

First measurement of the Sivers asymmetry for gluons from SIDIS data

The COMPASS Collaboration

Abstract

The Sivers function describes the correlation between the transverse spin of a nucleon and the transverse motion of its partons. It was extracted from measurements of the azimuthal asymmetry of hadrons produced in semi-inclusive deep inelastic scattering of leptons off transversely polarised nucleon targets, and it turned out to be non-zero for quarks. In this letter the evaluation of the Sivers asymmetry for gluons in the same process is presented. The analysis method is based on a Monte Carlo simulation that includes three hard processes: photon-gluon fusion, QCD Compton scattering and leading-order virtual-photon absorption process. The Sivers asymmetries of the three processes are simultaneously extracted using the LEPTO event generator and a neural network approach. The method is applied to samples of events containing at least two hadrons with large transverse momentum from the COMPASS data taken with a 160 GeV/c muon beam scattered off transversely polarised deuterons and protons. With a significance of more than two standard deviations a negative value is obtained for the gluon Sivers asymmetry. The result of a similar analysis for a Collins-like asymmetry for gluons is consistent with zero.

(to be submitted to Phys. Lett. B)

1 Introduction

An interesting and recently examined property of the quark distribution in a nucleon that is polarised transversely to its momentum is the fact that it is not left-right symmetric with respect to the plane defined by the directions of nucleon spin and momentum. This asymmetry of the distribution function is called the Sivers effect and was first suggested [1] as an explanation for the large left-right single transverse spin asymmetries observed for pions produced in the reaction $p^\uparrow p \rightarrow \pi X$ [2, 3, 4]. On the basis of T-invariance arguments the existence of such an asymmetric distribution, known as Sivers distribution function, was originally excluded [5]. Ten years later it was recognised however that it was indeed possible [6]. At that time it was also predicted that the Sivers function in semi-inclusive measurements of hadron production in DIS (SIDIS) and in the Drell-Yan process have opposite sign [7], a property referred to as “restricted universality”. A few years later the Sivers effect was experimentally observed in SIDIS experiments on transversely polarised proton targets, first by the HERMES Collaboration [8] and then by the COMPASS Collaboration [9]. Using the first HERMES data and the early COMPASS data taken with a transversely polarised deuteron target [10], a combined analysis soon allowed for first extractions of the Sivers function for u and d-quarks [11, 12, 13]. More precise measurements of the Sivers effect were performed since by the HERMES [14] and COMPASS [15, 16, 17] Collaborations, and new measurements with a transversely polarised ^3He target were also carried out at JLab [18, 19]. More information can be found in recent reviews [20, 21, 22].

At this point the question arises whether the gluon distribution in a transversely polarised nucleon is left-right symmetric or exhibits a Sivers effect similar to the quark distributions. Recently, the issue has been discussed repeatedly in the literature and the properties of the gluon Sivers distributions have been studied in great detail [23, 24]. While it was found that a non-zero Sivers function implies motion of partons in the nucleon, presently the connection between the Sivers function and the parton orbital angular momentum in the nucleon can only be described in a model-dependent way [25]. The correspondence between the Sivers effect and the transverse motion of partons has been originally proposed by M. Burkardt [26, 27, 28]. Hence it is of great interest to know whether there exists a gluon Sivers effect or not.

Presently, little is known on the gluon Sivers function. An important theoretical constraint comes from the so-called Burkardt sum rule [29]. It states, based on the presence of QCD colour-gauge links, that the total transverse momentum of all partons inside a transversely polarised proton should vanish. Fits to the Sivers asymmetry using SIDIS data [13] almost fulfil, within uncertainties, the Burkardt sum rule, leaving little space for a gluon contribution. From the null result of the COMPASS experiment for the Sivers asymmetry of positive and negative hadrons produced on a transversely polarised deuteron target [10], together with additional theoretical considerations, Brodsky and Gardner [30] stated that the gluon contribution to the parton orbital angular momentum should be negligible, and consequently that the gluon Sivers effect should be small. Also, using the so-called transverse momentum dependent (TMD) generalised parton model and the most recent phenomenological information on the quark Sivers distributions coming from SIDIS data, interesting constraints on our knowledge of the gluon Sivers function were derived [31] from the recent precise data on the transverse single spin asymmetry $A_N(p^\uparrow p \rightarrow \pi^0 X)$ measured at central rapidity by the PHENIX Collaboration at RHIC [32].

In DIS, the leading-order virtual-photon absorption process (LP) does not provide direct access to the gluon distribution since the virtual-photon does not couple to the gluon, so that higher-order processes have to be studied, *i.e.* QCD Compton scattering (QCDC) and Photon-Gluon Fusion (PGF). It is well known that in lepton-proton scattering one of the most promising processes to directly probe the gluon is open charm production, $\ell p^\uparrow \rightarrow \ell' c\bar{c}X$. This is the channel that the COMPASS Collaboration has investigated at length in order to measure $\Delta g/g$, the gluon polarisation in a longitudinally polarised nucleon [33]. Tagging the charm quark by identifying D-mesons in the final state has the advantage that in the lowest order of the strong coupling constant there are no other contributions to the cross section and one becomes essentially sensitive to the gluon distribution function. An alternative method to tag the gluon in DIS, which has the advantage of higher statistics, has also been developed and used by COMPASS, *i.e.* the production of high- p_T hadrons [34, 35]. In the LP, the hadron transverse momentum p_T with respect to the virtual photon direction (in the frame where the nucleon momentum is parallel to this direction) originates from the intrinsic transverse momentum k_T of the struck quark in the nucleon and its fragmentation, which both lead to a small transverse component. On the contrary, both the QCDC and PGF hard processes can provide hadrons with high transverse momentum. Therefore, tagging events with hadrons of high transverse momentum p_T enhances the contribution of higher-order processes. Nevertheless, although in the high- p_T sample the PGF fraction is enriched, in order to single out the contribution of the PGF process to the measured asymmetry the contributions from LP and QCDC have to be subtracted [36].

In this letter, the gluon Sivers effect is investigated using COMPASS data collected by scattering a 160 GeV/c muon beam off transversely polarised deuterons and protons. The experimental set-up and the data selection are described in Section 2. In Section 3 the measurement is described. The details of the analysis are given in Section 4. The procedure of neural network (NN) training with a Monte Carlo data sample is shown in Section 5. Section 6 contains the overview of the systematic studies. In Section 7 the results are presented. Summary and conclusions are given in Section 8.

2 Experimental set-up and data samples

The COMPASS experiment uses a fixed target set-up and a polarised muon beam delivered by the M2 beam line of the CERN SPS. The transversely polarised deuteron target used for the 2003 and 2004 data taking consisted of two oppositely polarised cylindrical cells situated along the beam, each 60 cm long with a 10 cm gap in between. In 2010 the transversely polarised proton target consisted of three cells: 30 cm, 60 cm and 30 cm long with the central cell oppositely polarised to the downstream and upstream cell and 5 cm gaps between the cells. During all data taking periods the polarisation was reversed once per week, in this way systematic effects due to acceptance are cancelled. For the deuteron runs the target was filled with ^6LiD . The ^6Li nucleus can be regarded as one quasi-free deuteron and a ^4He core. The average dilution factor f_d , defined as the ratio of the DIS cross section on polarisable nucleons in the target to the cross section on all target nucleons, amounts to 0.36 and includes also electromagnetic radiative corrections. The average polarisation of the deuteron was 0.50. For the asymmetry measurements on the proton, NH_3 was used as a target. Its average dilution factor f_p amounts to 0.15 and the proton polarisation to 0.80. In both cases, the naturally polarised muon beam of 160 GeV/c was used. The basic features of the COMPASS spectrometer, as described in Ref. [37], are the same for 2003-4 and 2010 data taking. Several upgrades were performed in 2005, the main one being the installation of a new target magnet, which allowed to increase the polar angle acceptance from 70 mrad to 180 mrad.

