The spin structure of the nucleon

Christine A. Aidala
Physics Department, University of Michigan, 450 Church Street, Ann Arbor, Michigan 48109-1040, USA

Steven D. Bass
Institute for Theoretical Physics, University of Innsbruck, Technikerstrasse 25, A-6020 Innsbruck, Austria

Delia Hasch
INFN-Frascati, via E. Fermi 40, 00044 Frascati (Rm), Italy

Gerhard K. Mallot
CERN, CH-1211 Genève 23, Switzerland

(published 12 April 2013)

This article reviews our present understanding of QCD spin physics: the proton spin puzzle and new developments aimed at understanding the transverse structure of the nucleon. Present experimental investigations of the nucleon's internal spin structure, the theoretical interpretation of the different measurements, and the open questions and challenges for future investigation are discussed.

DOI: 10.1103/RevModPhys.85.655 PACS numbers: 12.38.-t, 13.88.+e, 14.20.Dh

CONTENTS
I. Introduction 655
II. Spin Structure Functions and Parton Distributions 656
III. Experiments 659
A. SLAC experiments 659
B. CERN experiments 660
1. The EMC and SMC experiments 660
2. The COMPASS experiment 660
C. The HERMES experiment at DESY 661
D. JLab experiments 662
E. Hadronic scattering experiments 662
1. The Relativistic Heavy Ion Collider 662
2. RHIC as a polarized $p + p$ collider 663
3. RHIC experiments 663
IV. The Proton Spin Puzzle 664
A. Spin sum rules 666
B. Proton spin puzzles 667
C. Spin and the singlet axial charge 667
V. Quark and Gluon Polarization from Data 668
A. Valence and sea polarization 668
B. Gluon polarization 670
C. NLO QCD-motivated fits to spin data 671
VI. Theoretical Understanding 673
VII. Transverse Nucleon Structure and Orbital Angular Momentum 675
A. Generalized parton distributions 676
1. Deeply virtual Compton scattering 677
2. The quest for orbital angular momentum and GPD parametrizations 680
B. Transversity, transverse-momentum-dependent distributions, and fragmentation functions 680
1. The Sivers and Boer-Mulders TMD distributions 681
2. The Collins TMD fragmentation function 683
3. Probing transversity 683
4. Current status and recent progress with TMD distributions 684
5. Proton-proton asymmetries and TMD-factorization breaking 684
VIII. Future Projects 685
IX. Conclusions and Outlook 686
Acknowledgments 687
References 687

I. INTRODUCTION

There has been a vigorous and global program of experiments and theoretical developments in the last 25 years aimed at understanding the internal spin structure of the proton. How is the proton's spin built up from the spin and orbital angular momentum (OAM) of the quarks and gluons inside? Tremendous progress has been made with unraveling the proton's spin structure with advances in experimental techniques, theoretical models, perturbative QCD, nonperturbative QCD, and lattice calculations.

This activity was inspired by the initial European Muon Collaboration (EMC) data which suggested the puzzling result that the quark intrinsic spin contributes little to the proton's spin (Ashman et al., 1988). Today there is good convergence of the theoretical and experimental understanding of the proton's longitudinal spin structure. Further...
puzzling data in measurements of transverse single-spin asymmetries revealed up to 40% asymmetries in proton-proton collisions (and 5%–10% in lepton-nucleon collisions with unpolarized leptons and transversely polarized nucleons) which persist to high energies. These single-spin asymmetries indicate significant spin-orbit coupling in the nucleon associated with quark transverse momentum and the bound state structure of the nucleon. The study of transverse momentum and associated orbital angular-momentum processes has spawned new programs to map out the three-dimensional structure of the nucleon. In this article we review these developments highlighting the considerable and exciting developments in QCD spin physics in recent years, together with an outlook to the future: What are the main open questions and the planned experiments to help answer them?

In 1988 EMC published their polarized deep inelastic measurement of the proton’s $g_1$ spin-dependent structure function and the flavor-singlet axial charge $g_A^{(0)}$ (the nucleon’s “quark spin content”) suggesting that quark spins summed over up-, down-, and strange-quark flavors contribute only a small fraction of the proton’s spin. This result inspired considerable theoretical activity and new experiments at CERN, SLAC, DESY, Jefferson Laboratory (JLab), and the Relativistic Heavy Ion Collider (RHIC) at Brookhaven National Laboratory (BNL) to understand the spin structure of the nucleon. The first task was to check the initial curious result from EMC and the second to resolve the spin-flavor structure of the proton. How is the spin content of the proton distributed among the valence and sea quarks and gluons? What about orbital angular momentum in the nucleon?

We now know that the nucleon’s flavor-singlet axial charge measured in polarized deep inelastic scattering is $g_A^{(0)} \sim 0.35$. This value was surprising from the viewpoint of early quark models. In the static quark model (the eigenday way picture of Gell-Mann) before inclusion of quark motion, quark spin contributes 100% of the proton’s spin. Relativistic quark models without gluonic or pion-cloud degrees of freedom generally predict about 60% of the proton’s spin should be carried by the quarks, with the remaining 40% in quark orbital angular momentum. Today data and theory point to a consistent picture where the proton spin puzzle is a valence quark effect. Valence quark contributions to $g_A^{(0)}$ approximately saturate the measured value. While polarized glue may contribute a significant fraction of the proton’s spin (perhaps up to 50% at the scale of present experiments), sea quark and QCD gluon corrections to the singlet axial charge are small and within the expectations of quark models. The pion cloud of the nucleon acts to shift angular momentum from spin to orbital angular momentum and induces SU(3) breaking in the nucleon’s axial charges. There is also a fascinating theoretical possibility that the valence quarks may polarize the QCD vacuum in a nucleon through gluon topological effects so that some fraction of the proton’s singlet axial charge resides at zero parton momentum (or Bjorken $x$). Nonzero orbital angular momentum of the valence quarks is expected, induced also by confinement which introduces a transverse scale in the physics. This orbital angular momentum through spin-orbit coupling is a prime candidate to explain the large single-spin asymmetries observed in proton-proton collisions. Information about quark total angular momentum in the proton can be extracted from deeply virtual Compton scattering (DVCS) and high-energy single-spin asymmetry data in model-dependent analyses. The results are consistent with QCD lattice calculations.

This review is organized as follows. In the first part (Secs. II and III) we give a brief introduction to nucleon spin physics and the experiments that have been performed to investigate it. Then, in Sec. IV, we discuss the proton spin puzzle and the small value of $g_A^{(0)}$ extracted from polarized deep inelastic scattering. In Sec. V we give an overview of the present global program aimed at disentangling the spin-flavor structure of the proton. Section VI covers the theoretical interpretation of longitudinal spin data and understanding of the proton spin puzzle. We next turn our attention to the transverse structure of the nucleon and manifestations of orbital angular momentum in the nucleon in Sec. VII. This discussion introduces generalized parton distributions (GPDs), which describe hard exclusive reaction processes, and transverse-momentum-dependent distributions (TMDs), which describe spin-momentum correlations and spin-orbit couplings in the nucleon. The TMDs are manifest in high-energy single-spin and azimuthal asymmetries. A summary of key issues and challenging questions for the next generation of experiments is given in Secs. VIII and IX.


II. SPIN STRUCTURE FUNCTIONS AND PARTON DISTRIBUTIONS

Our knowledge about the high-energy spin structure of the nucleon comes from both polarized deep inelastic scattering (pDIS) experiments and high-energy polarized proton-proton collisions. pDIS experiments involve scattering a longitudinally polarized high-energy lepton beam from a longitudinally or transversely polarized nucleon at large momentum transfer. Inclusive measurements, where only the scattered lepton is observed in the final state, and semi-inclusive measurements, where one tagg on at least one high-energy final-state hadron in coincidence with the scattered lepton, have been performed. The experiments were performed with an electron beam at SLAC and JLab, with electron and positron beams at DESY and with muon beams at CERN. In proton-proton scattering the protons are either longitudinally or transversely polarized. Polarized deep inelastic scattering experiments have so far all been performed using a fixed target. A future polarized electron-ion collider is in planning. Details of the experiments are given in Sec. III. Historically, information about the proton’s internal spin structure came first from measuring the proton’s $g_1$ and $g_2$ spin structure functions in inclusive deep inelastic scattering and, more recently, from semi-inclusive reactions in both
lepton-nucleon and proton-proton collisions and hard exclusive processes in lepton-nucleon scattering.

Measurements with longitudinally polarized targets and beams tell us about the helicity distributions of quarks and gluons in the nucleon, which at leading order can be thought of as the difference in probability of finding a parton with longitudinal polarization parallel or antiparallel to that of the nucleon. Measurements with transversely polarized targets are particularly sensitive to quark and gluon transverse and orbital angular momentum. Studies of transverse degrees of freedom in the nucleon and in fragmentation processes are a current subject of experimental investigation with sensitivity to spin-orbit couplings in QCD.

For polarized lepton-proton scattering, specialize to the target rest frame and let $E$ denote the energy of the incident lepton which is scattered through an angle $\theta$ to emerge in the final state with energy $E'$. Let $\parallel$ denote the longitudinal polarization of the lepton beam. In photon-nucleon scattering the spin-dependent structure functions $g_1$ and $g_2$ are defined through the imaginary part of the forward Compton scattering amplitude. The structure functions contain all of the target-dependent information in the deep inelastic process. Consider the amplitude for forward scattering of a photon carrying momentum $q$ from a polarized nucleon with momentum $p$, mass $M$, and spin $s$. We work with the kinematic Bjorken variable $x = Q^2/2p \cdot q = Q^2/2M \nu$, where $\nu = p \cdot q/M = E - E'$, and let $y = p \cdot q/p \cdot k = E/E'$. For a longitudinally polarized proton target (with spin denoted $\parallel$) the unpolarized and polarized differential cross sections are

$$
\frac{d^2\sigma}{dx dy} = \frac{2\pi\alpha^2}{MEx^2y^2}\left[(1 - y - \frac{Mxy}{2E})F_2(x, Q^2) + xy^2F_1(x, Q^2)\right]
$$

(1)

and

$$
\frac{d^2\sigma}{dx dy} = \frac{4\alpha^2}{MEx^2y^2}\left[(2 - y - \frac{Mxy}{E})g_1(x, Q^2) - 2\frac{Mx}{E}g_2(x, Q^2)\right].
$$

(2)

where the mass of the lepton is neglected. The relation between the structure functions in deep inelastic lepton-nucleon scattering and the virtual-photon nucleon cross sections is discussed and derived in various textbooks, e.g., Roberts (1990). One finds

$$
A_1 = \frac{\sigma_{1/2} - \sigma_{3/2}}{\sigma_{1/2} + \sigma_{3/2}} = \frac{g_1 - (Q^2/x^2)g_2}{F_1} - \frac{g_1}{F_1},
$$

(3)

where $\sigma_{1/2}$ and $\sigma_{3/2}$ are the cross sections for the absorption of a transversely polarized photon with spin polarized parallel and antiparallel to the spin of the longitudinally polarized nucleon. For a longitudinal polarized target the $g_2$ contribution to the differential cross section and the longitudinal spin asymmetry is suppressed relative to the $g_1$ contribution by the kinematic factor $M/E \ll 1$. For a transverse polarized target this kinematic suppression factor for $g_2$ is missing implying that transverse polarization is vital to measure $g_2$. We refer to Roberts (1990) and Windmolders (2002) for the procedure of how the spin-dependent structure functions are extracted from the spin asymmetries measured in polarized deep inelastic scattering.

In high-$Q^2$ deep inelastic scattering the structure functions $F_1, F_2, g_1$, and $g_2$ exhibit approximate scaling. They are to a very good approximation independent of $Q^2$ and depend only on Bjorken $x$. (The small $Q^2$ dependence which is present in these structure functions is logarithmic and determined by perturbative QCD evolution.)

In the (pre-QCD) parton model the deep inelastic structure functions $F_1$ and $F_2$ are written as

$$
F_1(x) = \frac{1}{2x}F_2(x) = \frac{1}{2} \sum_q e_q^2(q + \bar{q})(x)
$$

(4)

and the polarized structure function $g_1$ is

$$
g_1(x) = \frac{1}{2} \sum_q e_q^2 \Delta q(x).
$$

(5)

Here $e_q$ denotes the electric charge of the struck quark and

$$
\{q + \bar{q}\}(x) = (q^1 + \bar{q}^1)(x) + (q^2 + \bar{q}^2)(x),
$$

$$
\Delta q(x) = (q^1 + \bar{q}^1)(x) - (q^2 + \bar{q}^2)(x)
$$

(6)

denote the spin-independent (unpolarized) and spin-dependent quark parton distributions which measure the distribution of quark momentum and spin in the proton. For example, $\bar{q}^1(x)$ is interpreted as the probability to find an antiquark of flavor $q$ with plus component of momentum $xp_+$ ($p_+ = p_0 + p_3$ is the plus component of the target proton’s momentum) and spin polarized in the same direction as the spin of the target proton. When we integrate out the momentum fraction $x$ the quantity $\Delta q = \int_0^1 dx \Delta q(x)$ is interpreted as the fraction of the proton’s spin which is carried by quarks (and antiquarks) of flavor $q$. Hence summing over the up-, down-, and strange-quark $\Delta q$ contributions gives the total fraction of the proton’s spin carried by the spins of these quarks.

