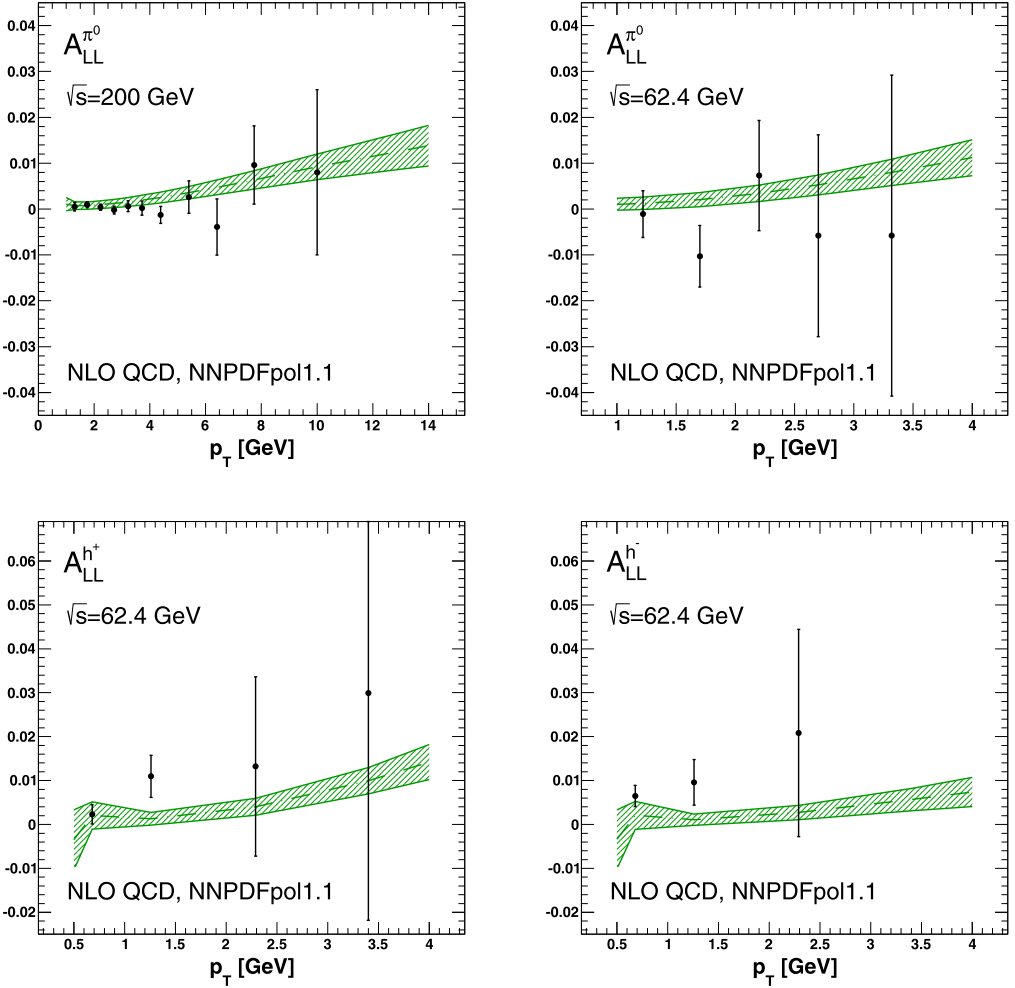


Fig. 15. NLO predictions for neutral-pion (upper plots) and positively and negatively charged hadrons (lower plots) spin asymmetries using NNPDFpol1.1 with the corresponding PDF uncertainty band, compared to available data from the PHENIX experiment presented in Refs. [24,25,27,76].

and RHIC collider [24–26] data. However, a consistent inclusion of these data in our framework requires a simultaneous determination of fragmentation functions using NNPDF methodology.