A crucial point of this analysis is the search for an observable that is strongly correlated with the gluon azimuthal angle ϕ_g . In the LEPTO generator [38], gluons are accessed via PGF with a quark-antiquark pair in the final state and the fragmentation process is described by the Lund model [39]. As a result of MC studies, the best correlation is found between ϕ_g and ϕ_P , where the latter denotes the azimuthal angle of the vector sum \mathbf{P} of the two hadron momenta. For the present analysis, two charged hadrons for each event are selected. If more than two charged hadrons are reconstructed in an event, only the hadron with the largest transverse momentum, p_{T1} , and the one with the second-largest transverse momentum, p_{T2} , are taken into account. In order to enhance the PGF fraction in the sample and at the same time the correlation between ϕ_g and ϕ_P , a further requirement is applied to the transverse momenta of the two hadrons: $p_{T1} > 0.7 \text{ GeV}/c$ and $p_{T2} > 0.4 \text{ GeV}/c$. Moreover, the fractional energies of the two hadrons must fulfil the following conditions: $z_i > 0.1$ ($i=1,2$) and $z_1 + z_2 < 0.9$, where the last requirement rejects events from diffractive vector meson production. Hadron pairs are selected with no charge constraint. With this choice the correlation coefficient is 0.54. The Sivers asymmetry is then obtained as the sine modulation in the Sivers angle, $\phi_{Siv} = \phi_P - \phi_S$. Here ϕ_S is the azimuthal angle of the nucleon spin vector.

The same kinematic data selection is used for both deuteron and proton data. The requirement on photon virtuality, $Q^2 > 1 \text{ (GeV}/c)^2$, selects events in the perturbative region and the one on the mass of the hadronic final state, $W > 5 \text{ GeV}/c^2$, removes the region of the exclusive nucleon resonance production. The Bjorken- x variable covers the range $0.003 < x_{Bj} < 0.7$. For the fractional energy of the virtual photon, y , the limit $y > 0.1$ removes a region sensitive to experimental biases and the requirement $y < 0.9$ rejects events with large electromagnetic radiative corrections.

3 Sivers asymmetry in two hadron production

In order to extract the gluon Sivers asymmetry, $\mu + N \rightarrow \mu' + 2h + X$ events are selected as described in Section 2. By labelling with the symbol \uparrow the cross section associated to a target cell polarised upwards in the laboratory and by \downarrow the cross section of a target cell polarised downwards in the laboratory, the Sivers asymmetry can be written as

$$A_T^{2h}(\vec{x}, \phi_{Siv}) = \frac{\Delta\sigma(\vec{x}, \phi_{Siv})}{\sigma(\vec{x})}, \quad (1)$$

where $\vec{x} = (x_{Bj}, Q^2, p_{T1}, p_{T2}, z_1, z_2)$, $\Delta\sigma \equiv d^7\sigma\uparrow - d^7\sigma\downarrow$ and $\sigma \equiv d^7\sigma\uparrow + d^7\sigma\downarrow$. All cross sections are integrated over the two azimuthal angles ϕ_S and ϕ_R , where ϕ_R is the azimuthal angle of the relative momentum of the two hadrons, $\mathbf{R} = \mathbf{P}_1 - \mathbf{P}_2$. The number of events in a ϕ_{Siv} bin is given by

$$N(\vec{x}, \phi_{Siv}) = \alpha(\vec{x}, \phi_{Siv}) (1 + f P_T A^{Siv}(\vec{x}) \sin \phi_{Siv}). \quad (2)$$

Here f is the dilution factor, P_T the target polarisation and $\alpha = an\Phi\sigma_0$ an acceptance-dependent factor, where a is the total spectrometer acceptance, n the density of scattering centres, Φ the beam flux and σ_0 the spin-averaged part of the cross section. From here on, the Siverson asymmetry $A_T^{2h}(\vec{x}, \phi_{Siv})$ is factorised into the azimuth-independent amplitude $A^{Siv}(\vec{x})$ and the modulation $\sin \phi_{Siv}$.

In order to extract the Siverson asymmetry of the gluon, the amplitude of the $\sin \phi_{Siv}$ modulation is extracted from data. The general expression for the cross section of SIDIS production with at least one hadron in the final state is well known [40]. It contains eight azimuthal modulations, which are functions of the single-hadron azimuthal angle and ϕ_S . In the absence of correlations possibly introduced by experimental effects, they are all orthogonal, so that the Siverson asymmetry can either be extracted as the amplitude of the $\sin \phi_{Siv}$ modulation or one can perform a simultaneous fit of all eight amplitudes. For the case of heavy-quark pair and dijet production in lepton-nucleon collisions, all azimuthal asymmetries associated to the gluon distribution function have been recently worked out in Ref. 41. There, the Siverson asymmetry is defined as the amplitude of the $\sin(\phi_T - \phi_S)$ modulation, where ϕ_T is the azimuthal angle of the transverse-momentum vector of the quark-antiquark pair, \mathbf{q}_T . In our analysis, ϕ_T is replaced by ϕ_P , due to its correlation with the gluon azimuthal angle ϕ_g , and the Siverson asymmetry is extracted taking into account only the $\sin(\phi_P - \phi_S)$ modulation in the cross section. It has been verified that including in the cross section the same eight transverse-spin modulations as in SIDIS single-hadron production [40] and extracting simultaneously all asymmetries gives the same result on the gluon Siverson asymmetry.

In order to determine the Siverson asymmetry for gluons from two-hadron production in SIDIS, it is necessary to assume that the main contributors to muon-nucleon DIS are the three processes (Fig. 1) as presented in Ref. [38]. This model is successful in describing the unpolarised data. At COMPASS kinematics, the leading process appears at zero-order QCD in the total DIS cross section and it is the dominant process, while the other two processes, photon-gluon fusion and QCD Compton, are first-order QCD processes and hence suppressed. However, their contribution can be enhanced by constraining the transverse momentum of the produced hadrons, as mentioned above.

Introducing the process fractions $R_j = \sigma_j/\sigma$ ($j \in \{\text{PGF}, \text{QCDC}, \text{LP}\}$), the amplitude of the Siverson asymmetry can be expressed in terms of the amplitudes of the three contributing processes:

$$\begin{aligned} f P_T A^{Siv} \sin \phi_{Siv} &= \frac{\Delta\sigma}{\sigma} = \frac{\sigma_{\text{PGF}}}{\sigma} \frac{\Delta\sigma_{\text{PGF}}}{\sigma_{\text{PGF}}} + \frac{\sigma_{\text{QCDC}}}{\sigma} \frac{\Delta\sigma_{\text{QCDC}}}{\sigma_{\text{QCDC}}} + \frac{\sigma_{\text{LP}}}{\sigma} \frac{\Delta\sigma_{\text{LP}}}{\sigma_{\text{LP}}} \\ &= f P_T (R_{\text{PGF}} A_{\text{PGF}}^{Siv} + R_{\text{QCDC}} A_{\text{QCDC}}^{Siv} + R_{\text{LP}} A_{\text{LP}}^{Siv}) \sin \phi_{Siv}, \end{aligned} \quad (3)$$

with $\sigma = \sum_j \sigma_j$, $\Delta\sigma = \sum_j \Delta\sigma_j$ and $f P_T A_j^{Siv} \sin \phi_{Siv} = \Delta\sigma_j/\sigma_j$. The determination of R_j is done on an event-by-event basis by using the neural networks (NN) trained on Monte Carlo data as described in Section 5.