What values should we expect for the $\Delta q$? First consider the static quark model. The simple SU(6) proton wave function

$$
|p\rangle = \frac{1}{\sqrt{2}}|u\rangle|u\rangle,|d\rangle_{S-0}\rangle + \frac{1}{\sqrt{18}}|u\rangle|u\rangle,|d\rangle_{S-1}\rangle
$$

$$
\frac{1}{3}|u\rangle|d\rangle_{S-1}\rangle - \frac{1}{3}|d\rangle|u\rangle_{S-1}\rangle + \sqrt{\frac{1}{3}}|d\rangle|u\rangle_{S-1}\rangle
$$

(7)

yields the values $\Delta u - \Delta d = \frac{2}{3}$ and $\Delta u + \Delta d = 1$. In relativistic quark models one has to take into account the four-component Dirac spinor

$$
\psi \sim \left(\begin{array}{c}
f \\
p_\| \cdot \gamma_\perp
\end{array}\right)
$$

The lower component of the Dirac spinor is a $p$ wave with intrinsic spin primarily pointing in the opposite direction to the spin of the proton (Jaffe and Manohar, 1990). Relativistic effects renormalize the axial charges by the depolarization factor 0.65 with a net transfer of angular momentum from intrinsic spin to orbital angular momentum. In QCD and
more sophisticated models further depolarization is induced by gluonic and pion-cloud degrees of freedom; see Sec. VI.

In QCD the flavor-singlet combination of the \( \Delta q(x) \) quark parton distributions mixes with the spin-dependent gluon distribution under \( Q^2 \) evolution (Altarelli and Parisi, 1977). This spin-dependent gluon distribution measures the momentum and spin dependence of glue in the proton. The second spin structure function \( g_2 \) vanishes without the effect of quark transverse momentum and has a nontrivial parton interpretation (Jaffe, 1990; Roberts, 1990).

The parton model description of polarized deep inelastic scattering involves writing the deep inelastic structure functions as the sum over the convolution of “soft” quark and gluon parton distributions with “hard” photon-parton scattering coefficients

\[
g_f^P(x) = \left\{ \frac{1}{12}(\Delta u - \Delta d) + \frac{1}{36}(\Delta u + \Delta d - 2\Delta s) \right\} \otimes C_{ns}^P + \frac{1}{9}(\Delta u + \Delta d + \Delta s) \otimes C_{t}^P + f \Delta g \otimes C_{ss}^P. \tag{8} \]

Here \( \Delta q(x) \) and \( \Delta g(x) \) denote the polarized quark and gluon parton distributions, \( C_t^P \) and \( C_{ss}^P \) denote the corresponding hard-scattering coefficients, and \( f \) is the number of quark flavors liberated into the final state (\( f = 3 \) below the charm production threshold). The parton distributions contain all the target-dependent information and describe a flux of quark and gluon partons into the (target-independent) interaction between the hard photon and the parton which is described by the coefficients \( C_t^P \) and \( C_{ss}^P \). These coefficients are calculated using perturbative QCD via the cross section for the hard photon scattering from a quark or gluon parton “target.” They are independent of infrared mass singularities (terms involving the quark mass or virtuality of the parton in the photon-parton collision) which are absorbed into the parton distributions (and softened by confinement related physics).

If the same recipe (“factorization scheme”) for separating hard and soft parts of the parton phase space is applied consistently to all hard processes then the factorization theorem asserts that the parton distributions that one extracts from different experiments are process independent. In other words, the same polarized quark and gluon distributions should be obtained from experiments involving polarized hard QCD processes in polarized proton-proton collisions and polarized deep inelastic scattering experiments. For example, colliding longitudinally polarized proton beams provides sensitivity to the gluon-helicity distribution function at leading order. For hadron production with transverse momentum \( p_T \), the helicity-dependent difference in hadron production is defined as

\[
\frac{d\Delta \sigma}{dp_T} = \frac{1}{2} \left[ \frac{d\sigma^{++}}{dp_T} - \frac{d\sigma^{+-}}{dp_T} \right]. \tag{9} \]

Here \( P_a \) and \( P_b \) denote the momenta of the incident protons; \( \Delta f_a(x_a, \mu) \) are the polarized parton distributions of the colliding partons carrying light-cone momentum fraction \( x \) evaluated at factorization and renormalization scale \( \mu \). The helicity-dependent difference in the cross section of the hard partonic scattering \( a + b \rightarrow \text{jet} + X \) is denoted by \( d\Delta \sigma \) and is calculable in perturbative QCD. Partonic cross-section calculations are carried out to finite order in \( \alpha_s \) and have a dependence on factorization and renormalization scales, denoted as \( \mu \). The final hadronic cross section is independent of the factorization and renormalization scales and the scheme used. The QCD parton model treatment readily generalizes to the production of high-energy hadrons in the final state, with the produced “fast” hadron carrying a significant fraction of the momentum of a “parent” parton. The parton-to-hadron process is parametrized by fragmentation functions which also obey process-independent factorization in perturbative QCD calculations.

Analogous to the helicity distributions measured with longitudinal polarization, transversity distributions describe the density of transversely polarized quarks inside a transversely polarized proton; see, e.g., Barone, Drago, and Ratcliffe (2002). The transversity distributions, which were introduced by Ralston and Soper (1979), Arturi and Mekhfi (1990), Cortes, Pire, and Ralston (1992), and Jaffe and Ji (1992) are interpreted in parton language as follows. Consider a nucleon moving with (infinite) momentum in the \( \hat{e}_3 \) direction, but polarized transverse to \( \hat{e}_3 \). Then \( \delta q(x) \) [also denoted \( \Delta T_q(x) \) and \( h_1^T(x) \) in the literature] counts the quarks with flavor \( q \), momentum fraction \( x \), and their spin parallel to the spin of a nucleon minus the number antiparallel. That is, in analogy with Eq. (6), \( \delta q(x) \) measures the distribution of partons with transverse polarization in a transversely polarized nucleon, viz.

\[
\delta q(x) = q^T(x) + q^T(x) - q^T(x) - q^T(x). \tag{11} \]

In a helicity basis transversity corresponds to helicity flip making it a probe of chiral symmetry breaking (Collins, 1993). There is no gluon analog of transversity in the nucleon so \( \delta g \) evolves in \( Q^2 \) like a valence or nonsinglet quark distribution, without mixing with glue. If quarks moved nonrelativistically in the nucleon, \( \delta q \) and \( \Delta g \) would be identical since rotations and Euclidean boosts commute and a series of boosts and rotations can convert a longitudinally polarized nucleon into a transversely polarized nucleon at infinite momentum. The difference between the transversity and helicity distributions reflects the relativistic character of quark motion in the nucleon.

Following the discovery that the quark spin contribution to the proton’s spin is small, there has been a vigorous program to measure the separate contributions of up-, down-, and strange-quark flavors as well as the gluon spin and the orbital contributions. This has inspired dedicated spin programs in semi-inclusive deep inelastic scattering (SIDIS) and polarized proton-proton collisions to measure the separate valence and
A. SLAC experiments

SLAC experiments pioneered polarized DIS measurements and set many standards in polarized beam and target technologies. Their spin program focused on high statistics measurements of the inclusive asymmetries. The first measurements of the proton spin structure were performed by the experiments E80 (Alguird et al., 1976, 1978) and E130 (Baum et al., 1980, 1983), followed by a series of high precision experiments E142 (Anthony et al., 1996), E143 (Abe et al., 1998), E154 (Abe et al., 1997), and E155 (Anthony et al., 1999, 2000) a decade later. These experiments utilized polarized electrons which were produced by laser photoemission and subsequently accelerated. The longitudinal polarization of the beam was frequently inverted and the polarization measured using Møller scattering. A rapid cycling of the beam and/or target polarization reduces systematic uncertainties in the measured spin asymmetries related to the stability of the experimental setup. Polarized target materials involved solid-state butanol and ammonia (NH₃) for the proton and D-butanol, ND₃ as well as ⁶LiD for the deuteron (Crabb and Meyer, 1997; Meyer, 2004). For the most recent E154 and E155 experiments the target polarization was typically 38% for ³He, 90% for NH₃, and 22% for LiD with beam polarization about 80%. The target material, doped with a paramagnetic substance or irradiated with electron beams, was polarized using dynamic nuclear polarization, which requires temperatures of about 1 K and strong magnetic holding fields. Such targets contain a considerable amount of nonpolarizable nucleons, which is parameterized by the so-called dilution factor. This factor depends on all kinematic variables relevant for the process under study and needs, in principle, to be determined for each type of measurement. Typical values for polarized solid-state targets range between 0.1 and 0.2 with the exception of ⁶LiD (0.4–0.5) and represent an important factor in the extraction of physical observables from the measured ones. Information on the neutron structure was obtained either from the

III. EXPERIMENTS

Experiments that have probed the nucleon spin structure are outlined in Table I. This includes both polarized deep inelastic lepton-nucleon scattering and proton-proton collision experiments. Considerable effort was invested in developing polarized beam and target technology, yielding physics results with ever increasing precision. The first experiments focused on inclusive deep inelastic measurements of nucleon spin structure. More recent experiments, described in detail below, were able to detect and identify hadrons in the final state leading to new probes of the nucleon in semi-inclusive and hard exclusive reactions. Future experimental programs (COMPASS-II, the 12 GeV upgrade of JLab, and experiments at Fermilab and RHIC) with high luminosity and acceptance are planned to explore the three-dimensional structure of the nucleon in spatial and transverse-momentum degrees of freedom. We discuss these future programs in Sec. VIII.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Year</th>
<th>Beam</th>
<th>Target</th>
<th>Energy (GeV)</th>
<th>Q² (GeV²)</th>
<th>x</th>
</tr>
</thead>
<tbody>
<tr>
<td>SLAC: E80, E130</td>
<td>1976–1983</td>
<td>e⁻</td>
<td>H-butanol</td>
<td>≤ 23</td>
<td>1–10</td>
<td>0.1–0.6</td>
</tr>
<tr>
<td>SLAC: E142/3</td>
<td>1992–1993</td>
<td>e⁻</td>
<td>NH₃, ND₃</td>
<td>≤ 30</td>
<td>1–10</td>
<td>0.03–0.8</td>
</tr>
<tr>
<td>SLAC: E154/5</td>
<td>1995–1999</td>
<td>e⁻</td>
<td>NH₃, ⁶LiD, ³He</td>
<td>≤ 50</td>
<td>1–35</td>
<td>0.01–0.8</td>
</tr>
<tr>
<td>CERN: EMC</td>
<td>1985</td>
<td>μ⁺</td>
<td>NH₃, ⁶LiD, ³He</td>
<td>100, 190</td>
<td>1–30</td>
<td>0.01–0.5</td>
</tr>
<tr>
<td>CERN: SMC</td>
<td>1992–1996</td>
<td>μ⁺</td>
<td>H/D-butanol, NH₃</td>
<td>100, 190</td>
<td>1–60</td>
<td>0.004–0.5</td>
</tr>
<tr>
<td>FNAL: E581/E704</td>
<td>1988–1997</td>
<td>p</td>
<td>p</td>
<td>200</td>
<td>~1</td>
<td>0.1 &lt; x_F &lt; 0.8</td>
</tr>
<tr>
<td>DESY: HERMES</td>
<td>1995–2007</td>
<td>e⁺, e⁻</td>
<td>H, D, ³He</td>
<td>~30</td>
<td>1–15</td>
<td>0.02–0.7</td>
</tr>
<tr>
<td>CERN: COMPASS</td>
<td>2002–2012</td>
<td>μ⁺</td>
<td>NH₃, ⁶LiD</td>
<td>160, 200</td>
<td>1–70</td>
<td>0.003–0.6</td>
</tr>
<tr>
<td>JLab6: Hall A</td>
<td>1999–2012</td>
<td>e⁻</td>
<td>³He</td>
<td>≤ 6</td>
<td>1–2.5</td>
<td>0.1–0.6</td>
</tr>
<tr>
<td>JLab6: Hall B</td>
<td>1999–2012</td>
<td>e⁻</td>
<td>NH₃, ND₃</td>
<td>≤ 6</td>
<td>1–5</td>
<td>0.05–0.6</td>
</tr>
<tr>
<td>RHIC: BRAHMS</td>
<td>2002–2006</td>
<td>p</td>
<td>p (beam)</td>
<td>2 × (31–100)</td>
<td>~1–6</td>
<td>~0.6 &lt; x_p &lt; 0.6</td>
</tr>
<tr>
<td>RHIC: PHENIX, STAR</td>
<td>2002+</td>
<td>p</td>
<td>p (beam)</td>
<td>2 × (31–250)</td>
<td>~1–400</td>
<td>~0.02–0.4</td>
</tr>
<tr>
<td>CERN: COMPASS-II</td>
<td>2014+</td>
<td>μ⁺, μ⁻</td>
<td>Unpolarized H₂</td>
<td>160</td>
<td>~1–15</td>
<td>~0.005–0.2</td>
</tr>
<tr>
<td>JLab12: Halls A/B/C</td>
<td>2014+</td>
<td>e⁻</td>
<td>HD, NH₃, ND₃, ³He</td>
<td>≤ 12</td>
<td>~1–10</td>
<td>~0.05–0.8</td>
</tr>
</tbody>
</table>

Table I. High-energy spin experiments: the kinematic ranges in x and Q² correspond to the average kinematic values of the highest statistics measurement of each experiment, which is typically the inclusive spin asymmetry; x denotes Bjorken x unless specified.
combination of measurements with proton and deuteron targets or by using a polarized $^3$He target which is dominated by the neutron since the two proton spins in $^3$He are antialigned. Here polarization was obtained from optical pumping and adiabatic spin exchange. The target polarization was measured using the NMR technique. Scattered electrons were detected with magnetic spectrometers optimized for high-momentum resolution and good electron identification.