However, present and future data from existing facilities are unlikely to substantially improve our knowledge of polarized first moments, i.e. of the proton spin structure. Indeed, the accuracy of present determinations of polarized first moments is already limited by the uncertainties due to extrapolation into the unmeasured small- $x$  region. This is especially true for the polarized gluon, for which we have now tantalizing evidence of a positive polarization in the measured region, which is however completely swamped by an uncertainty from the extrapolation region which is larger by one order of magnitude. Improved accuracy requires a widening of the kinematic coverage: this could be achieved at a polarized Electron–Ion Collider [34–36], which would probe polarized PDFs down to much smaller values of  $x$ , as shown quantitatively in Refs. [39,40],

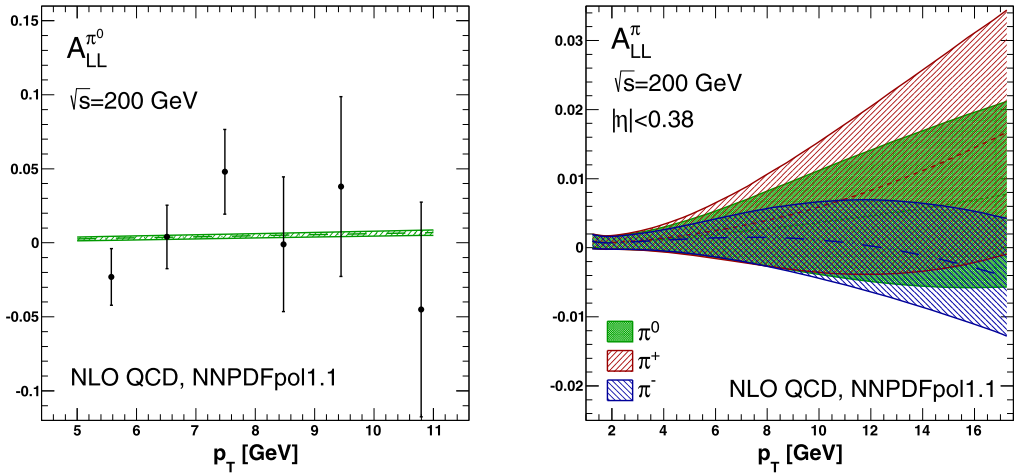


Fig. 16. (Left panel) Predictions for the neutral-pion spin asymmetry compared to data measured by STAR [26]. (Right panel) Prediction for the neutral- and charged-pion spin asymmetries in the kinematic range accessed by upcoming PHENIX measurements.

and would also provide further constraints on flavor separation from polarized charged-current DIS [94].

The NNPDFpol1.1 polarized PDFs, with  $N_{\text{rep}} = 100$  replicas, are available from the NNPDF HepForge web site, <http://nnpdf.hepforge.org/> in a format compliant with the LHAPDF interface [83,84]. In addition, stand-alone Fortran77, C++ and Mathematica driver codes are also available there.

## Acknowledgements

We would like to thank M. Stratmann and D. de Florian for useful discussions and for providing us with code for jet and  $W$  boson production at polarized hadron colliders [80,89], and S. Frixione for providing us with the code of Ref. [77] for jets. We also would like to thank C. Aidala, C. Gagliardi, A. Vossemer and J. Stevens for useful information about the PHENIX and STAR data, and R. Fatemi for providing us with the STAR data used in this publication. This research was supported by an Italian Ministry of Education, Research and Universities PRIN2010 grant (S.F., E.R.N., G.R.), by a UK STFC grant ST/J000329/1 (R.D.B.), by a European Investment Bank EIBURS grant (S.F., E.R.N.), and by the European Commission through the HiggsTools Initial Training Network PITN-GA-21012-316704.