4 Asymmetry extraction using the methods of weights

The method adopted in the present analysis was already applied to extract the gluon polarisation from the longitudinal double-spin asymmetry in the SIDIS measurement of single-hadron production [36]. Both for the deuteron data (two target cells) and the proton data (three target cells), four target configurations can be introduced. In the case of the two-cell target: 1 - upstream, 2 - downstream, 3 - upstream', 4 - downstream'. In the case of the three-cell target: 1 - (upstream+downstream), 2 - centre, 3 - (upstream'+downstream'), 4 - centre'. Here upstream', centre' and downstream' denote the cells after

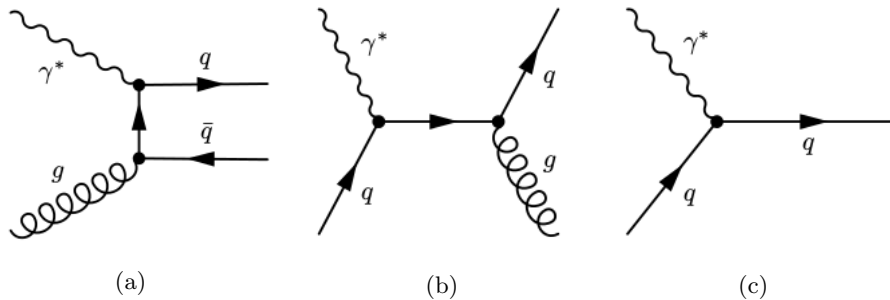


Figure 1: Feynman diagrams considered for $\gamma^* N$ scattering: a) photon-gluon fusion (PGF) , b) gluon radiation (QCD Compton scattering), c) Leading order process (LP).

the polarisation reversal and configuration 1 has the polarisation pointing upwards in the laboratory frame. Decomposing the Sivers asymmetry into the asymmetries of the contributing processes (Eq. (3)) and introducing the Sivers modulation $\beta_j^t(\vec{x}, \phi_{\text{Siv}}) = R_j(\vec{x})f(\vec{x})P_T^t \sin \phi_{\text{Siv}}$, which is specific for process j , one can rewrite Eq. (2):

$$N^t(\vec{x}, \phi_{\text{Siv}}) = \alpha^t(\vec{x}, \phi_{\text{Siv}}) \left(1 + \beta_{\text{PGF}}^t(\vec{x}, \phi_{\text{Siv}}) A_{\text{PGF}}^{\text{Siv}}(\vec{x}) + \beta_{\text{QCDC}}^t(\vec{x}, \phi_{\text{Siv}}) A_{\text{QCDC}}^{\text{Siv}}(\vec{x}) + \beta_{\text{LP}}^t(\vec{x}, \phi_{\text{Siv}}) A_{\text{LP}}^{\text{Siv}}(\vec{x}) \right), \quad (4)$$

where $t = 1, 2, 3, 4$ denotes the target configuration.

In order to minimise statistical uncertainties for each process, a weighting factor is introduced. It is known [42] that the choice $\omega_j = \beta_j$ for the weight optimises the statistical uncertainty but variations of the target polarisation P_T in time may introduce a bias to the final result. Therefore, the weighting factor $\omega_j \equiv \beta_j/P_T$ is used instead. Each of the four equations (4) is weighted three times with ω_j depending on the process $j \in \{\text{PGF}, \text{QCDC}, \text{LP}\}$ and integrated over ϕ_{Siv} and \vec{x} , yielding twelve observed quantities q_j^t :

$$q_j^t = \int d\vec{x} d\phi_{\text{Siv}} \omega_j(\vec{x}, \phi_{\text{Siv}}) N^t(\vec{x}, \phi_{\text{Siv}}) = \tilde{\alpha}_j^t \left(1 + \{\beta_{\text{PGF}}^t\}_{\omega_j} \{A_{\text{PGF}}^{\text{Siv}}\}_{\beta_{\text{PGF}}\omega_j} + \{\beta_{\text{QCDC}}^t\}_{\omega_j} \{A_{\text{QCDC}}^{\text{Siv}}\}_{\beta_{\text{QCDC}}\omega_j} + \{\beta_{\text{LP}}^t\}_{\omega_j} \{A_{\text{LP}}^{\text{Siv}}\}_{\beta_{\text{LP}}\omega_j} \right), \quad (5)$$

where $\tilde{\alpha}_j^t$ is the ω_j -weighted acceptance-dependent factor. The quantities $\{\beta_i^t\}_{\omega_j}$ and $\{A_i^{\text{Siv}}\}_{\beta_i^t\omega_j}$ are weighted averages, where the weight factor is denoted in the subscript ¹⁾.

The acceptance factors $\tilde{\alpha}_j^t$ cancel when for asymmetry extraction one uses the double ratio

$$r_j := \frac{q_j^1 q_j^4}{q_j^2 q_j^3} \quad (6)$$

as the data taking was performed such that $\tilde{\alpha}_j^1 \tilde{\alpha}_j^4 / \tilde{\alpha}_j^2 \tilde{\alpha}_j^3 = 1$. If this condition is not fulfilled, false asymmetries may occur. It is checked that this is not the case (see Section 6).

In the analysis, the quantities q_j and $\{\beta_i^t\}_{\omega_j}$ are approximated as follows:

$$q_j^t \approx \sum_{k=1}^{N_t} \omega_j^k, \quad (7)$$

$$\{\beta_i^t\}_{\omega_j} \approx \frac{\sum_{k=1}^{N_t} \beta_i^{t,k} \omega_j^k}{\sum_{k=1}^{N_t} \omega_j^k}. \quad (8)$$

The latter approximation holds for small observed raw asymmetries, *i.e.* $\omega A \ll 1$. In order to avoid numerical inconsistencies in Eq. (8) due to a zero-pole when integrating over the full range of ϕ_{Siv} ²⁾, two bins in ϕ_{Siv} ($[0; \pi]$, $[\pi; 2\pi]$) are introduced. In the aforementioned three double ratios given in Eq. (6) only asymmetries are unknown. However, in order to solve the system of equations one needs to assume that the weighted asymmetry for a given process i is the same for the three different weights $\omega_j \beta_i$, *i.e.* $\{A_i\}_{\beta_i \omega_{\text{PGF}}} = \{A_i\}_{\beta_i \omega_{\text{QCDC}}} = \{A_i\}_{\beta_i \omega_{\text{LP}}}$. This means that the values of ω_j and A_i must be uncorrelated. For example, since ω_j is proportional to R_j , which strongly depends on the hadron transverse momentum, one has to use a kinematic region where the asymmetries A_i are expected to be independent of p_T . Under these assumptions, the number of unknown weighted asymmetries is three, which exactly corresponds to the number of equations of type (6). These equations are solved by a χ^2 fit that includes simultaneously both bins in ϕ_{Siv} .

1) $\{\beta\}_{\omega} = \frac{\int \alpha \beta \omega d\vec{x} d\phi_{\text{Siv}}}{\int \alpha \omega d\vec{x} d\phi_{\text{Siv}}}$, $\{A\}_{\beta\omega} = \frac{\int A \alpha \beta \omega d\vec{x} d\phi_{\text{Siv}}}{\int \alpha \beta \omega d\vec{x} d\phi_{\text{Siv}}}$

2) Note that ω_j^k , which contains $\sin \phi_{\text{Siv}}$, is integrated in the region 0 to 2π .

Assuming that A_i can be approximated by a linear function of x_i and that x_i is not correlated with ω_j , results in

$$\{A_i\}_{\beta_i \omega_i} = A_i(\{x_i\}_{\beta_i \omega_i}). \quad (9)$$

This approximation allows to interpret the obtained results as an asymmetry value measured at the weighted value of x_i . For each process, the weighted value of x_i is obtained from MC using the relation

$$\{x_i\}_{\beta_i \omega_i} = \frac{\sum_{k=1}^{N_i} x_i^k \beta_i^k \omega_i^k}{\sum_{k=1}^{N_i} \beta_i^k \omega_i^k}. \quad (10)$$

Here, N_i is the number of events of type i in MC data. The assumption that the values of x_i are not correlated with ω_j , which allows us to consider only $\{x_i\}_{\omega_i \beta_i}$, was verified using MC data. The details of the analysis are given in [43].