B. CERN experiments

1. The EMC and SMC experiments

Following the early measurements at SLAC, the EMC experiment performed at CERN in 1985 the first polarized DIS measurements at $x < 0.1$ down to $x = 0.01$ after a series of measurements of unpolarized nucleon and nuclear structure functions. The experiment used the polarized CERN muon beam up to momenta of 200 GeV and a solid-state irradiated ammonia target. Their low-$x$ measurements, accessible due to the high energy of the muons, suggested the breakdown of the naive parton picture that quarks provide essentially all of the spin of the nucleon (Ashman et al., 1988, 1989).

This triggered more detailed and precise measurements by the Spin Muon Collaboration (SMC) in 1992–1996, and by COMPASS (since 2002). The beam line and the principal ideas of the CERN muon experiments are described in the COMPASS Sec. III.B.2. The EMC spectrometer is described in Aubert et al. (1981). The polarization of the CERN muon beam was measured by SMC (Adeva et al., 1994b). A detailed description of the SMC deuteron target polarization is given in Adeva et al. (1994a). The COMPASS experiment used the SMC target in the initial period of data taking up to 2005 as reported by Ball et al. (2003). A new target has been used since 2006 (Gautheron, 2007).

After 1987 the focus was on the region $x < 0.1$ and the flavor-singlet axial charge (Ellis-Jaffe sum rule) for the neutron. The latter must deviate from the naive prediction in a similar way as for the proton in order to preserve the fundamental isovector Bjorken sum rule for $g_1^n - g_1^p$. (These sum rules are discussed later.) The SMC experiment extended the measurement $x$ range down to $x = 0.004$ (for $Q^2 > 1$ GeV$^2$) and established the validity of the Bjorken $x$ sum rule with measurements using polarized proton (butanol and ammonia) and deuteron (D-butanol) targets (Adeva et al., 1993, 1998b). The large acceptance of the SMC spectrometer in the forward direction allowed them to present the first determination of individual quark distributions for different flavors (Adeva et al., 1996, 1998a) from semi-inclusive DIS. A dedicated polarimeter confirmed the validity of the beam polarization obtained from Monte Carlo simulations (Adeva et al., 1994b; Adams et al., 2000) used in the EMC, SMC, and COMPASS analyses.

2. The COMPASS experiment

The COMPASS spectrometer (Abbon et al., 2007; see Fig. 1) is installed at the muon beam line of the CERN Super Proton Synchrotron accelerator. A polarized muon beam of energy 160–200 GeV and with a polarization of about 80% impinges on a solid-state polarized target consisting of two or three cells with proton or deuteron target material polarized in opposite directions. The usable beam intensity is typically $2 \times 10^7$/s during a 9.6 s long spill. The repetition rate varies and is typically about 1/40 s. The muon polarization arises naturally from the weak decay of the parent pions produced by the primary proton beam of 400 GeV. The momentum of each beam muon is measured in the beam momentum station. Downstream of the target, the scattered muon and produced hadrons are detected in a two-stage magnetic spectrometer with the two dipole magnets (SM1, SM2).

Charged particles are tracked in the beam regions by scintillating fiber stations (SciFi) and by silicon detectors. In the inner region close to the beam, gaseous detectors of the micromegas and gas-electron-multiplier (Gem) types with high rate capabilities are deployed. The backbone of tracking in the intermediate region is multiwire proportional chambers (MWPCs). Finally, the large area tracking away from the beam region is covered by drift chambers (DC, W45) and drift tubes ([Straws, Rich Wall (RW), Muon Wall (MW)].

The velocity of charged particles is measured in a ring-imaging Cherenkov detector (RICH), which can separate pions and kaons from 9 up to 50 GeV. The inner quarter of the photon detector is made of multianode-photomultiplier tubes, while the outer part relies on MWPCs with a photosensitive CsI cathode.

The energy of charged particles is measured in sampling hadron calorimeters (HCAL), while neutral particles, in particular, high-energy photons, are detected in electromagnetic calorimeters (ECAL). They comprise lead-glass modules as well as scintillator-lead “shashlik” modules in the inner high-radiation region.

Event recording is triggered by the scattered muon, which is “identified” by its ability to traverse thick hadron absorbers, located just upstream of the MW detectors. The event selection is based on various systems of scintillator hodoscopes and logic modules applying selection criteria such as target pointing and energy loss in the scattering. The patterns causing a trigger were optimized by Monte Carlo simulations. The spectrometer has about 250,000 readout channels, which can be recorded with a frequency of 20 kHz for an event size of the order of 40 kbyte.

The heart of the experiment is the polarized target system. While the muon beam comes naturally polarized due to the parity violation in the decay of the parent pions, polarizing
protons and deuterons is very difficult. Gas targets cannot be used with the muon beam due to the low beam intensity compared to electron beams. An advantage of muon beams is the high muon energy, which presently cannot be reached by electron beams. The polarized target system comprises a 2.5 T solenoid magnet, a 0.6 T dipole magnet, a \(^3\)He/\(^4\)He dilution refrigerator, a 70 GHz microwave system, and an NMR system to measure the target polarization. The target material is cooled down to about 60 mK in frozen spin mode. The nucleons/nuclei are polarized by dynamic nuclear polarization which is applicable only for particular materials. In COMPASS irradiated ammonia (NH\(_3\)) and lithium-6 deuteride (\(^6\)LiD) were selected as proton and deuteron targets, respectively. Lithium-6 is very close to a system of a free deuteron and a helium-4 core and has essentially the same magnetic moment as the deuteron. Thus \(^6\)LiD corresponds to two deuterons plus a helium nucleus. Typically, polarizations of 85\% for protons and 50\% for deuterons were reached. A key feature of COMPASS is that both target polarizations are present simultaneously in separate target cells along the beam, e.g., “→, →” for the two-cell configuration until 2004 and “→, →, →” for the three-cell configuration from 2006 onward. In the former configuration the length of the cells was twice 60 cm while in the latter it was 30, 60, 30 cm, respectively. Thus in an asymmetry measurement most systematic uncertainties cancel. Using the dipole and solenoid magnet, the magnetic field can be rotated from, e.g., pointing downstream to transverse to upstream. The spin follows the magnetic field adiabatically and thus the spin orientations can be changed within 30 min. Such a field rotation is performed typically once per day for the longitudinal polarization in order to cancel potentially remaining systematic effects. The field can also be kept transverse for measurements with transverse target polarization. Here the polarization is inverted by repolarizing typically once per week. In the shutdown year 2005 the superconducting target magnet was replaced by a new one, increasing the angular acceptance from ±70 to ±180 mrad.

The experiment is taking data since 2002. The main focus has been on inclusive and semi-inclusive polarized deep inelastic scattering. As schematically depicted in Fig. 2, the detection of a hadron in the final state provides information about the flavor of the struck quark, while the kinematics of the DIS event is fixed by the incoming and scattered lepton. The years 2008 and 2009 were dedicated to the hadron spectroscopy program of COMPASS with pion, kaon, and proton beams. In 2012 the pion polarizability was measured using a negative pion beam and a thin nickel target. A pilot run for deeply virtual Compton scattering and hard exclusive meson production was successfully completed in 2012.

C. The HERMES experiment at DESY

The HERMES experiment employed an innovative technique for the polarized target, which is very different from all other polarized DIS experiments. Gas targets of pure nuclear-polarized atoms of hydrogen or deuterium were used, which permit essentially background-free measurements from highly polarized nucleons with little or no dilution of the signal from unpolarized nucleons in the target. This choice eliminates one of the main systematic sources in polarized DIS, the uncertainty in the determination of the dilution factor.

The HERMES gas targets were highly longitudinally (~ 85\%) or transversely (75\%) polarized with the ability to invert the direction of the spin of the nucleons within milliseconds. Because of the low densities, however, such targets are practicable only in the high currents of storage rings. HERMES was operating from 1995 to 2007 at the Hadron Electron Ring Accelerator (HERA) lepton storage ring, which provided electron or positron beams of typically 40 mA and with an energy of 27.5 GeV. In order to enhance the target density, the novel technique of a storage cell was used (Baumgarten et al., 2002, 2003a, 2003c; Airapetian et al., 2004b). Here the gas was fed into a T-shaped open-ended elliptical cell coaxial to the lepton beam. The gas atoms underwent several hundred wall bounces before escaping from the ends where they were differentially pumped away by a large system of turbopumps. This increased the density by a factor of about 100 compared to free gas jet targets.

The polarized atoms were injected into the cell from an atomic beam source based on Stern-Gerlach polarization filtering and radio-frequency transitions between atomic substates in a magnetic field (Airapetian et al., 2005c). A small sample gas diffused from the middle of the cell into a Breit-Rabi polarimeter which measured the atomic polarization (Baumgarten et al., 2002) or into a target gas analyzer which measured the atomic and the molecular content of the sample (Baumgarten et al., 2003b). A magnet surrounding the storage cell provided a holding field defining the polarization axis and prevented spin relaxation via spin exchange or wall collisions. The cell temperature was kept at about 100 K, the value for which atomic recombination and spin relaxation during wall collisions are minimal.

Stored high-energy electron beams may become spontaneously transversely polarized via a small polarization asymmetry in the emission of synchrotron radiation by the beam particles as they are deflected by the magnetic fields of the ring (Sokolov-Ternov effect) (Sokolov and Ternov, 1964). The beam polarization grows and asymptotically approaches an equilibrium value with a time constant depending on the characteristics of the ring, for HERA typically one-half hour.
Polarizations as large as 60% were achieved. Spin rotators and polarimeters were essential components of the HERA lepton beam (Buon and Steffen, 1986; Barber et al., 1993, 1994; Beckmann et al., 2002). Spin rotators in front of and behind the experiment provided longitudinal polarization at the interaction point and at one of the two beam polarimeters. The two beam polarimeters were based on Compton backscattering of circularly polarized laser light. They continuously monitored the transverse and longitudinal polarization of the lepton beam.

The HERMES spectrometer was designed to detect the scattered lepton and produced hadrons within a wide angular acceptance and with good momentum resolution. Particular emphasis was given to the particle identification capabilities which allowed for pion, kaon, and proton separation over almost the whole momentum range (Akopov et al., 2002). The HERA beam lines passed through the noninstrumented horizontal midplane of the spectrometer. A horizontal iron plate shielded the beam lines from the 1.5 Tm dipole field of the spectrometer magnet, thus dividing the spectrometer into two identical halves. The geometrical acceptance of \( \pm 170 \text{ mrad} \) horizontally and \( \pm (40\text{–}140) \text{ mrad} \) vertically resulted in detected scattering angles ranging from 40 to 220 mrad. Tracking was provided by several stages of drift chambers before and after the spectrometer magnet. The combination of signals from a lead-glass calorimeter, a preshower detector, a transition radiation detector, and a ring-imaging Cherenkov provided lepton identification with very high efficiency and purity better than 99% as well as pion, kaon, and proton separation over almost the whole momentum range of 2–15 GeV. All components are described in detail in Ackerstaff et al. (1999a).

A recoil detector was installed in the target region for the last 1.5 years of HERMES data taking with unpolarized hydrogen and deuterium targets in order to enhance access to hard exclusive processes, in particular, to deeply virtual Compton scattering.

**D. JLab experiments**

Experiments at Jefferson Laboratory utilized the highest polarization electron beams (85%) with energy ranging from 0.8 close to 6 GeV. The technologies of polarizing beam and target follow those pioneered and further developed at SLAC.