## References

- [1] NNPDF Collaboration, R.D. Ball, et al., Nucl. Phys. B 874 (2013) 36, arXiv:1303.7236.
- [2] NNPDF Collaboration, L. Del Debbio, et al., J. High Energy Phys. 0503 (2005) 080, arXiv:hep-ph/0501067.
- [3] NNPDF Collaboration, L. Del Debbio, et al., J. High Energy Phys. 0703 (2007) 039, arXiv:hep-ph/0701127.
- [4] NNPDF Collaboration, R.D. Ball, et al., Nucl. Phys. B 809 (2009) 1, arXiv:0808.1231.
- [5] NNPDF Collaboration, R.D. Ball, et al., Nucl. Phys. B 838 (2010) 136, arXiv:1002.4407.
- [6] D. de Florian, et al., Phys. Rev. D 80 (2009) 034030, arXiv:0904.3821.
- [7] E. Leader, A.V. Sidorov, D.B. Stamenov, Phys. Rev. D 82 (2010) 114018, arXiv:1010.0574.
- [8] Asymmetry Analysis Collaboration, M. Hirai, S. Kumano, Nucl. Phys. B 813 (2009) 106, arXiv:0808.0413.

- [9] J. Blumlein, H. Bottcher, Nucl. Phys. B 841 (2010) 205, arXiv:1005.3113.
- [10] P. Jimenez-Delgado, A. Accardi, W. Melnitchouk, Phys. Rev. D 89 (2014) 034025, arXiv:1310.3734.
- [11] HERMES Collaboration, K. Ackerstaff, et al., Phys. Lett. B 464 (1999) 123, arXiv:hep-ex/9906035.
- [12] Spin Muon Collaboration, B. Adeva, et al., Phys. Lett. B 420 (1998) 180, arXiv:hep-ex/9711008.
- [13] HERMES Collaboration, A. Airapetian, et al., Phys. Rev. D 71 (2005) 012003, arXiv:hep-ex/0407032.
- [14] COMPASS Collaboration, M. Alekseev, et al., Phys. Lett. B 660 (2008) 458, arXiv:0707.4077.
- [15] COMPASS Collaboration, M. Alekseev, et al., Eur. Phys. J. C 64 (2009) 171, arXiv:0907.0388.
- [16] HERMES Collaboration, A. Airapetian, et al., Phys. Rev. Lett. 84 (2000) 2584, arXiv:hep-ex/9907020.
- [17] HERMES Collaboration, A. Airapetian, et al., J. High Energy Phys. 1008 (2010) 130, arXiv:1002.3921.
- [18] Spin Muon Collaboration (SMC), B. Adeva, et al., Phys. Rev. D 70 (2004) 012002, arXiv:hep-ex/0402010.
- [19] COMPASS Collaboration, C. Adolph, et al., Phys. Lett. B 718 (2013) 922, arXiv:1202.4064.
- [20] COMPASS Collaboration, C. Adolph, et al., Phys. Rev. D 87 (2013) 052018, arXiv:1211.6849.
- [21] PHENIX Collaboration, S. Adler, et al., Phys. Rev. Lett. 93 (2004) 202002, arXiv:hep-ex/0404027.
- [22] PHENIX Collaboration, S. Adler, et al., Phys. Rev. D 73 (2006) 091102, arXiv:hep-ex/0602004.
- [23] PHENIX Collaboration, A. Adare, et al., Phys. Rev. D 76 (2007) 051106, arXiv:0704.3599.
- [24] PHENIX Collaboration, A. Adare, et al., Phys. Rev. D 79 (2009) 012003, arXiv:0810.0701.
- [25] PHENIX Collaboration, A. Adare, et al., Phys. Rev. Lett. 103 (2009) 012003, arXiv:0810.0694.
- [26] STAR Collaboration, L. Adamczyk, et al., Phys. Rev. D 89 (2014) 012001, arXiv:1309.1800.
- [27] PHENIX Collaboration, A. Adare, et al., arXiv:1402.6296, 2014.
- [28] PHENIX Collaboration, A. Adare, et al., Phys. Rev. D 84 (2011) 012006, arXiv:1009.4921.
- [29] STAR Collaboration, L. Adamczyk, et al., Phys. Rev. D 86 (2012) 032006, arXiv:1205.2735.
- [30] STAR Collaboration, L. Adamczyk, et al., arXiv:1405.5134, 2014.