5 Monte Carlo optimisation and Neural Network training

The present analysis is very similar to the one used for the $\Delta g/g$ extraction from high- p_T hadron pairs [35] and single hadrons [36]. For the NN training with custom input, output and target vector the package NetMaker [44] is used. The NN is trained with a Monte Carlo sample with process identification. As input vector the following six kinematic variables are chosen: x_{Bj} , Q^2 , p_{T1} , p_{T2} , p_{L1} , p_{L2} . The latter two are the longitudinal components of the hadron momenta. The trained neural network is applied to the data by taking the vector of the aforementioned six variables, and its output is interpreted as probabilities that the given event is a result of one of the three contributing processes. Hence the simulated distributions of these variables need to be in agreement with the corresponding distributions in the data samples.

Using the LEPTO generator (version 6.5) [38], two separate MC data samples were produced to simulate the deuteron and proton data. The generator is tuned to the COMPASS data sample obtained with the high- p_T hadron-pair selection as described in Ref. [35]. The MSTW08 parameterisation of input PDFs [45] was chosen as it gives a good description of the F_2 structure function in the COMPASS kinematic range and is valid down to $Q^2 = 1$ (GeV/c)². Electromagnetic radiative corrections [46] were applied as a weighting factor to the MC distributions shown in Figures 2 and 3 but not in the MC samples used in NN training. This difference was studied and it was estimated to be negligible.

The generated events were processed by COMGEANT, the COMPASS detector simulation program based on GEANT3. The MC samples for the proton and deuteron data differ in the target material and in the spectrometer set-up. The FLUKA package [47] is used in order to simulate secondary interactions. As the next step, the COMPASS reconstruction program CORAL was applied. The same data selection as for real events was used for MC data. Figures 2 and 3 show the comparison between experimental and MC data for the case of the deuteron and proton data, respectively.

The main goal of the NN parameterisation is the estimation of the process fractions R_j . In the typical case of signal and background separation, the expected NN output would be set to one for the signal and zero for the background. The output value returned by the NN would then correspond to the fraction of signal events in the sample in the given phase space point of the input parameter vector. In the present analysis, the process fractions were estimated simultaneously. In order to have a closure relation on the process probabilities, the sum of them must add up to one, hence only two independent output variables from the NN are needed. The estimation of the process fractions R^j from the NN output is accomplished by assigning to each event the probabilities $P_{\text{NN}}^{\text{PGF}}$, $P_{\text{NN}}^{\text{QCDC}}$ and $P_{\text{NN}}^{\text{LP}}$. The distribution of the NN output after training is shown in Fig. 4 on the “Mandelstam representation”, i.e. as points in an equilateral triangle with unit height. Points outside of the triangle refer to one estimator being negative, which is possible because in the training the estimators are not bound to be positive. The direct separation of the PGF process using this distribution is statistically less efficient than weighting each event by the three probabilities obtained from the NN output values. These probabilities are used as values of the process fractions (R_{PGF} , R_{QCDC} and R_{LP}) in the data analysis described in Section 4.

For the validation of the NN training, a statistically independent MC sample is used to check how the NN works on a sample different from the one used for the training. In each bin of P_{NN} (the value assigned to every MC event by the trained NN), the true fraction obtained from LEPTO based on the process ID, P_{MC} , is calculated. The results for the NN trained with a MC sample for the proton data are presented in Fig. 5. Altogether, the agreement between P_{NN} and P_{MC} for all three processes is satisfactory. However, for the two last bins of PGF, the two last bins of QCDC and the last bin of LP the neural network output does not coincide with the true fraction of the given process. This discrepancy concerns

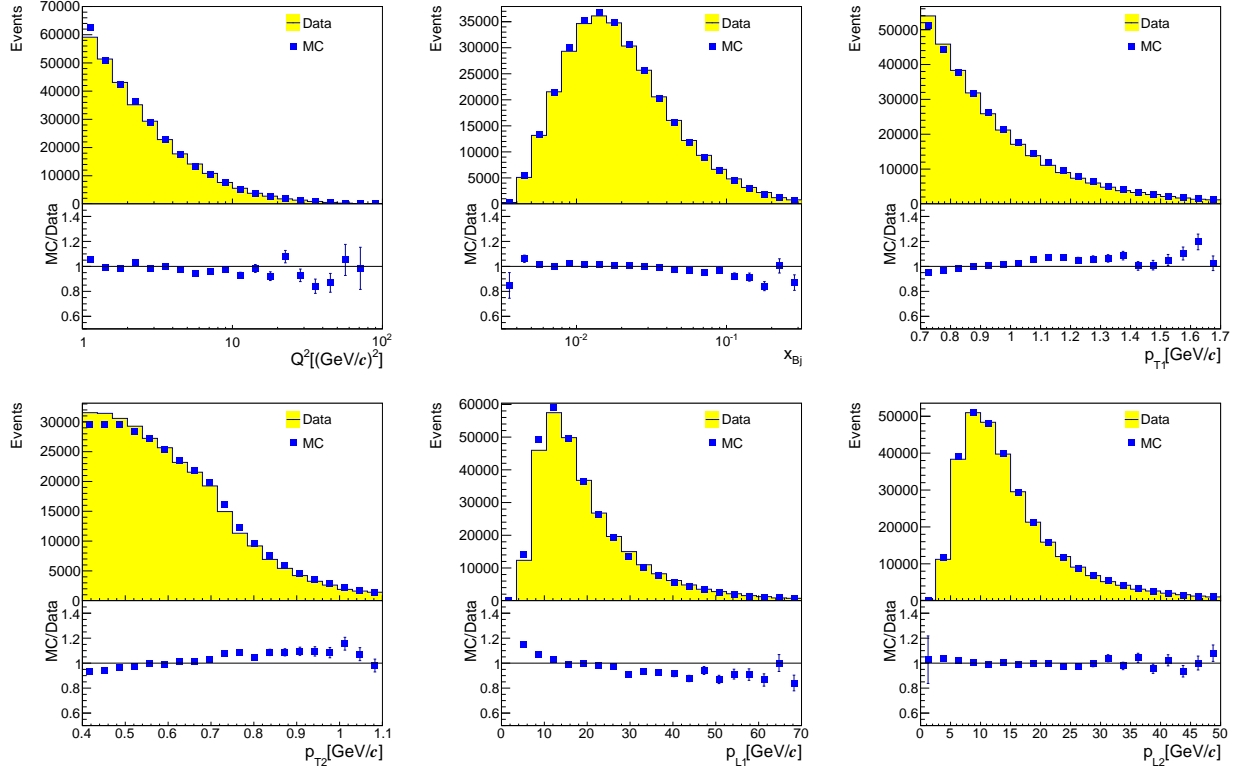


Figure 2: Comparison of distributions of kinematic variables between experimental and MC high- p_T deuteron data.