The beam was provided by the Continuous Electron Beam Accelerator Facility (CEBAF) (Leemann, Douglas, and Krafft, 2001), which used polarized electron guns based on a "super-lattice" of a thin gallium arsenide (GaAs) layer on top of GaAs-phospide bulk matter illuminated by circularly polarized photons from high intensity lasers (Sinclair et al., 2007; Stutzman et al., 2007). Subsequently, the polarized electrons passed up to 5 times the two linear accelerators based on superconducting radio-frequency technology and connected by two recirculation arcs. The spin direction of the electrons was manipulated using the crossed electric and magnetic fields of Wien filters, which allow for rapid spin rotation. Their direction was inverted about every 30 ms. Beam polarimetry was employed at several stages of the acceleration process. CEBAF delivered polarized beams simultaneously to the three experimental halls (Halls A, B, and C) with the option to independently dial the energy and intensity. Typical beam intensities ranged from a few nA in Hall B to over 100 \( \mu \text{A} \) in the other two halls (Kazimi et al., 2004).

Longitudinal polarized solid-state ammonia (\( \text{NH}_3 \)) targets for the proton and \( \text{ND}_3 \) for the deuteron were employed at Hall B (Keith et al., 2003). These targets are based on similar techniques as discussed before for the SLAC and CERN experiments for both polarization and polarimetry. Hall A used a polarized \( ^3\text{He} \) target. The target polarization was measured by both the NMR technique of adiabatic fast passage and a technique based on electron paramagnetic resonance (Romalis et al., 1998). Average target polarizations of about 55% were obtained.

Halls A and C were both instrumented with small acceptance but high resolution spectrometers that could cope with the highest beam intensities but measured at fixed scattering angles. These spectrometers are equipped for high resolution tracking, precise time-of-flight measurements, and lepton and hadron separation (Alcorn et al., 2004).

Hall B was instrumented with the CEBAF large acceptance spectrometer (CLAS) (Mecking et al., 2003). The CLAS design was based on a toroidal field, generated by six superconducting coils arranged around the beam line. The six coils naturally divided the detector into six independent spectrometers, each of them containing a set of drift chambers for tracking, a gas Cerenkov counter for electron and pion separation, an array of scintillator counters for particle identification using time-of-flight measurements, and electromagnetic calorimeters for neutral particle identification. For charged particles, CLAS covered polar angles between \( 8^\circ \) and \( 142^\circ \) in the laboratory frame and between 60% and 80% of the azimuthal angles.

**E. Hadronic scattering experiments**

While deep inelastic lepton-nucleon scattering has long been a standard tool of the trade in the study of unpolarized and polarized nucleon structure, much has been learned from polarized hadronic scattering as well. The first high-energy primary polarized proton beams were achieved at the Zero-Gradient Synchrotron at Argonne National Laboratory in 1973. Proton beams there were initially accelerated to 6 GeV with a polarization of about 60%, and shortly thereafter polarized beams up to 12 GeV were achieved. In the 1990s at Fermilab, secondary beams of polarized protons or antiprotons from lambda or antilambda decays opened up new kinematic regions for polarized hadronic scattering, with polarized beams of up to 200 GeV (\( \sqrt{s} = 19 \text{ GeV} \)). Polarized hadronic scattering experiments at center-of-mass energies more than an order of magnitude higher were achieved with the inauguration of the BNL Relativistic Heavy Ion Collider for polarized protons in 2001.

1. **The Relativistic Heavy Ion Collider**

RHIC is located at Brookhaven National Laboratory in New York. RHIC was built to collide heavy ions at center-of-mass energies of up to 200 GeV per colliding nucleon pair and polarized protons at center-of-mass energies ranging from 50 to 500 GeV. Collision of asymmetric species, i.e., different species in the two beams, is also possible due to
independent rings with independent steering magnets. The first polarized proton collisions were achieved at a center-of-mass energy of 200 GeV in December 2001.

The RHIC storage ring is 3.83 km in circumference and is designed with six interaction points (IPs) at which beam collisions are possible. Up to 112 particle bunches per ring can be injected, in which case the time between bunch crossings at the IPs is 106 ns. Polarizations of up to 65% for 100 GeV proton beams and about 60% for 255 GeV beams have been achieved. The maximum luminosities achieved thus far are $5 \times 10^{31}$ cm$^{-2}$s$^{-1}$ at $\sqrt{s} = 200$ GeV and $2 \times 10^{32}$ cm$^{-2}$s$^{-1}$ at $\sqrt{s} = 510$ GeV.

Three experiments have studied polarized proton collisions at RHIC. There are two ongoing large experiments, STAR (Ackermann et al., 2003) and PHENIX (Adcox et al., 2003), each of which have more than 500 collaborators total working on both the heavy ion and polarized proton programs, and the smaller BRAHMS (Adamczyk et al., 2003) experiment, with fewer than 100 collaborators, which took data through 2006. In addition to the program of proton spin structure measurements, the transverse single-spin asymmetry in elastic proton-proton scattering has also been measured to constrain the hadronic spin-flip amplitude in this reaction (Adamczyk et al., 2012b).

**2. RHIC as a polarized p + p collider**

RHIC is the first and only high-energy polarized proton-proton collider in the world. A number of technological developments and advances over the past several decades have made it possible to create a high-current polarized proton source, maintain the beam polarization throughout acceleration and storage, and obtain accurate measurements of the degree of beam polarization at various stages of acceleration in order to identify and address possible origins of depolarization at each step. Only RHIC polarimetry will be discussed here. There are two types of polarimeters installed in RHIC. The fast proton-carbon ($pC$) polarimeter (Nakagawa et al., 2008) takes advantage of a known analyzing power $A_{pC}^p = 0.01$ in the elastic scattering of polarized protons with carbon atoms ($p^1 + C \rightarrow p^1 + C$), which originates from interference between electromagnetic and hadronic elastic scattering amplitudes. The $pC$ polarimeter can make measurements in less than 10 s and provide immediate information on the stability or decay of the beam polarization from a few data points taken over the several hours of a fill. Calibration of the $pC$ polarimeter to within an absolute beam polarization of less than 5% can then be provided by measuring polarized elastic $p + p$ scattering with a polarized hydrogen-jet-target polarimeter (Zelenksi et al., 2005). With the hydrogen-jet-target polarization of greater than 90% known to better than 2% in absolute polarization (Okada et al., 2006), the absolute beam polarization can be determined by exploiting the symmetry of the process.

The naturally stable spin direction through acceleration and storage in RHIC is transverse to the proton’s momentum, in the vertical direction. Spin rotator dipole magnets have been used to achieve both radial and longitudinal spin (MacKay et al., 2003). The rotators are located outside the interaction regions of the PHENIX and STAR experiments, giving both experiments the ability to choose independently whether they want longitudinally or transversely polarized collisions. The BRAHMS experiment, having no spin rotators available, focused on transverse spin measurements. The local nature of the spin rotator magnets means that the STAR and PHENIX experiments must each have their own way of checking the direction of the spin vector at their respective interaction regions.

Observed azimuthal transverse single-spin asymmetries in the production of forward neutrons (Bazilevsky et al., 2003) and forward charged particles can be used to provide local polarimetry. These asymmetries are exploited by the experiments to measure the degree to which the beam polarization is vertically transverse, radially transverse, or longitudinal. More information on local polarimetry at PHENIX and STAR can be found in Kiryluk (2005) and Adare et al. (2007).

**3. RHIC experiments**

a. The PHENIX detector

PHENIX was designed as a large, multipurpose experiment with fast data acquisition and high granularity over a limited acceptance. See Fig. 3 for beam and side views of the PHENIX detector as configured for data taking in 2012. There are two central arms with an acceptance covering a pseudorapidity range $|\eta| < 0.35$ and $\Delta \phi = \pi/2$ each in azimuth.
The central arms include drift and pad chambers (DC, PC1, PC2, and PC3) for momentum and position measurements, a RICH detector primarily for electron identification, small-acceptance time of flight, and aerogel counters (TOF-E, TOF-W, aerogel) for charged hadron identification, and electromagnetic calorimetry (PbSc, PbGl). Electronics-level triggering in the central arms uses information from the calorimetry and ring-imaging Cherenkov detector. There are two muon spectrometers covering a pseudorapidity of $<1.2 < |\eta| < 2.4$, consisting of tracking chambers and muon identifier panels (MuTr, MuID). Resistive plate chambers (RPC3) were added in 2011 and 2012 to improve triggering on high-momentum muons for $W$ boson measurements. Forward electromagnetic calorimetry [Muon Piston Calorimeter (MPC)] covering $3.0 < |\eta| < 3.9$ was added in 2006 and 2007, and silicon vertex detectors [Forward Silicon Vertex Tracker (F)VTX] for heavy flavor measurements over $|\eta| < 2.4$ were added in 2011 and 2012.

For luminosity measurements, identical zero-degree hadronic calorimeters (ZDC) are located in the RHIC tunnel at $\pm 18$ m from the nominal IP for all RHIC experiments. PHENIX also uses quartz Cherenkov beam-beam counters (BBC) positioned around the beam pipe at $\pm 1.44$ m from the nominal interaction point as a minimum-bias trigger detector and for polarization-averaged as well as spin-dependent luminosity measurements. Collision rates for 500 GeV $p+p$ running reach $\sim 3$ MHz, and the electronics-level triggers select events to reduce this rate to approximately 7 kHz of recorded data.

b. The STAR detector

The Solenoidal Tracker at RHIC (STAR) was designed as a large, multipurpose detector with wide acceptance, making it well suited for correlation measurements. The core of STAR is a time-projection chamber, which covers $2\pi$ in azimuth and has tracking capabilities over $|\eta| < 1.3$ and good particle identification for $|\eta| < 1$. There is electromagnetic calorimetry for $-1 < \eta < 2$. In the forward direction, there is additional electromagnetic calorimetry for $2.5 < \eta < 4.0$. Recent upgrades include a time-of-flight detector with 100 ps resolution for additional particle identification, and tracking based on Gem detectors for $1 < \eta < 2$ was partially installed for 2012 data taking to enable charge-sign discrimination of forward electrons from $W$ boson decays.

In addition to the zero-degree hadronic calorimeters identical among the RHIC experiments, STAR has scintillator beam-beam counters positioned around the beam pipe covering $3.4 < |\eta| < 5.0$, which provide a minimum-bias trigger as well as spin-averaged and spin-dependent luminosity measurements along with the ZDCs.

c. The BRAHMS detector

The BRAHMS detector was a smaller experiment at RHIC designed for excellent momentum measurement and charged particle identification over a very broad range of rapidities. It consisted of two movable spectrometer arms covering small solid angles, the forward spectrometer, which could be positioned as close as $2.3^\circ$ from the beam pipe, and the midrapidity spectrometer, which could be moved to cover an angular range from $30^\circ < \theta < 95^\circ$. The spectrometer arms included five dipole magnets, time-projection chambers, multiwire drift chambers, time-of-flight hodoscopes, and threshold as well as ring-imaging Cherenkov detectors. Global detectors consisted of a silicon array for multiplicity measurements, threshold Cherenkov beam-beam counters for event vertex and timing determination as well as luminosity measurements, and ZDCs identical to those used by PHENIX and STAR.

IV. THE PROTON SPIN PUZZLE

We begin our discussion of physics results by first describing how the small value of the quark spin content $g_A^{(0)}$ is obtained from polarized deep inelastic scattering and the first moment of the $g_1$ spin structure function.

In QCD the first moment of $g_1$ is determined from the dispersion relation for polarized photon-nucleon scattering and the light-cone operator product expansion. One finds that the first moment of $g_1$ is related to the scale-invariant axial charges of the target nucleon by

$$\int_0^1 dx g_1^\nu(x, Q^2) = \left( \frac{1}{12} g_A^{(3)} + \frac{1}{36} g_A^{(8)} \right) \left[ 1 + \sum_{\ell \geq 1} c_{\ell S} a_\ell^I(Q) \right] + \frac{1}{9} g_A^{(0)} |_{\text{inv}} \left[ 1 + \sum_{\ell \geq 1} c_{\ell S} a_\ell^I(Q) \right] + O \left( \frac{1}{Q^2} \right) + \beta_\omega. \quad (12)$$

Here $g_A^{(3)}$, $g_A^{(8)}$, and $g_A^{(0)} |_{\text{inv}}$ are the isovector, SU(3) octet, and scale-invariant flavor-singlet axial charges, respectively. The
flavor nonsinglet $c_{NSF}$ and singlet $c_{S}$ Wilson coefficients are calculable in $\ell$-loop perturbative QCD. These perturbative QCD coefficients have been calculated to $O(\alpha_s^\ell)$ precision (Larin, van Ritbergen, and Vermaseren, 1997). For $\alpha_s = 0.3$ typical of the deep inelastic experiments one finds $\{1 + \sum_{i=1}^{\ell} c_{NSF}^i \alpha_s^i(Q)\} = 0.85$ and $\{1 + \sum_{i=1}^{\ell} c_{S}^i \alpha_s^i(Q)\} = 0.96$. The term $\beta_\infty$ represents a possible leading-twist subtraction constant from the circle at infinity when one closes the contour in the complex plane in the dispersion relation (Bass, 2005). The subtraction constant affects just the first moment and corresponds to a contribution at Bjorken $x$ equal to zero.