- [31] STAR Collaboration, M. Aggarwal, et al., Phys. Rev. Lett. 106 (2011) 062002, arXiv:1009.0326.
- [32] PHENIX Collaboration, A. Adare, et al., Phys. Rev. Lett. 106 (2011) 062001, arXiv:1009.0505.
- [33] STAR Collaboration, L. Adamczyk, et al., arXiv:1404.6880, 2014.
- [34] A. Deshpande, et al., Annu. Rev. Nucl. Part. Sci. 55 (2005) 165, arXiv:hep-ph/0506148.
- [35] A. Accardi, et al., arXiv:1212.1701, 2012.
- [36] D. Boer, et al., arXiv:1108.1713, 2011.
- [37] M.L. Mangano, et al., arXiv:hep-ph/0105155, 2001.
- [38] S. Forte, M.L. Mangano, G. Ridolfi, Nucl. Phys. B 602 (2001) 585, arXiv:hep-ph/0101192.
- [39] E.C. Aschenauer, R. Sassot, M. Stratmann, Phys. Rev. D 86 (2012) 054020, arXiv:1206.6014.
- [40] NNPDF Collaboration, R.D. Ball, et al., Phys. Lett. B 728 (2013) 524, arXiv:1310.0461.
- [41] S. Kretzer, Phys. Rev. D 62 (2000) 054001, arXiv:hep-ph/0003177.
- [42] B.A. Kniehl, G. Kramer, B. Potter, Nucl. Phys. B 582 (2000) 514, arXiv:hep-ph/0010289.
- [43] L. Bourhis, et al., Eur. Phys. J. C 19 (2001) 89, arXiv:hep-ph/0009101.
- [44] M. Hirai, et al., Phys. Rev. D 75 (2007) 094009, arXiv:hep-ph/0702250.
- [45] D. de Florian, R. Sassot, M. Stratmann, Phys. Rev. D 75 (2007) 114010, arXiv:hep-ph/0703242.
- [46] S. Albino, B. Kniehl, G. Kramer, Nucl. Phys. B 803 (2008) 42, arXiv:0803.2768.
- [47] S. Albino, et al., arXiv:0804.2021, 2008.
- [48] D. de Florian, et al., Prog. Part. Nucl. Phys. 67 (2012) 251, arXiv:1112.0904.
- [49] NNPDF Collaboration, R.D. Ball, et al., Nucl. Phys. B 849 (2011) 112, arXiv:1012.0836.
- [50] NNPDF Collaboration, R.D. Ball, et al., Nucl. Phys. B 855 (2012) 608, arXiv:1108.1758.
- [51] European Muon Collaboration, J. Ashman, et al., Nucl. Phys. B 328 (1989) 1.
- [52] Spin Muon Collaboration, B. Adeva, et al., Phys. Rev. D 58 (1998) 112001.
- [53] Spin Muon Collaboration, B. Adeva, et al., Phys. Rev. D 60 (1999) 072004.
- [54] COMPASS Collaboration, V.Y. Alexakhin, et al., Phys. Lett. B 647 (2007) 8, arXiv:hep-ex/0609038.
- [55] COMPASS Collaboration, M. Alekseev, et al., Phys. Lett. B 690 (2010) 466, arXiv:1001.4654.
- [56] E143 Collaboration, K. Abe, et al., Phys. Rev. D 58 (1998) 112003, arXiv:hep-ph/9802357.
- [57] E154 Collaboration, K. Abe, et al., Phys. Rev. Lett. 79 (1997) 26, arXiv:hep-ex/9705012.
- [58] E155 Collaboration, P. Anthony, et al., Phys. Lett. B 493 (2000) 19, arXiv:hep-ph/0007248.
- [59] HERMES Collaboration, K. Ackerstaff, et al., Phys. Lett. B 404 (1997) 383, arXiv:hep-ex/9703005.
- [60] HERMES Collaboration, A. Airapetian, et al., Phys. Rev. D 75 (2007) 012007, arXiv:hep-ex/0609039.
- [61] T. Kluge, K. Rabbertz, M. Wobisch, arXiv:hep-ph/0609285, 2006.
- [62] T. Carli, et al., Eur. Phys. J. C 66 (2010) 503, arXiv:0911.2985.
- [63] J. Alwall, et al., arXiv:1405.0301, 2014.
- [64] S. Frixione, G. Ridolfi, Phys. Lett. B 383 (1996) 227, arXiv:hep-ph/9605209.