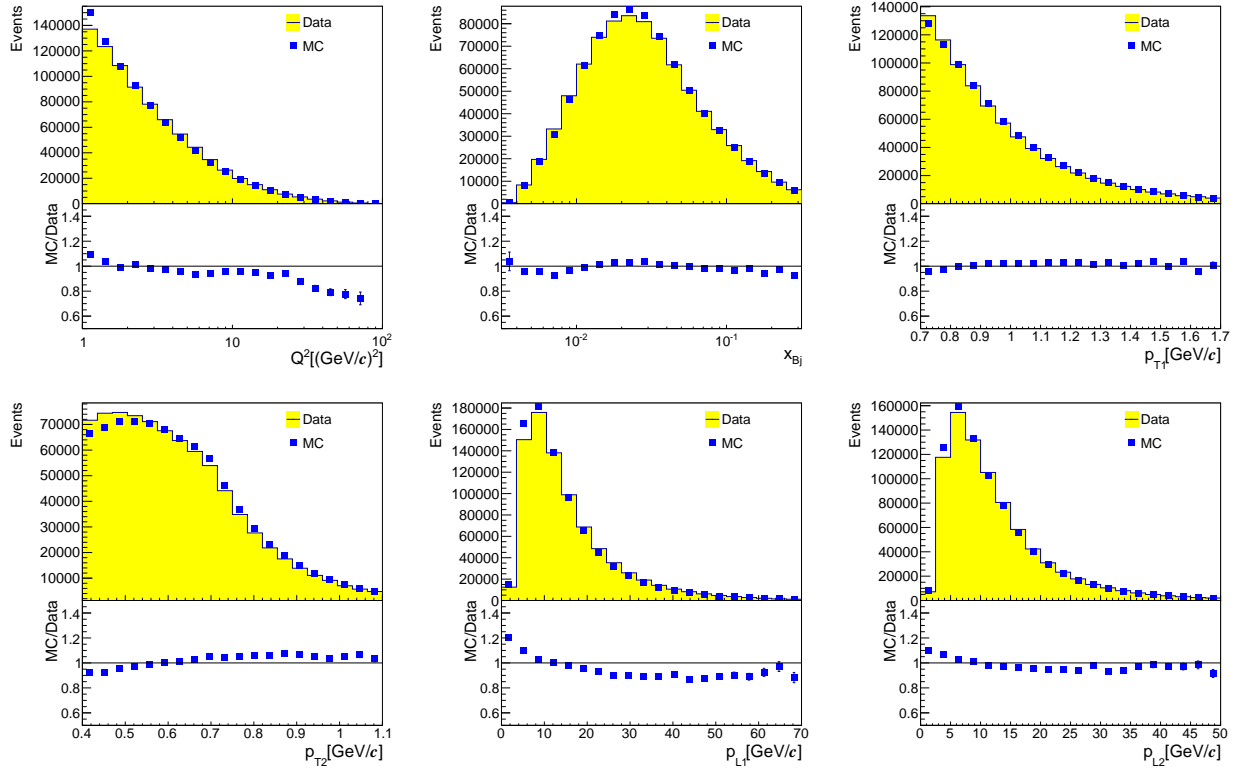


Figure 3: Comparison of distributions of kinematic variables between experimental and MC high- p_T proton data.

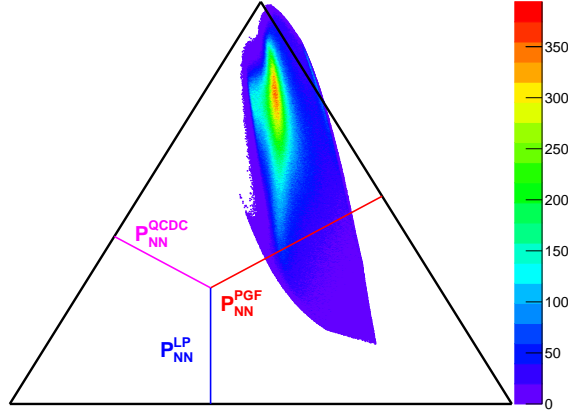


Figure 4: Neural network output after training the MC sample for the proton analysis. The probability of each process for a point representing an event is proportional to the distance to the opposite side of the triangle as in the “Mandelstam representation”. An exemplary set of probabilities assigned to a point is shown by the colour lines.

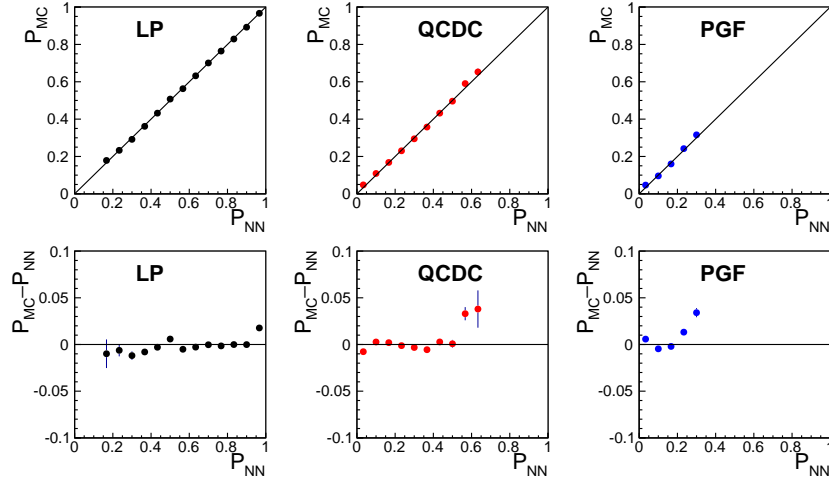


Figure 5: Top panels: Neural network validation. Here P_{NN} is the fraction of the process given by the NN and P_{MC} is the true fraction of each process from MC in a given P_{NN} bin. Bottom panels: Difference $P_{MC} - P_{NN}$ per bin.

a small part of the event sample and is included as a part of the MC-dependent systematic uncertainty. Because of the good agreement between MC and real data, it is assumed that the fractions in Eq. 3 can be taken from the trained NN, which means that on average $R_j = P_{NN}^j$.

6 Systematic uncertainties

The main source of systematic uncertainties is the dependence of the final results on the Monte Carlo settings and tuning. In order to estimate this uncertainty, different MC settings were used in the process of neural network training. Different combinations of fragmentation parameters were used, the default LEPTO tuning or the COMPASS tuning for the high- p_T selected sample. The event generation was done with and without the ‘Parton Shower’ [48]. Two PDF sets were used (MSTW08 or CTEQ5L [49]). Two different parameterisations of the longitudinal structure function F_L are used, either from LEPTO or the $R = \sigma_L/\sigma_T$ parameterisation of Ref. [50]. For secondary interactions, either the FLUKA or the GHEISHA [51] package were used.

Figure 6 shows the results for the gluon Sivers asymmetry when using eight different MC productions of the deuteron and proton data. The final result is presented on the top. These two MC productions, using FLUKA and Parton Shower, yield the best comparison between experimental and MC data, which is shown in Fig. 2 and 3. The systematic uncertainty originating from different MC tunings is calculated as $(A_{max}^{PGF} - A_{min}^{PGF})/2$.

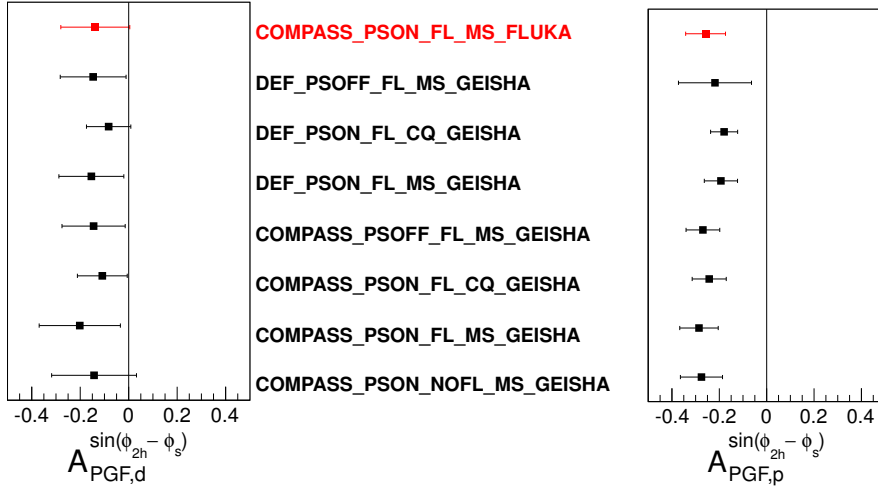


Figure 6: Systematic changes in the final result caused by using different MC settings. Besides the final result shown on the top, seven other results are shown that are obtained with MC samples that differ by the choice of COMPASS or default LEPTO tuning, ‘Parton Shower’ on or off, F_L from LEPTO or from $R = \sigma_L/\sigma_T$, MSTW or CTEQ5L PDF sets, FLUKA or GHEISHA for secondary interactions. The results for deuteron data are shown in the left panel and for the proton data on the right panel.