In terms of the flavor-dependent axial charges

$$2M s_{\mu} \Delta q = \langle p, s | \bar{s} \gamma_\mu \gamma_5 q | p, s \rangle,$$

(13)

the isovector, octet, and singlet axial charges are

$$g_A^{(3)} = \Delta u - \Delta d, \quad g_A^{(8)} = \Delta u + \Delta d - 2\Delta s, \quad g_A^{(0)}_{\text{inv}}/E(\alpha_s) = g_A^{(0)} = \Delta u + \Delta d + \Delta s.$$  

(14)

Here

$$E(\alpha_s) = \exp \int_0^{\alpha_s} d\tilde{\alpha}_s \gamma(\tilde{\alpha}_s)/\beta(\tilde{\alpha}_s)$$  

(15)

is a renormalization group factor which corrects for the (two-loop) nonzero anomalous dimension $\gamma(\alpha_s)$ of the singlet axial-vector current

$$J_{\mu 5} = \bar{u} \gamma_\mu \gamma_5 u + \bar{d} \gamma_\mu \gamma_5 d + \bar{s} \gamma_\mu \gamma_5 s$$  

(16)

which is close to 1 and which goes to 1 in the limit $Q^2 \to \infty$. The symbol $\beta$ denotes the QCD beta function $\beta(\alpha_s) = -\left(1 - \frac{1}{2} f(\alpha_s^2/2\pi) + \cdots \right)$ and $\gamma$ is given by $\gamma(\alpha_s) = f(\alpha_s^2/\pi^2) + \cdots$ where $f = 3$ is the number of active flavors (Kodaira, 1980). The singlet axial charge $g_A^{(0)}_{\text{inv}}$ is independent of the renormalization scale $\mu$ and corresponds to $g_A^{(0)}(Q^2)$ evaluated in the limit $Q^2 \to \infty$. The flavor nonsinglet axial charges $g_A^{(3)}$ and $g_A^{(8)}$ are renormalization group invariants. We are free to choose the QCD coupling $\alpha_s(\mu)$ at either a hard or a soft scale $\mu$. The perturbative QCD expansion of $E(\alpha_s)$ remains close to 1—even for large values of $\alpha_s$. If we take $\alpha_s \sim 0.6$ as typical of the infrared then $E(\alpha_s) \approx 1 - 0.13 - 0.03 + \cdots = 0.84 + \cdots$ where $-0.13$ and $-0.03$ are the $O(\alpha_s)$ and $O(\alpha_s^2)$ corrections, respectively.

In the naive parton model $g_A^{(0)}$ is interpreted as the fraction of the proton’s spin which is carried by the intrinsic spin of its quark and antiquark constituents. The experimental value of $g_A^{(0)}$ is obtained through measuring $g_1$ and combining the first moment integral in Eq. (12) with knowledge of $g_A^{(3)}$ and $g_A^{(8)}$ from other processes plus theoretical calculation of the perturbative QCD Wilson coefficients.

The isovector axial charge is measured independently in neutron $\beta$ decays [$g_A^{(3)} = 1.270 \pm 0.003$ (Beringer et al., 2012)] and the octet axial charge is commonly taken to be the value extracted from hyperon $\beta$ decays assuming a two-parameter SU(3) fit [$g_A^{(8)} = 0.58 \pm 0.03$ (Close and Roberts, 1993)]. However, it should be noted that the uncertainty quoted for $g_A^{(8)}$ has been a matter of some debate (Jaffe and Manohar, 1990; Ratcliffe, 2004). SU(3) symmetry may be badly broken and some have suggested that the error on $g_A^{(8)}$ should be as large as 25% (Jaffe and Manohar, 1990). A recent reevaluation of the nucleon’s axial charges in the cloudy bag model taking into account the effect of the one-gluon-exchange (OGE) hyperfine interaction and the pion cloud plus kaon loops led to the value $g_A^{(8)} = 0.46 \pm 0.05$ (Bass and Thomas, 2010). The model reduction of $g_A^{(8)}$ from the SU(3) value comes primarily from the pion cloud with $g_A^{(3)}$ taking its physical value.

Deep inelastic measurements of $g_1$ have been performed in experiments at CERN, DESY, JLab, and SLAC. An overview of the world data on the nucleon’s $g_1$ spin structure function is shown in Fig. 4. These data are published in EMC (Ashman et al., 1989), SMC (Adeva et al., 1998), E142 (Anthony et al., 1996), E143 (Abe et al., 1998), E154 (Abe et al., 1997), E155 (Anthony et al., 2000), E155 (Anthony et al., 1999), HERMES (Airapetian et al., 2007), JLab (Zheng

![FIG. 4. World data on $xg_1$ as a function of $x$ for the proton (top), the deuteron (middle), and the neutron (bottom) at the $Q^2$ of the measurement. Only data points for $Q^2 > 1$ GeV$^2$ and $W > 2.5$ GeV are shown. Error bars are statistical errors only.](image-url)
et al., 2004; Dharmawardane et al., 2006), and COMPASS (Alexakhin et al., 2007; Alekseev et al., 2010d). There is a general consistency among all data sets. The kinematic reach of the different experiments is visible in Fig. 5. COMPASS have the smallest-\(x\) data, down to \(x \sim 0.004\).

There are several striking features in the data. COMPASS measurements of the deuteron spin structure function \(g_1^d\) show the remarkable feature that \(g_1^d\) is consistent with zero in the small-\(x\) region between 0.004 and 0.02 (Alexakhin et al., 2007). In contrast, the isovector part of \(g_1\) is observed to rise at small \(x\) as \(g_1^{p-n} \sim x^{-0.22\pm0.07}\) (Alekseev et al., 2010d) and is much larger than the isoscalar \(g_1^d\). This compares to the situation in the unpolarized structure function \(F_2\), where the small-\(x\) region is dominated by isoscalar gluonic exchanges.

### A. Spin sum rules

To test deep inelastic sum rules it is necessary to have all data points at the same value of \(Q^2\). Next-to-leading order (NLO) QCD-motivated fits taking into account the scaling violations associated with perturbative QCD are used to evolve all the data points to the same \(Q^2\). First moment sum rules are then evaluated by extrapolating these fits to \(x = 0\) and to \(x = 1\), or using a Regge-motivated extrapolation of the data. NLO QCD-motivated fits discussed in Sec. V.C are used to extract from these scaling violations the parton distributions and, in particular, the gluon polarization.

Polarized deep inelastic scattering experiments are interpreted in terms of a small value for the flavor-singlet axial charge. For example, COMPASS found using the SU(3) value for \(g_A^{(8)}\) (Alexakhin et al., 2007) and no leading-twist subtraction constant

\[
\frac{g_A^{(0)}_{\text{pDIS}, Q^2\rightarrow\infty}}{g_A^{(8)}} = 0.33 \pm 0.03(\text{stat}) \pm 0.05(\text{syst}).
\]

(17)

(This deep inelastic quantity misses any contribution to \(g_A^{(0)}_{\text{low}}\) from a possible delta function at \(x = 0\).) When combined with \(g_A^{(8)} = 0.58 \pm 0.03\), the value of \(g_A^{(0)}_{\text{DIS}}\) in Eq. (17) corresponds to a negative strange-quark polarization

\[
\Delta s_{Q^2\rightarrow\infty} = \frac{1}{3}(g_A^{(0)}_{\text{pDIS}, Q^2\rightarrow\infty} - g_A^{(8)}) = -0.08 \pm 0.01(\text{stat}) \pm 0.02(\text{syst}).
\]

(18)

The nonzero value of \(\Delta s_{Q^2\rightarrow\infty}\) in Eq. (18) is known as the violation of the Ellis-Jaffe sum rule (Ellis and Jaffe, 1974).

The extracted value of \(g_A^{(0)}_{\text{pDIS}}\) required to be understood by theory and the corresponding polarized strangeness depend on the value of \(g_A^{(8)}\). If we instead use the value \(g_A^{(8)} = 0.46 \pm 0.05\), the corresponding experimental value of \(g_A^{(0)}_{\text{DIS}}\) would increase to \(g_A^{(0)}_{\text{DIS}} = 0.36 \pm 0.03 \pm 0.05\) with

\[
\Delta s \sim -0.03 \pm 0.03.
\]

(20)

We discuss the value of \(\Delta s\) in more detail in Secs. V and VI in connection with more direct measurements from semi-inclusive deep inelastic scattering plus global fits to spin data, models, and recent lattice calculations with disconnected diagrams (quark sea contributions) included.

The Bjorken sum rule (Bjorken, 1966, 1970) for the isovector part of \(g_1\) follows from current algebra and is a fundamental prediction of QCD. The first moment of the isovector part of \(g_1\) is determined by the nucleon’s isovector axial charge

\[
\int_0^1 dx g_1^{p-n} = \frac{1}{6} g_A^{(3)} \left\{ 1 + \sum_{\ell=1} \alpha_s^2(Q) \right\}
\]

(21)

up to a 1% correction from charge symmetry violation suggested by a recent lattice calculation (Cloet et al., 2012). It has been confirmed in polarized deep inelastic scattering at the level of 5%. The value of \(g_A^{(3)}\) extracted from the most recent COMPASS data is \(1.28 \pm 0.07(\text{stat}) \pm 0.01(\text{syst})\) (Alekseev et al., 2010d) and compares well with the Particle Data Group value 1.270 \pm 0.003 deduced from neutron beta decays (Beringer et al., 2012).

The evolution of the Bjorken integral \(\int_0^1 dx g_1^{p-n}\) as a function of \(x_{\text{min}}\) and the isosinglet integral \(\int_0^{x_{\text{min}}} dx g_1^{p-n}\) are shown in Fig. 6. The Bjorken sum rule and isosinglet integral converges in the measured \(x\) region. Note that a large contribution, about 50%, of the Bjorken sum rule comes from \(x \leq 0.15\). The integral for the first moment of \(g_1^{p-n}\) saturates at \(x \sim 0.05\): the isosinglet part of \(g_1\) is close to zero in this measured range of small Bjorken \(x\).
The nucleon’s second spin structure function $g_2$ is believed to satisfy the Burkhardt-Cottingham sum rule $\int_0^1 dx_2 g_2 = 0$ (Burkhardt and Cottingham, 1970). The most precise measurements to date in polarized deep inelastic scattering come from the SLAC E155 and E143 experiments, which report $\int_{0.8}^{0.9} dx_2 g_2^p = -0.042 \pm 0.008$ for the proton and $\int_{0.8}^{0.9} dx_2 g_2^n = -0.006 \pm 0.011$ for the deuterium at $Q^2 = 5$ GeV$^2$ (Anthony et al., 2003). E155 estimate a contribution about 0.02 to the first moment of the proton $g_2$ comes from the $x$ range between 0 and 0.02 from the twist-2 (Wandzura-Wilczek) part of $g_2$: $g_2^{WW}(x) = \int_1^x (dy/y) g_1(y) - g_1(x)$.

B. Proton spin puzzles

The results from polarized deep inelastic scattering pose the following questions:

- How is the spin $1/2$ of the proton built up from the spin and orbital angular momentum of the quarks and gluons inside?
- Why is the quark spin content $g_A^{(0)}|_{pDIS}$ so small?
- How about $g_A^{(0)} \neq g_A^{(8)}$? What separates the values of the octet and singlet axial charges? How reliable is the SU(3) value of $g_A^{(8)}$?
- Is the proton spin puzzle a valence quark or a sea or glue effect?
- Can we extract information about the quark and gluon orbital angular-momentum contributions from experiments and with minimal model dependence?

We next discuss the theoretical development and experimental work that has been performed to address these questions and the physics interpretation of present measurements.

C. Spin and the singlet axial charge

There are two key issues involved in understanding the small value of $g_A^{(0)}|_{pDIS}$: the physics interpretation of the flavor-singlet axial charge $g_A^{(0)}$ and possible SU(3) breaking in the extraction of $g_A^{(8)}$ from hyperon $\beta$ decays. How big really is the Okubo-Zweig-Iizuka (OZI) violation $\Delta s = \frac{1}{2} \times (g_A^{(0)}|_{pDIS} - g_A^{(8)})$?