- [65] C. Simolo, arXiv:0807.1501, 2006.
- [66] J. Riedl, M. Stratmann, A. Schafer, Eur. Phys. J. C 73 (2013) 2360, arXiv:1212.1319.
- [67] C. Peterson, et al., Phys. Rev. D 27 (1983) 105.
- [68] G. Colangelo, P. Nason, Phys. Lett. B 285 (1992) 167.
- [69] J. Pumplin, et al., J. High Energy Phys. 0207 (2002) 012, arXiv:hep-ph/0201195.
- [70] A. Martin, W. Stirling, R. Thorne, Phys. Lett. B 636 (2006) 259, arXiv:hep-ph/0603143.
- [71] NNPDF Collaboration, R.D. Ball, et al., Nucl. Phys. B 867 (2013) 244, arXiv:1207.1303.
- [72] C. Bourrely, J. Soffer, Phys. Lett. B 314 (1993) 132.
- [73] G. Bunce, et al., Annu. Rev. Nucl. Part. Sci. 50 (2000) 525, arXiv:hep-ph/0007218.
- [74] C. Bourrely, J. Guillet, J. Soffer, Nucl. Phys. B 361 (1991) 72.
- [75] P. Chiappetta, et al., Z. Phys. C 59 (1993) 629.
- [76] PHENIX Collaboration, A. Adare, et al., Phys. Rev. D 86 (2012) 092006, arXiv:1202.4020.
- [77] D. de Florian, et al., Nucl. Phys. B 539 (1999) 455, arXiv:hep-ph/9808262.
- [78] S. Frixione, Z. Kunszt, A. Signer, Nucl. Phys. B 467 (1996) 399, arXiv:hep-ph/9512328.
- [79] S. Frixione, Nucl. Phys. B 507 (1997) 295, arXiv:hep-ph/9706545.
- [80] B. Jager, M. Stratmann, W. Vogelsang, Phys. Rev. D 70 (2004) 034010, arXiv:hep-ph/0404057.
- [81] A. Mukherjee, W. Vogelsang, Phys. Rev. D 86 (2012) 094009, arXiv:1209.1785.
- [82] C. Gagliardi, private communication.
- [83] M. Whalley, D. Bourilkov, R. Group, arXiv:hep-ph/0508110, 2005.
- [84] <https://lhapdf.hepforge.org/>.
- [85] E. Aschenauer, et al., arXiv:1304.0079, 2013.
- [86] D. de Florian, et al., arXiv:1404.4293, 2014.
- [87] Z.B. Kang, J. Soffer, Phys. Rev. D 83 (2011) 114020, arXiv:1104.2920.
- [88] E.R. Nocera, Nuovo Cimento C 036 (2013) 143, arXiv:1302.6409.
- [89] D. de Florian, W. Vogelsang, Phys. Rev. D 81 (2010) 094020, arXiv:1003.4533.
- [90] See <https://indico.cern.ch/getFile.py/access?contribId=304&sessionId=9&resId=0&materialId=slides&confId=184503>.
- [91] A. Martin, et al., Eur. Phys. J. C 28 (2003) 455, arXiv:hep-ph/0211080.
- [92] B. Jager, et al., Phys. Rev. D 67 (2003) 054005, arXiv:hep-ph/0211007.
- [93] B. Fuks, et al., J. High Energy Phys. 1405 (2014) 045, arXiv:1403.2383.
- [94] E.C. Aschenauer, et al., Phys. Rev. D 88 (2013) 114025, arXiv:1309.5327.