The systematic uncertainty due to false asymmetries was studied by extracting the asymmetries between the two parts of the same target cell. The results are found to be compatible with zero. Furthermore it was checked how a small artificial false Siverts asymmetry influences the final result. When a false asymmetry of 1% is introduced, for both proton and deuteron data the final result changes by 25% of the statistical error. No systematic uncertainty is assigned to account for false asymmetries.

The final state of the photon-gluon-fusion process is a quark-antiquark pair. Thus most of the hadron pairs produced from this subprocess should have opposite charge. Although a selection $q_1 q_2 = -1$ slightly increases the (ϕ_g, ϕ_P) correlation, it also reduces the statistics. The results with and without this requirement are statistically consistent. The requirement of opposite charges of the two hadrons is hence not included in the data selection.

Radiative corrections were not included in the MC production that is used in the main analysis of this letter. In order to estimate the systematic uncertainty introduced by this omission, a separate MC sample is used that was produced for the 2006 COMPASS set-up including radiative corrections based on RADGEN [52]. The difference in the final value for the gluon Siverts asymmetry for the proton is only 0.018, which corresponds to 21% of the statistical uncertainty. A corresponding systematic uncertainty is assigned due to the fact that radiative corrections are not included in the MC simulations and hence in the NN training.

Our results are obtained in only one x_g bin for $A_{\text{PGF}}^{\text{Siv}}$, one x_C bin for $A_{\text{QDC}}^{\text{Siv}}$ and one x_{Bj} bin for $A_{\text{LP}}^{\text{Siv}}$. As the asymmetries are strongly correlated binning in x_{Bj} affects the values of $A_{\text{PGF}}^{\text{Siv}}$ and $A_{\text{QDC}}^{\text{Siv}}$, which are extracted in a single bin as before. The $A_{\text{PGF}}^{\text{Siv}}$ result changes by 0.07 for deuteron data and 0.011 for proton data when two x_{Bj} bins are introduced, and these values are taken as an estimate of the related systematic uncertainty (see Table 1).

The asymmetries A_j^{Siv} of Eq. (4) were also extracted using the unbinned maximum likelihood method that, as expected, yields the same results of A_{PGF} as the above described analysis. Concerning the orthogonality of different modulations of the cross section, it was checked by what amount the Siverts asymmetry changes, when also the other seven asymmetries were included in the fit (see above). The change in the final result of the PGF asymmetry is negligible for both deuteron and proton data.

The systematic uncertainties on target polarisation and dilution factor are multiplicative and estimated to be about 5% and 2% of the statistical uncertainty, respectively. The final systematic uncertainty is obtained by summing all components in quadrature. All above mentioned contributions and the final systematic uncertainty are listed in Table 1.

7 Results

The method presented in Section 4 with the use of trained neural networks was applied to the two data sets described in Section 2. The gluon Siverts asymmetry as extracted from lepton nucleon DIS, in

source	deuteron data		proton data	
	uncertainty	fraction of σ_{stat}	uncertainty	fraction of σ_{stat}
Monte Carlo settings	0.060	40%	0.054	64%
radiative corrections	0.018	12%	0.018	21%
one or two x_{Bj} bins	0.07	47%	0.011	13%
include 7 other asymmetries	0.003	2%	0.005	6%
target polarisation	0.0075	5%	0.0043	5%
dilution factor	0.003	2%	0.0018	2%
total $\sqrt{\sum \sigma_i^2}$	0.10	63%	0.06	69%

Table 1: Summary on systematic uncertainties of the final values of the gluon Sivers asymmetry for deuteron and proton data.

which at least two high- p_T hadrons are detected, is shown in Fig. 7 and presented in Table 2 together with the contribution of the two other hard processes, *i.e.* QCD Compton and leading process. The result of the analysis of the deuteron data is $A_{PGF}^{Siv,d} = -0.14 \pm 0.15(\text{stat.}) \pm 0.10(\text{syst.})$ measured at $\langle x_g \rangle = 0.13$. The proton result, $A_{PGF}^{Siv,p} = -0.26 \pm 0.09(\text{stat.}) \pm 0.06(\text{syst.})$ obtained at $\langle x_g \rangle = 0.15$, is consistent with the deuteron result within less than one standard deviation of the combined statistical uncertainty. The two results are expected to be consistent, as presumably the transverse motion of gluons is the same in neutron and proton. Combining the proton and deuteron results, the measured effect is negative, $A_{PGF}^{Siv} = -0.23 \pm 0.08(\text{stat}) \pm 0.05(\text{syst})$, which is away from zero by more than two standard deviations of the quadratically combined uncertainty. This result is particularly interesting in view of the gluon contribution to the proton spin. A non-zero gluon Sivers effect is a signature of gluon transverse motion in the proton [25]. The recent analysis of the PHENIX data [31] gives a gluon Sivers effect for protons, which is compatible with zero. The COMPASS result for the proton target is negative and more than two standard deviations below zero, but it should be noted that the two results are obtained for different centre of mass energy and x_g values. Even more important, one has to recall that the existence of colour gauge links complicates the picture, as they lead to two different universal gluon Sivers functions, which in the different processes combine with process-dependent calculable factors [53]. As a result, the gluon Sivers function that appears in one process can be different from the one appearing in a different process, and assessment of compatibility requires a deeper theoretical analysis.

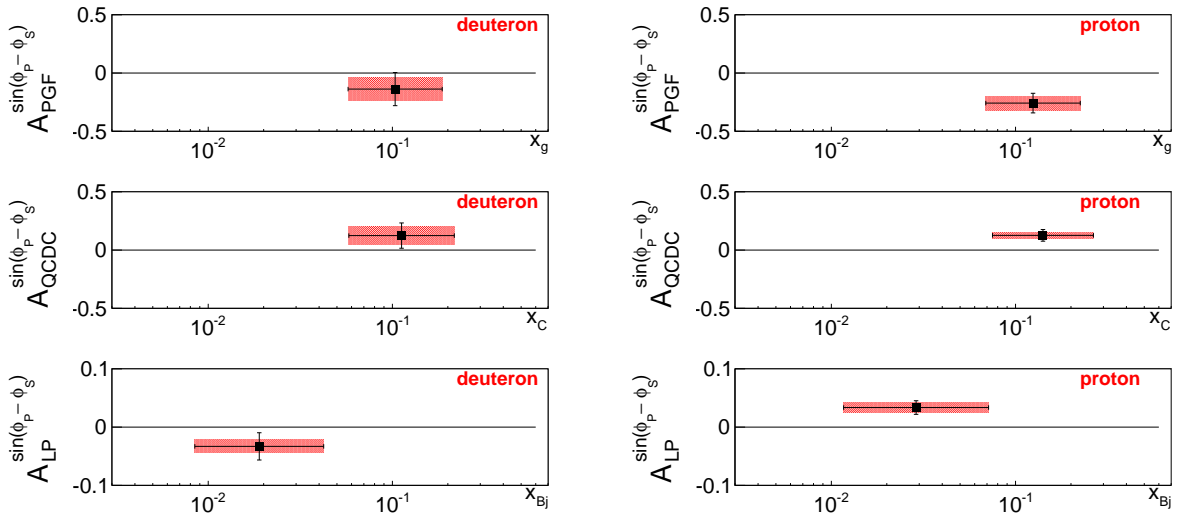


Figure 7: Sivers two-hadron asymmetry extracted for Photon-Gluon fusion (PGF), QCD Compton (QCDC) and Leading Process (LP) from the COMPASS high- p_T deuteron (left) and proton (right) data. The x range is the RMS of the logarithmic distribution of x in the MC simulation. The red bands indicate the systematic uncertainties. Note the different ordinate scale used in the third row of panels.

	deuteron data			proton data		
process	asymmetry	statistical error	systematic uncertainty	asymmetry	statistical error	systematic uncertainty
PGF	-0.14	0.15	0.10	-0.26	0.09	0.06
QCDC	0.12	0.11	0.08	0.13	0.05	0.03
LP	-0.03	0.02	0.01	0.03	0.01	0.01

Table 2: Summary of Sivers asymmetries, A_{PGF}^{Siv} , A_{QCDC}^{Siv} , A_{LP}^{Siv} , obtained for deuteron and proton data.