Theoretical QCD analysis based on the axial anomaly leads to

$$g_A^{(0)} = \left( \sum_q \Delta q - \frac{3}{2\pi} \Delta g \right)_{\text{partons}} + C_{\text{iso}},$$

see Altarelli and Ross (1988), Carlitz, Collins, and Mueller (1988), Efremov and Teryaev (1988), Bass et al. (1991), and Bass (2005). Here $\Delta g_{\text{partons}}$ is the amount of spin carried by polarized gluon partons in the polarized proton with $\alpha_s \Delta g \sim \text{const as } Q^2 \to \infty$ (Altarelli and Ross, 1988; Efremov and Teryaev, 1988); $\Delta g_{\text{partons}}$ measures the spin carried by quarks and antiquarks carrying soft transverse momentum $k_t^2 \sim O(P^2, m^2)$, where $P^2$ is a typical gluon virtuality in the nucleon and $m$ is the light quark mass. The polarized gluon term is associated with events in polarized deep inelastic scattering where the hard photon strikes a quark or antiquark generated from photon-gluon fusion and carrying $k_t^2 \sim Q^2$ (Carlitz, Collins, and Mueller, 1988). It is associated with the QCD axial anomaly in perturbative QCD. $C_{\text{iso}}$ denotes a potential nonperturbative gluon topological contribution with support only at Bjorken $x = 0$ (Bass, 2005). This term is discussed in Sec. VI on theoretical understanding. It is associated with the possible subtraction constant in the dispersion relation for $g_1$. If nonzero it would mean that $\lim_{x \to 0} \int_1^x (dy/y) g_1(y) - g_1(x)$.

Possible explanations for the small value of $g_A^{(0)}|_{pDIS}$ extracted from polarized deep inelastic experiments that have been suggested in the theoretical literature include screening from positive gluon polarization, possible SU(3) breaking in the isosinglet axial charges $g_A^{(8)}$ and $g_A^{(0)}$, negative strangeness polarization in the nucleon, a possible topological contribution at $x = 0$, plus connections to axial U(1) dynamics discussed by Fritzsch (1989), Narison, Shore, and Veneziano (1995), Bass (1999b), and Shore (2008).

The two-loop QCD evolution factor $E(\alpha_s)$ in Eq. (15) is associated with the polarized gluon term which carries all the scale dependence. The quark spin contribution $\Delta q_{\text{partons}}$ and the subtraction constant in Eq. (22) are QCD scale invariant. The quark spin term $\Delta q_{\text{partons}}$ is also known as the JET and chiral scheme (Cheng, 1996; Leader, Sidorov, and Stamenov, 1998) and Adler-Bardeen (AB) scheme (Ball, Forte, and Ridolfi, 1996) version of quark polarization; see Sec. V.C. In an alternative approach, called the minimal subtraction (MS) scheme (Bodwin and Qiu, 1990), $\sum_q \Delta q_{\text{mS}}$ is defined as the total matrix element of the flavor-singlet axial current [including the gluonic terms in Eq. (22)]. We return to this issue in Sec. V.C with discussion of QCD fits to experimental data. The growth in the gluon polarization $\Delta g / \alpha_s$ at large $Q^2$ is compensated by growth with opposite sign in the gluon orbital angular momentum.

One would like to understand the dynamics which yield a small value of the singlet axial charge extracted from polarized deep inelastic scattering and also the sum rule for the longitudinal spin structure of the nucleon.
\[ \frac{1}{2} = \frac{1}{2} \sum_{q} \Delta q + \Delta g + L_q + L_g, \]  
where \( L_q \) and \( L_g \) denote the orbital angular-momentum contributions. Operator definitions of the different terms or combinations of terms in this equation have been discussed by Jaffe and Manohar (1990), Ji (1997b), Shore and White (2000), Bakker, Leader, and Trueman (2004), Bass (2005), Chen et al. (2008), Wakamatsu (2010), and Leader (2011) and most recently by Hatta (2012), Ji, Xiong, and Yuan (2012), and Lorce (2012). We discuss orbital angular momentum and attempts to measure it in Secs. VI and VII.

There is presently a vigorous global program to disentangle the different contributions. Key experiments include semi-inclusive polarized deep inelastic scattering (COMPASS and HERMES) and polarized proton-proton collisions (PHENIX and STAR), as well as deeply virtual Compton scattering and hard exclusive meson production to learn about total angular momentum (COMPASS, HERMES, and JLab). Single-spin observables in semi-inclusive scattering from transversely polarized targets is sensitive to orbital angular momentum in the proton.

V. QUARK AND GLUON POLARIZATION FROM DATA

Key observables needed to understand the small value of the singlet axial charge \( g_A^{01} \) of the nucleon are the polarized strangeness and polarized glue in the nucleon. The search for polarized strangeness has inspired a dedicated experimental program to measure polarized glue in the nucleon. The search for polarized glue in the nucleon is motivated by the discovery of Altarelli and Ross (1987) and Efremov and Teryaev (1988) that polarized glue makes a scaling contribution to the first moment of the polarized structure function. Further, much activity was motivated by the discovery of Altarelli and Ross (1987) and Efremov and Teryaev (1988) that polarized glue makes a scaling contribution to the first moment of the polarized structure function. If there were a large negative contribution \(-3(\alpha_s/2\pi)\Delta g\) with, e.g., gluon polarization of the order of \(\Delta g \approx 2.5\) at \(Q^2 = 10\) GeV\(^2\), then this could reconcile the small measured value of \(g_A^{01}\) of the naive parton model expectation of about 0.6 through Eq. (22). This suggests a vigorous and ambitious program to measure \(\Delta g\).

Interesting channels include gluon mediated processes in semi-inclusive polarized deep inelastic scattering (COMPASS and HERMES) and hard QCD processes in high-energy polarized proton-proton collisions at RHIC.

A. Valence and sea polarization

Semi-inclusive measurements of fast pions and kaons in the current fragmentation region with final-state particle identification can be used to reconstruct the individual up-, down-, and strange-quark contributions to the proton’s spin. In contrast to inclusive polarized deep inelastic scattering where the \(g_1\) structure function is deduced by detecting only the scattered lepton, the detected particles in the semi-inclusive experiments are high-energy (greater than 20% of the energy of the incident photon) charged pions and kaons in coincidence with the scattered lepton. For large energy fraction \(z = E_p/E_T \approx 1\) the most probable occurrence is that the detected \(\pi^+\) and \(K^-\) contain the struck quark or antiquark in their valence Fock state. They therefore act as a tag of the flavor of the struck quark (Close, 1979).

In leading order the virtual-photon–proton double-spin (cross-section) asymmetry is

\[ A_{1p}^h(x, Q^2) = \frac{\sum_{q,h} e_q^2 f_q(x, Q^2) f_{1q}^h dz D^h_q(z, Q^2)}{\sum_{q,h} e_q^2 f_q(x, Q^2) f_{1q}^h dz D^h_q(z, Q^2)}, \]

where \(z_{\text{min}} \approx 0.2\). Here \(\Delta f_q(x, Q^2)\) is the quark (or antiquark) polarized parton distribution, \(f_q(x, Q^2)\) is the unpolarized distribution, and \(e_q\) is the quark charge; \(D^h_q(z, Q^2)\) is the fragmentation function for the struck quark or antiquark to produce a hadron \(h(=\pi^+, K^\pm)\) carrying energy fraction \(z = E_h/E_T\) in the target rest frame. Note the integration over the transverse momentum \(p_t\) (Close and Milner, 1991). Since pions and kaons have spin zero, the fragmentation functions are the same for both polarized and unpolarized lepton production. The fragmentation functions for \(u \to \pi^-\) and \(d \to \pi^+\) are known as “favored” (where the fragmenting quark has the same flavor as a valence quark in the final-state hadron); the fragmentation functions for \(u \to \pi^+\) and \(d \to \pi^-\) are known as “unfavored.”

This program for polarized deep inelastic scattering was pioneered by the SMC (Adeva et al., 1996, 1998a) and the HERMES (Ackerstaff et al., 1999b; Aipetian et al., 2004a, 2005a) experiments. The most recent measurements from HERMES are reported in Aipetian et al. (2008c) and from COMPASS in Alekseev et al. (2010c).

The experimental strategy has been to measure the asymmetries \(A_1^h\) for charged hadron production and separated charged pion and kaon production from proton and deuteron targets. There is good agreement between the COMPASS and HERMES data in the kinematic region of overlap; see Fig. 7. Flavor-separated polarized quark distributions for valence and sea quarks are then extracted from the data using fragmentation functions that have been fitted to previous hadron production data, with the most accurate taken to be those from the DSS group (de Florian, Sassot, and Stratmann, 2007) from a global fit to single-hadron production in \(e^+e^-\), \(ep\), and \(pp\) collisions.

The polarizations of the up and down quarks are positive and negative, respectively, while the extracted sea polarization data are consistent with zero; see Table II which includes measurements from COMPASS (Alekseev et al., 2010c), HERMES (Aipetian et al., 2005a), and SMC (Adeva et al., 1998a).

The COMPASS and HERMES determinations of the sum of strange and antistrange polarization \(\Delta s(x)\) are shown together in Fig. 8, plotted in the combination \(x\Delta s(x)\). There is no evidence in the semi-inclusive data for large negative strange-quark polarization in the nucleon. The HERMES data cover the region \(0.02 < x < 0.6\), where the extracted \(\Delta s\) is consistent with a zero or small positive value. These data integrate to \(\int_{0.02}^{0.6} dx x\Delta s = 0.037 \pm 0.019 \pm 0.027\) (Aipetian et al., 2008c) in contrast to the negative value for polarized strangeness, Eq. (18), extracted from inclusive measurements of \(g_1\). COMPASS measurements (Alekseev et al., 2009b, 2010c) show no evidence of strangeness polarization in the region \(x > 0.004\) with the integrated \(\Delta s = -0.02 \pm 0.02\pm 0.02\).

The precise value of \(\Delta s\) extracted from semi-inclusive scattering may be affected by any possible future improvement in the accuracy of the kaon fragmentation functions \(D_{q}^{K}(z)\).
However, a drastic change in the ratio $\int dx^0 D_{1,p}^u / \int dx^0 D_{1,p}^d$ would be needed to bring the first moment of $\Delta s$ extracted from semi-inclusive scattering in agreement with the inclusive value, Eq. (18), obtained using the $SU(3)$ value of $\beta_{A}^{(8)}$ (Alekseev et al., 2010c). More experimental data, especially on kaon fragmentation processes, are needed for improved precision on strangeness polarization in the nucleon.

Semi-inclusive data are consistent with a small positive or zero isospin asymmetry in the polarized sea $\Delta u - \Delta d$. For the COMPASS data at 3 GeV$^2$, one finds $\int_{0.023}^{0.1} dx (\Delta u - \Delta d) = 0.06 \pm 0.04$ (stat) $\pm 0.02$ (syst). For HERMES data at 2.5 GeV$^2$, $\int_{0.023}^{0.1} dx (\Delta u - \Delta d) = 0.048 \pm 0.057$ (stat) $\pm 0.028$ (syst) (Airapetian et al., 2005a). These values compare with the unpolarized sea measurement $\int_{0.2}^{0.8} dx (\Delta u - \Delta d) = -0.118 \pm 0.012$ from the E866 experiment at FNAL (Towell et al., 2001). A compilation of theoretical predictions is given in Peng (2003). Meson cloud models predict small negative isospin asymmetries in the polarized sea (Cao and Signal, 2001; Kumano and Miyama, 2002) whereas statistical (Bourrely, Soffer, and Buccella, 2002) and chiral quark soliton models (ChQSM) (Wakamatsu, 2003) predict positive values. The COMPASS and HERMES results are consistent with these predictions within uncertainties.

The $W$ boson production process at RHIC (Bunce et al., 2000) provides additional flavor-separated measurements of polarized up and down quarks and antiquarks. At RHIC the polarization of the $u$, $\bar{u}$, $d$, and $\bar{d}$ quarks in the proton is being measured directly using $W$ boson production in $u\bar{d} \to W^+$ and $\bar{d}u \to W^-$. The charged weak boson is produced through a pure $V-A$ coupling and the chirality of the quark and antiquark in the reaction is fixed. The $W$ is observed through its leptonic decay $W \to l\nu$, and the charged lepton is measured. Measurement of the flavor-separated antiquark helicity distributions via $W$ production in $p+p$ collisions is complementary to measurements via SIDIS in that there is no dependence on details of the fragmentation process, and the process scale $Q^2 = M_W^2$ is significantly higher than any data from existing fixed-target polarized DIS experiments. A parity-violating asymmetry for $W^+$ production in $p+p$ collisions at $\sqrt{s} = 500$ GeV consistent with predictions based on antiquark helicity distributions extracted from SIDIS has already been observed by both PHENIX (Adare et al., 2011b) and STAR (Aggarwal et al., 2011) based on data collected in 2009. Considerably improved results are expected from data taken in 2011 and 2012 with higher luminosities and polarization. Preliminary results for both $W^+$ and $W^-$ asymmetries from STAR, based on data taken at the beginning of 2012, are consistent with results from SIDIS and suggest the possible asymmetry $\Delta \bar{u} > \Delta \bar{d}$ for $x$ from 0.05 to 1 (Aschenauer et al., 2012).