For the asymmetry of the leading process, the high- p_T sample of the COMPASS proton data has provided a positive value (see Fig.7 right-bottom panel). It can be compared with the COMPASS results on the Sivers asymmetry for charged hadrons produced in SIDIS $\ell p \rightarrow \ell' h^\pm X$ single-hadron production [17], which for negative hadrons was found to be about zero and for positive hadrons different from zero and positive, so that for the two-hadron final state a positive value may indeed be expected.

The same analysis method was also applied to extract the Collins-like asymmetry for charged hadrons, *i.e.* the cross section dependence on the sine of the Collins angle $(\phi_P + \phi_S - \pi)$. To this purpose, the asymmetries $A_{PGF}^{\sin(\phi_P + \phi_S - \pi)}$, $A_{QCDC}^{\sin(\phi_P + \phi_S - \pi)}$, $A_{LP}^{\sin(\phi_P + \phi_S - \pi)}$ were determined for the same COMPASS high- p_T deuteron and proton data samples. The results are shown in Fig. 8. The amplitude of the Collins modulation for gluons is found to be consistent with zero, in agreement with the naive expectation that is based on the fact that there is no gluon transversity distribution [54]. Recently it was suggested that a transversity-like TMD gluon distribution h_1^g could generate a $\sin(\phi_S + \phi_T)$ modulation in lepton production of two jets or heavy quarks [41]. In this case the results shown in Fig. 8 provide a bound to the size of h_1^g . The results given in the present letter can also be interpreted such that no false systematic asymmetry is introduced by the rather complex analysis method used, and that the result obtained for the gluon Sivers asymmetry, which is definitely different from zero, is strengthened. It should also be noted that the Collins-like asymmetry of the leading process for the proton is found to be consistent with zero for high- p_T hadron pairs, in qualitative agreement with the measurement of the Collins asymmetry in single-hadron SIDIS measurement [55], where opposite values of about equal size were observed for positive and negative hadrons.

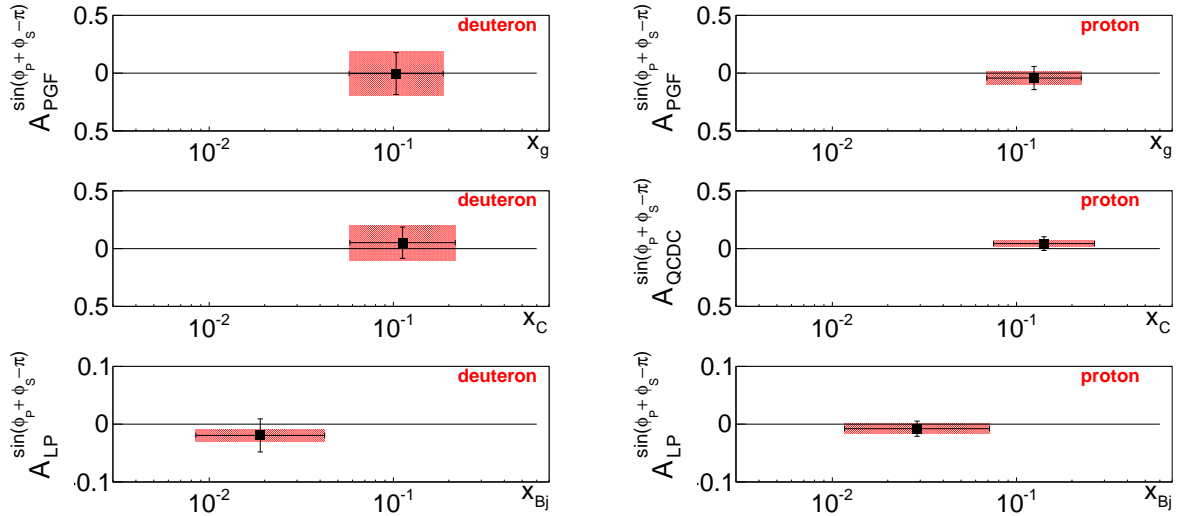


Figure 8: Collins-like two-hadron asymmetry extracted for Photon-Gluon fusion (PGF), QCD Compton (QCDC) and Leading Process (LP) from the COMPASS high- p_T deuteron (left) and proton (right) data. The x range is the RMS of the logarithmic distribution of x in the MC simulation. The red bands indicate the systematic uncertainties.

8 Summary and conclusions

The Sivers asymmetry for gluons is extracted from the measurement of high- p_T hadron pairs in SIDIS at COMPASS off transversely polarised deuterons and protons. The analysis is very similar to the one already used by the COMPASS collaboration in order to measure $\Delta g/g$, the gluon polarisation in a longitudinally polarised nucleon. The large kinematic acceptance and the high energy of the muon beam make the sample containing two high- p_T hadrons sufficiently large for the present analysis, which is limited to a small part of the accessible phase space. The criteria applied to select hadron pairs allow to enhance the contribution of the photon-gluon fusion process with respect to the leading-order

virtual-photon absorption process. The Sivers asymmetry was then obtained as the amplitude of the sine modulation in the Sivers angle, $\phi_{Siv} = \phi_P - \phi_S$.

In spite of the enrichment of the PGF fraction in the high- p_T hadron pair sample, in order to single out the contribution of the PGF process to the asymmetry it is necessary to subtract the contributions from the other two processes, LP and QCDC. In this analysis, the fractions of the three processes were determined from MC algorithms, and the three corresponding asymmetries were extracted from the data using a NN technique. Therefore, the analysis requires a precise MC description of the data, so that these quantities can be calculated reliably. Since the results derived from a NN approach strongly depend on the Monte Carlo sample on which the network was trained, much effort was devoted to obtain a good description of the experimental data by MC simulations, and the analysis was repeated using eight different MCs and the (small) differences in the results were included in the systematic uncertainties.

Averaging results obtained from the deuteron and proton data, the measured gluon Sivers asymmetry turns out to be $-0.23 \pm 0.08(\text{stat.}) \pm 0.05(\text{syst.})$, which is away from zero by more than two standard deviations of the total experimental uncertainty. This result supports the existence of a transverse motion of gluons in a transversely polarised nucleon, although the quantification of this motion is model-dependent.

In addition, another result obtained in this work from the same data is the extraction of the Collins-like gluon asymmetry, *i.e.* the amplitude of the sine modulation of the Collins angle $\phi_{Col} = \phi_P + \phi_S - \pi$. Recent developments have hypothesised a non-zero Collins-like gluon asymmetry that however is not related to transversity. Our result on the Collins-like asymmetry, which is obtained from the same hadron-pair data that we used to extract the non-zero result on the gluon Sivers asymmetry, is found to be compatible with zero.

Acknowledgements

This work was made possible thanks to the financial support of our funding agencies. We also acknowledge the support of the CERN management and staff, as well as the skills and efforts of the technicians of the collaborating institutes.

References

- [1] D. W. Sivers, Phys. Rev. **D41**, 83 (1990).
- [2] J. Antille *et al.*, Phys. Lett. **B94**, 523 (1980).
- [3] FNAL-E704, D. L. Adams *et al.*, Phys. Lett. **B264**, 462 (1991).
- [4] E704, E581, D. L. Adams *et al.*, Phys. Lett. **B261**, 201 (1991).
- [5] J. C. Collins, Nucl. Phys. **B396**, 161 (1993), hep-ph/9208213.
- [6] S. J. Brodsky, D. S. Hwang, and I. Schmidt, Phys. Lett. **B530**, 99 (2002), hep-ph/0201296.
- [7] J. C. Collins, Phys. Lett. **B536**, 43 (2002), hep-ph/0204004.
- [8] HERMES Collaboration, A. Airapetian *et al.*, Phys. Rev. Lett. **94**, 012002 (2005), hep-ex/0408013.
- [9] COMPASS Collaboration, M. G. Alekseev *et al.*, Phys. Lett. **B692**, 240 (2010), hep-ex/1005.5609.
- [10] COMPASS Collaboration, V. Yu. Alexakhin *et al.*, Phys. Rev. Lett. **94**, 202002 (2005), hep-ex/0503002.
- [11] W. Vogelsang and F. Yuan, Phys. Rev. **D72**, 054028 (2005), hep-ph/0507266.
- [12] A. V. Efremov, K. Goeke and P. Schweitzer, Eur. Phys. J. ST **162** (2008) 1, hep-ph/0801.2238.
- [13] M. Anselmino *et al.*, Eur. Phys. J. **A39**, 89 (2009), hep-ph/0805.2677.
- [14] HERMES Collaboration, A. Airapetian *et al.*, Phys. Rev. Lett. **103**, 152002 (2009), hep-ex/0906.3918.
- [15] COMPASS Collaboration, E. S. Ageev *et al.*, Nucl. Phys. **B765**, 31 (2007), hep-ex/0610068.
- [16] COMPASS Collaboration, M. Alekseev *et al.*, Phys. Lett. **B673**, 127 (2009), hep-ex/0802.2160.
- [17] COMPASS Collaboration, C. Adolph *et al.*, Phys. Lett. **B717**, 383 (2012), hep-ex/1205.5122.
- [18] Jefferson Lab Hall A, X. Qian *et al.*, Phys. Rev. Lett. **107**, 072003 (2011), nucl-ex/1106.0363.
- [19] Jefferson Lab Hall A Collaboration, Y. X. Zhao *et al.*, Phys. Rev. **C90**, 055201 (2014), nucl-ex/1404.7204.
- [20] V. Barone, F. Bradamante, and A. Martin, Prog. Part. Nucl. Phys. **65**, 267 (2010), hep-ph/1011.0909.
- [21] C. A. Aidala, S. D. Bass, D. Hasch, and G. K. Mallot, Rev. Mod. Phys. **85**, 655 (2013), hep-ph/1209.2803.
- [22] H. Avakian, A. Bressan, and M. Contalbrigo, Eur. Phys. J. **A52**, 150 (2016), [Erratum: Eur. Phys. J. A52, no. 6, 165 (2016)].
- [23] P. J. Mulders and J. Rodrigues, Phys. Rev. **D63**, 094021 (2001), hep-ph/0009343.
- [24] D. Boer, C. Lorcé, C. Pisano, and J. Zhou, Adv. High Energy Phys. **2015**, 371396 (2015), hep-ph/1504.04332 and the reference therein.

- [25] M. Burkardt, Phys. Rev. **D66**, 114005 (2002), hep-ph/0209179.
- [26] M. Burkardt, Int. J. Mod. Phys. **A18**, 173 (2003), hep-ph/0207047.
- [27] M. Burkardt, Nucl. Phys. **A735**, 185 (2004), hep-ph/0302144.
- [28] M. Burkardt and D. S. Hwang, Phys. Rev. **D69**, 074032 (2004), hep-ph/0309072.
- [29] M. Burkardt, Phys. Rev. **D69**, 091501 (2004), hep-ph/0402014.
- [30] S. J. Brodsky and S. Gardner, Phys. Lett. **B643**, 22 (2006), hep-ph/0608219.
- [31] U. D'Alesio, F. Murgia, and C. Pisano, JHEP **09**, 119 (2015), hep-ph/1506.03078.
- [32] PHENIX Collaboration, A. Adare *et al.*, Phys. Rev. **D90**, 012006 (2014), hep-ex/1312.1995.
- [33] COMPASS Collaboration, C. Adolph *et al.*, Phys. Rev. **D87**, 052018 (2013), hep-ex/1211.6849.
- [34] A. Bravar, D. von Harrach, and A. Kotzinian, Phys. Lett. **B421**, 349 (1998), hep-ph/9710266.
- [35] COMPASS Collaboration, C. Adolph *et al.*, Phys. Lett. **B718**, 922 (2013), hep-ex/1202.4064.
- [36] COMPASS Collaboration, C. Adolph *et al.*, sub. Eur. Phys. J. C, hep-ex/1512.05053.
- [37] COMPASS Collaboration, P. Abbon *et al.*, Nucl. Instrum. Meth. **A577**, 455 (2007), hep-ex/0703049.
- [38] G. Ingelman, A. Edin, and J. Rathsmann, Comput. Phys. Commun. **101**, 108 (1997), hep-ph/9605286.
- [39] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjostrand, Phys. Rept. **97**, 31 (1983).
- [40] A. Bacchetta *et al.*, JHEP **02**, 093 (2007), hep-ph/0611265.
- [41] D. Boer, P. J. Mulders, C. Pisano, and J. Zhou, JHEP (2016), hep-ph/1605.07934.
- [42] J. Pretz, Nucl. Instrum. Meth. **A659**, 456 (2011), physics.data-an/1104.1038.
- [43] A. Szabelski, PhD Thesis, National Centre for Nuclear Research, Warsaw 2016.
- [44] R. Sulej, Netmaker, <http://www.ire.pw.edu.pl/~rsulej/NetMaker/>.
- [45] New Muon Collaboration, M. Arneodo *et al.*, Nucl. Phys. **B483**, 3 (1997), hep-ph/9610231.
- [46] A. A. Akhundov, D. Y. Bardin, L. Kalinovskaya, and T. Riemann, Fortsch. Phys. **44**, 373 (1996), hep-ph/9407266.
- [47] A. Ferrari, P. R. Sala, A. Fasso, and J. Ranft, Fluka, <http://www.fluka.org>, 2005.
- [48] M. Bengtsson and T. Sjostrand, Z. Phys. **C37**, 465 (1988).
- [49] CTEQ Collaboration, H. Lai *et al.*, Eur. Phys. J. **C12**, 375 (2000), hep-ph/9903282.
- [50] E143 Collaboration, K. Abe *et al.*, Phys. Lett. **B452**, 194 (1999), hep-ex/9808028.
- [51] H. Fesefeldt, The Simulation of Hadronic Showers: Physics and Applications, 1985.
- [52] I. Akushevich, H. Bottcher, and D. Ryckbosch, in *Monte Carlo generators for HERA physics. Proceedings, Workshop, Hamburg, Germany, 1998-1999*, pp. 554–565, 1998, hep-ph/9906408.
- [53] M. G. A. Buffing, A. Mukherjee, and P. J. Mulders, Phys. Rev. **D88**, 054027 (2013), hep-ph/1306.5897.
- [54] V. Barone, A. Drago, and P. G. Ratcliffe, Phys. Rept. **359**, 1 (2002), hep-ph/0104283.
- [55] COMPASS Collaboration, C. Adolph *et al.*, Phys. Lett. **B717**, 376 (2012), hep-ex/1205.5121.