An independent measurement of the strange-quark axial charge could be made through neutrino-proton elastic scattering. This process measures the combination $\frac{1}{2}(\Delta u - \Delta d - \Delta s)_{inv} = -0.01 g^{(0)}_{\Lambda} |_{inv}$, where the small last term is a correction from heavy quarks which has been calculated to leading-order (LO) (Kaplan and Manohar, 1988) and NLO.

**TABLE II.** First moments for valence quark and light-sea polarization from SMC, HERMES, and COMPASS. For each experiment the integrated sea quark is evaluated from data up to $x = 0.3$ and, for SMC, assuming an isospin symmetric polarized sea.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>$x$ range</th>
<th>$Q^2$ (GeV$^2$)</th>
<th>$\Delta u$</th>
<th>$\Delta d$</th>
<th>$\Delta \bar{u}$</th>
<th>$\Delta \bar{d}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMC</td>
<td>0.003–0.7</td>
<td>10</td>
<td>$0.73 \pm 0.10 \pm 0.07$</td>
<td>$-0.47 \pm 0.14 \pm 0.08$</td>
<td>$0.01 \pm 0.04 \pm 0.03$</td>
<td>$0.01 \pm 0.04 \pm 0.03$</td>
</tr>
<tr>
<td>HERMES</td>
<td>0.023–0.6</td>
<td>2.5</td>
<td>$0.60 \pm 0.07 \pm 0.04$</td>
<td>$-0.17 \pm 0.07 \pm 0.05$</td>
<td>$0.00 \pm 0.04 \pm 0.02$</td>
<td>$-0.05 \pm 0.03 \pm 0.01$</td>
</tr>
<tr>
<td>COMPASS</td>
<td>0.006–0.7</td>
<td>10</td>
<td>$0.67 \pm 0.03 \pm 0.03$</td>
<td>$-0.28 \pm 0.06 \pm 0.03$</td>
<td>$0.02 \pm 0.02 \pm 0.01$</td>
<td>$-0.05 \pm 0.03 \pm 0.02$</td>
</tr>
</tbody>
</table>
(Bass et al., 2002) accuracy. The axial charge measured in $\nu p$ elastic scattering is independent of any assumptions about possible $SU(3)$ breaking, the presence or absence of a subtraction at infinity in the dispersion relation for $g_1$, and the $x \sim 0$ behavior of $g_1$. A recent suggestion for an experiment using low-energy neutrinos produced from pion decay at rest is discussed by Pagliaroli using low-energy neutrinos produced from pion decay at rest (de Florian, Sassot, and Stratmann, 2007). The inner error bar represents the statistical uncertainty; the full bar represents the quadratic sum of statistical and systematic uncertainties.

In a recent analysis (Pate, McKee, and Papavassiliou, 2008) of parity-violating quasielastic electron and neutrino scattering data between 0.45 and 1 GeV$^2$ (from the JLab experiments G0 and HAPPEX and the Brookhaven experiment E734), the axial form factor was extrapolated to $Q^2 = 0$ and favored negative or zero values of $\Delta s$ with large uncertainty.

**B. Gluon polarization**

Polarized proton-proton scattering is sensitive to the ratio of polarized to unpolarized glue $\Delta g/g$, via leading-order interactions of gluons, as illustrated in Fig. 9. The first experimental attempt to look at gluon polarization was made by the FNAL E581/704 Collaboration using a 200 GeV polarized proton beam and a polarized proton target. They measured a longitudinal double-spin asymmetry $A_{LL}$ for inclusive multi-$\gamma$ and $\pi^0\pi^0$ production consistent with zero within their sensitivities, suggesting that $\Delta g/g$ is not so large in the region of $0.05 \leq x_g \leq 0.35$ (Adams et al., 1994). COMPASS was conceived to measure $\Delta g$ via the study of the photon-gluon fusion process, as shown in Fig. 10. The cross section for this process is directly related to the (polarized) gluon distribution at the Born level. The experimental technique consists of the reconstruction of charmed mesons (Alekseev et al., 2009c; Adolph et al., 2012d) or high-$p_T$ hadrons (Ageev et al., 2006) in the final state to access $\Delta g$. For the charmed meson case COMPASS also performed a NLO analysis which shifts probed $x_g$ to larger values. The high-$p_T$ particle method leads to samples with larger statistics, but these have higher background contributions from QCD Compton processes and fragmentation. High-$p_T$ hadron production was also used in early attempts to access gluon polarization by HERMES (Airapetian et al., 2000a) and SMC (Adeva et al., 2004) and the most recent HERMES determination (Airapetian et al., 2010c) and COMPASS measurement (Adolph et al., 2012e).

These measurements in lepton-nucleon scattering are listed in Table III for the ratio of the polarized to unpolarized glue $\Delta g/g$ and shown in Fig. 11 for LO analyses of the data. The data cluster is around $x_g \sim 0.1$ with the exception of the COMPASS NLO point from open charm. There is no evidence in the data for nonzero gluon polarization at this value of $x_g$.

The chance to measure $\Delta g$ was a main physics drive for polarized RHIC. Experiments using the PHENIX and STAR detectors are investigating polarized glue in the proton. Measurements of $\Delta g/g$ from RHIC are sensitive to gluon polarization in the range $0.02 \leq x_g \leq 0.3$ ($\sqrt{s} = 200$ GeV) and $0.06 \leq x_g \leq 0.4$ ($\sqrt{s} = 62.4$ GeV) for the neutral pion $\Lambda_{LL}$ measurement by PHENIX (Adare et al., 2009a, 2009b) and inclusive jet production measured by STAR at 200 GeV center-of-mass energy (Abelev et al., 2008b; Adamczyk et al., 2012a). Additional channels sensitive to $\Delta g$ at RHIC have been published as well (Abelev et al., 2009; Adare et al., 2011a, 2012).

Combined preliminary results from PHENIX and STAR using more recent 200 GeV data than those published in...
Abelev et al. (2008b) and Adare et al. (2009b) are shown in Fig. 12. The longitudinal double-spin asymmetry in neutral pion production measured by PHENIX based on combined data from 2005, 2006, and 2009 is shown as a function of pion $p_T$ (upper scale) (Manion, 2011). Figure 12 also shows the asymmetry in single-inclusive jet production as a function of jet $p_T$ (lower scale) measured by STAR based on data taken in 2009 (Djawotho, 2011), providing the first evidence for nonzero gluon polarization in the proton. The relationship between the pion and jet $p_T$ scales is given by the mean $z$ value of $\sim 0.5$ (Adler et al., 2006). The data are shown with a calculation using helicity distributions extracted from a global fit to polarized world data from DIS, semi-inclusive DIS, and proton-proton collisions of de Florian–Sassot–Stratmann–Vogelsang (DSSV) (de Florian et al., 2008, 2009) that was updated to include these results (Aschenauer et al., 2012); see Sec. V.C for more details about fits to helicity distributions. A given $p_T$ bin for single-inclusive jet or hadron production generally samples a wide range of $x_g$ values. However, dijet measurements in $p + p$ collisions provide better constraints on the $x_g$ values probed. Preliminary STAR results for dijet production have also been released (Walker, 2011) and confirm the nonzero double-spin asymmetry seen in single jet production.

While there is now evidence in the RHIC data that gluon polarization in the proton is nonzero, the measurements indicate that polarized glue, by itself, is not sufficient to resolve the difference between the small value of $g_A^{(0)}$pDIS and the naive constituent quark model prediction $-0.6$ through the polarized gluon term $-3(\alpha_s/\pi)\Delta g$. Note, however, that gluon polarization $\Delta g \sim 0.2$–$0.3$ would still make a significant contribution to the spin of the proton in Eq. (23).

C. NLO QCD-motivated fits to spin data

Global NLO perturbative QCD analyses are performed on polarization data sets including both lepton-nucleon and proton-proton collision data. The aim is to extract the polarized quark and gluon parton distributions. These analyses, starting from Ball, Forte, and Ridolfi (1996) and Altarelli et al. (1997), frequently use Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) evolution and are performed in a given factorization scheme. This QCD fit approach has more recently been extended to a global analysis of data from polarized DIS, semi-inclusive polarized DIS, and high-energy

![FIG. 11 (color online). Gluon polarization $\Delta g/g$ from leading-order analyses of hadron or hadron-pair production as a function of the probed gluon momentum fraction $x_g$. Also shown are NLO fits from de Florian et al. (2009) and Leader, Sidorov, and Stamnoven (2010). The inner error bar represents the statistical uncertainty; the full bar represents the quadratic sum of statistical and systematic uncertainties. The horizontal bar indicates the $x_g$ range of the measurement. Adapted from Adolph et al., 2012e.](image1)

![FIG. 12 (color online). The longitudinal double-spin asymmetry in $\pi^0$ production measured by PHENIX (Manion, 2011) and in jet production measured by STAR (Djawotho, 2011), shown with calculations based on the DSSV polarized parton distributions that were updated to include these results. The relationship between the pion and jet $p_T$ scales is given by the mean $z$ value of $\sim 0.5$ (Adler et al., 2006). Error bars represent the statistical uncertainty. From Aschenauer et al., 2012.](image2)
polarized proton-proton collisions (de Florian et al., 2008, 2009).

The separation of $g_1$ into hard and soft contributions is not unique and depends on the choice of factorization scheme. For example, one might use a kinematic cutoff on the partons’ transverse momentum squared ($k_t^2 > \mu^2$) to define the factorization scheme and thus separate the hard and soft parts of the phase space for the photon-parton collision. The cutoff $\mu^2$ is called the factorization scale. The coefficients in Eq. (8) have the perturbative expansion

$$ C^q = \delta(1-x) + \frac{\alpha_s}{2\pi} f^q(x, Q^2/\mu^2) $$

and

$$ C^g = \frac{\alpha_s}{2\pi} f^g(x, Q^2/\mu^2), $$

where the strongest singularities in the functions $f^q$ and $f^g$ as $x \to 1$ are $\ln(1-x)/(1-x)$, and $\ln(1-x)$, respectively; see, e.g., Lampe and Reya (2000). The deep inelastic structure functions are dependent on $Q^2$ and independent of the factorization scale $\mu^2$ and the scheme used to separate the $\gamma^*$-parton cross section into hard and soft contributions.

Examples of different schemes used in the literature are the modified minimal subtraction ($\overline{\text{MS}}$) (t’Hooft and Veltman, 1972; Bodwin and Qiu, 1990) to regulate the mass singularities which arise in scattering from massless partons, the AB (Ball, Forte, and Ridolfi, 1996) and CI (chiral invariant) (Cheng, 1996) or JET (Leader, Sidorov, and Stamenov, 1998) schemes. In the $\overline{\text{MS}}$ scheme the polarized gluon distribution does not contribute explicitly to the first moment of $g_1$. In the AB, CI, and JET schemes, on the other hand, the polarized gluon (axial-anomaly contribution) $\alpha_s\Delta g$ contributes explicitly to the first moment since $\int_0^1 dx x C^g_{\text{CI}} = -\alpha_s/2\pi$; see the spin decomposition in Eq. (22).

The $\mu^2$ dependence of the parton distributions is given by the DGLAP equations (Altarelli and Parisi, 1977)

$$ \frac{d}{dt} \Delta \Sigma(x, t) = \left[ \int \frac{dy}{y} \Delta P_{qq}(x/y, \alpha_s(t)) \Delta \Sigma(y, t) \right] \chi^2 + \frac{2f}{\alpha_s} \int \frac{dy}{y} \Delta P_{qs}(x/y, \alpha_s(t)) \Delta g(y, t), $$

$$ \frac{d}{dt} \Delta g(x, t) = \left[ \int \frac{dy}{y} \Delta P_{ss}(x/y, \alpha_s(t)) \Delta \Sigma(y, t) \right] \chi^2 + \frac{1}{\alpha_s} \int \frac{dy}{y} \Delta P_{qs}(x/y, \alpha_s(t)) \Delta g(y, t), $$

where $\Delta \Sigma(x, t) = \sum_q \Delta q(x/t, t) = \ln \mu^2$, and $f$ is the number of active flavors. The splitting functions $P_{ij}$ in Eq. (25) have been calculated at leading order by Altarelli and Parisi (1977) and at next-to-leading order by Zijlstra and van Neerven (1994), Mertig and van Neerven (1996), and Vogelsang (1996).

The largest uncertainties in the QCD fits are associated with the ansatz chosen for the shape of the spin-dependent quark and gluon distributions at a given input scale. Further, the SU(3) value of $g_A^{(0)} (= 0.58 \pm 0.03)$ is assumed in present fits although no significant change in the $\chi^2$ quality of the fits should be expected if one instead took a value of $g_A^{(0)}$ with possible 20% SU(3) breaking included.1 The values for the quark and gluon spin contents ($\Delta \Sigma$ and $\Delta g$) obtained in recent NLO fits are listed in Table IV with results quoted from Blümlein and Böttcher (2010) (BB10: DIS data), Nocera et al. (2012) (NFRR12: DIS data), Leader, Sidorov, and Stamenov (2010) (LSS10: DIS and SISID data), and de Florian et al. (2008, 2009) (DSSV08: DIS, SISID, and proton-proton collision data).

The most complete fits in terms of maximum included data are from the DSSV group, which take the SU(3) value for $g_A^{(0)}$. One finds need for a large negative contribution to $\Delta s$ from small $x$, outside the measured $x$ range when SISID data are included. The values obtained in this approach for $\int_0^{x_{\text{min}}} dx x \Delta s(x)$ are $\Delta s = -0.057$ with $x_{\text{min}} = 0$ and about $-0.001$ with $x_{\text{min}} = 0.001$. That is, to reproduce the SU(3) value of the octet axial charge, the negative polarized strangeness obtained from inclusive $g_1$ measurements gets pushed into the unmeasured small-$x$ range $x < 0.004$. It is interesting here to note that, historically (before COMPASS, HERMES, and RHIC spin), the proton spin puzzle was assumed to be associated with strangeness, sea, or glue polarization in the newly opened kinematics of EMC, SLAC, and SMC, $x$ between 0.1 and 0.01. We now have accurate SISID measurements down to $x \sim 0.004$ which show no evidence for large sea or glue polarization effects. With the SISID measurements of $\Delta s$, one needs either SU(3) breaking in the octet axial charge or strangeness or glue effects at very small $x$. Without including the most recent data from 2009 or later, de Florian et al. (2008, 2009) found a best-fit full first moment $\int_0^{x_{\text{min}}} dx x \Delta g(x) = -0.084$ at $Q^2 = 10 \text{ GeV}^2$. With a very large $\Delta \chi^2/\chi^2 = 2\%$ allowed range, the truncated first moment $\int_0^{x_{\text{min}}} dx x \Delta g = 0.013^{+0.702}_{-0.314}$ was obtained. With these errors $\Delta g$ is still not precise.

A recent attempt to extract polarized parton distributions from inclusive polarized deep inelastic data using neural network techniques is reported by Nocera et al. (2012). In this approach no assumption is made about the functional form of the input distributions, greatly reducing the primary (and notoriously difficult to quantify) systematic uncertainty on parton distribution fits. The neural network method has already been used quite successfully in the parametrization of the unpolarized parton distributions, with results published at

---

1We thank S. Taneja and R. Windmolders for discussion on this issue.

---

Table IV. First moments of the polarized singlet-quark and gluon distributions at the scale $4 \text{ GeV}^2$ in the $\overline{\text{MS}}$ scheme; values quoted from Nocera et al. (2012).

<table>
<thead>
<tr>
<th></th>
<th>DSSV08</th>
<th>BB10</th>
<th>LSS10</th>
<th>AAC08</th>
<th>NFRR12</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta \Sigma(Q^2)$</td>
<td>$0.25 \pm 0.02$</td>
<td>$0.19 \pm 0.08$</td>
<td>$0.21 \pm 0.03$</td>
<td>$0.24 \pm 0.07$</td>
<td>$0.31 \pm 0.10$</td>
</tr>
<tr>
<td>$\Delta g(Q^2)$</td>
<td>$-0.10 \pm 0.16$</td>
<td>$0.46 \pm 0.43$</td>
<td>$0.32 \pm 0.19$</td>
<td>$0.63 \pm 0.19$</td>
<td>$-0.2 \pm 1.4$</td>
</tr>
</tbody>
</table>
NLO in 2010 (Ball et al., 2010) and now NNLO in 2012 (Ball et al., 2012). However, in the case of the application of the neural network method in the extraction of polarized parton distributions, thus far only inclusive polarized DIS data have been incorporated, similar to the BB10 fit (Blümlein and Böttcher, 2010).

VI. THEORETICAL UNDERSTANDING

In relativistic quark models some of the proton’s spin is carried by quark orbital angular momentum. One has to take into account the four-component Dirac spinor for the quarks

$$\psi = \frac{N}{\sqrt{4\pi}} \left( f \sigma \cdot \gamma_R \right)$$

where \( f \) and \( g \) are functions of the spatial coordinates and \( N \) is a normalization factor. The lower component of the Dirac spinor is a \( \psi \) wave with intrinsic spin primarily pointing in the opposite direction to spin of the nucleon. In the MIT bag model, where quarks are confined in an infinite square well potential with radius \( R \), one finds the depolarization factor \( N^2 \int_0^\infty dR^2 (f^2 - \frac{1}{4}g^2) = 0.65 \) for \( \Delta q \) in the proton with all quarks in the (1s) ground state (Jaffe and Manohar, 1990). That is, 35% of the proton’s spin content is shifted into orbital angular momentum through the confinement potential.

More detailed calculations of nonsinglet axial charges in relativistic constituent quark models are sensitive to the confinement potential, effective color-hyperfine interaction, pion and kaon clouds, plus additional wave-function corrections (associated with center-of-mass motion) chosen to reproduce the physical value of \( g_A^{(3)} \).

This physics was recently investigated within the cloudy bag model (Myhrer and Thomas, 2008; Thomas, 2008; Bass and Thomas, 2010). The cloudy bag was designed to model confinement and spontaneous chiral symmetry breaking, taking into account pion physics and the manifest breakdown of chiral symmetry at the bag surface in the MIT bag. If we wish to describe proton spin data including matrix elements of \( J_{\mu 5}, J_{\mu 5}^{(3)}, \) and \( J_{\mu 5}^{(8)} \), then we would like to know that the model versions of these currents satisfy the relevant Ward identities (the divergence equations for these currents). For the scale-invariant nonsinglet axial charges \( g_A^{(3)} \) and \( g_A^{(8)} \), corresponding to the matrix elements of partially conserved currents, the model is well designed to make a solid prediction.

The effective color-hyperfine interaction has the quantum numbers of OGE. In models of hadron spectroscopy this interaction plays an important role in the nucleon-\( \Lambda \) and \( \Sigma-\Lambda \) mass differences, as well as the nucleon magnetic moments (Close, 1979) and the spin and flavor dependence of parton distribution functions (Close and Thomas, 1988). It shifts total angular momentum between spin and orbital contributions and, therefore, also contributes to model calculations of the octet axial charges (Myhrer and Thomas, 1988).

With OGE included (together with a phenomenological wave-function renormalization to ensure \( g_A^{(3)} \) takes the physical value), the model is in very good agreement with the SU(3) fit to the nucleon and hyperon axial charges extracted from \( \beta \) decays with \( g_A^{(8)} \) predicted to be around 0.6.

Next, the pion cloud induces SU(3) breaking in the nucleon’s axial charges. The pion cloud further shifts intrinsic spin into orbital angular momentum (Schreiber and Thomas, 1988; Tsushima et al., 1988). Including pion and kaon cloud corrections gives the model result \( g_A^{(8)} = 0.46 \pm 0.05 \) (with the corresponding semiclassical singlet axial charge or spin fraction being 0.42 \( \pm 0.07 \) before inclusion of gluonic effects) (Bass and Thomas, 2010). With this cloudy bag value for \( g_A^{(8)} \), the corresponding experimental value of \( g_A^{(8)}_{\text{pDIS}} \) increases to \( g_A^{(8)}_{\text{pDIS}} = 0.36 \pm 0.03 \pm 0.05 \), considerably reducing the apparent OZI violation \( \frac{1}{2} (g_A^{(8)}_{\text{pDIS}} - g_A^{(8)}) \) that one needs to explain.

A recent lattice calculation with disconnected diagrams included (Bali et al., 2012) gave \( \Delta s = -0.02 \pm 0.01 \) in the MS scheme at 7.4 GeV\(^2\). This value compares with the cloudy bag prediction \( \Delta s \approx -0.01 \) before gluonic degrees of freedom are included. These numbers are in good agreement with the values extracted from polarized SIDIS data by COMPASS and HERMES with the DSS fragmentation functions.

Gluon polarization has been investigated in bag and light-cone models and in studies of heavy-quark axial charges. The nucleon’s charm-quark axial charge was interpreted to give an estimate of gluon polarization \( |\Delta g(m^2)| \lesssim 0.3 \) with \( \alpha_s(m^2) = 0.4 \) (Bass, Casey, and Thomas, 2011). This upper bound corresponds to \( |\bar{q}\alpha_s/2\pi|\Delta g| \lesssim 0.06 \). Values of \( \Delta g \approx 0.3 \) and \( \approx 0.5 \) at 1 GeV\(^2\) were obtained in the MIT bag model (Chen and Ji, 2008) and in a light-cone model (Brodsky, Burkardt, and Schmidt, 1995), respectively. These theoretical values are consistent with the extractions of gluon polarization from COMPASS, HERMES, and RHIC spin data.

To understand \( C_{\text{eq}} \),\(^2\) deep inelastic sum rules are derived using the operator product expansion and the dispersion relation for deeply virtual photon-nucleon scattering. Two important issues with the dispersion are the convergence of the first moment integral at the highest energies and any contribution from closing the circle in the complex momentum plane. The subtraction constant, if finite, corresponds to a constant real term in the forward Compton scattering amplitude. It affects just the first moment integral and thus behaves as a \( \delta(x) \) term with support only at \( x = 0 \). A subtraction constant yields a finite correction to the sum rule obtained from integrating only over finite nonzero values of Bjorken \( x \). One can show (Bass, 2005) that a nonlocal gluon topological structure requires consideration of a possible \( \delta(x) \) subtraction constant. Whether it has finite value or not is sensitive to the realization of axial U(1) symmetry breaking by instantons (Crewther, 1978; ‘t Hooft, 1986) and the importance of topological structure in the proton. The QCD vacuum is a Bloch superposition of states characterized by nonvanishing topological winding number and nontrivial chiral properties. When we put a valence quark into this vacuum it can act as a source which polarizes the QCD vacuum with net result that the spin “dissolves.” Some fraction of the spin of the constituent quark is shifted from moving partons into the vacuum at \( x = 0 \).

\(^2\)In the notation of Eq. (12): \( \beta_{\text{eq}} = -\frac{1}{2} C_{\text{eq}} \left[ 1 + \sum_{i=1}^{n} \epsilon_{i} \alpha_{i}(Q) \right] \).
contribution becomes associated with nonlocal gluon topology with support only at Bjorken $x = 0$. Valuable information about the spin puzzle also follows from looking at the $x$ dependence of $g_1$. The small value of $g_A^{[0]}$ or "missing spin" is associated with a "collapse" in the isosinglet part of $g_1$ to something close to zero instead of a valence-like rise at $x$ less than about 0.05; see, e.g., the $g_1$ data in Fig. 4 and the convergence of $\int_{x_{\text{min}}}^{1} dx g_1^{[0]}$ and $\int_{x_{\text{min}}}^{1} dx q g_1^{[0]}$ in Fig. 6. This isosinglet part is the sum of SU(3) flavor singlet and octet contributions. If there were a large positive polarized gluon contribution to the proton’s spin, this would act to drive the small $x$ part of the isosinglet part of $g_1$ negative (Bass and Thomas, 1993), that is, acting in the opposite direction to any valence-like rise at small $x$. However, gluon polarization measurements constrain this spin contribution to be small in measured kinematics meaning that the sum of valence and sea-quark contributions is suppressed at small $x$. Neither the SU(3) flavor singlet nor the octet contribution breaks free in the measured small $x$ region. Hence, the suppression of $g_1^{[0]}$ at small $x$ should be either an isosinglet effect or a delicate cancellation between octet and singlet contributions over an order of magnitude in small Bjorken $x$.

The $g_1^{[0]}$ data are consistent with quark model and perturbative QCD counting rule predictions in the valence region $x > 0.2$ (Bass, 1999a). The size of $g_A^{[0]}$ forces us to accept a large contribution from small $x$ (a nonperturbative constraint) and the rise in $g_1^{[0]}$ is in excellent agreement with the prediction $g_1^{[0]} \sim x^{-0.22}$ of hard Regge exchange (Bass, 2007a), in particular, a possible $a_1$ hard-pomeron cut involving the hard pomeron which seems to play an important role in unpolarized deep inelastic scattering (Cudell, Donnachie, and Landshoff, 1999) and in the proton-proton total cross section measured at the CERN Large Hadron Collider (Donnachie and Landshoff, 2011). (Soft) Regge theory predicts that the singlet term should behave as $\sim N \ln x$ in the small $x$ limit, with the coefficient $N$ to be determined from experiment (Bass and Landshoff, 1994; Close and Roberts, 1994). From the data, this normalization seems to be close to zero.

Where are we in our understanding of the spin structure of the proton and the small value of $g_A^{[0]}$ in DIS? Measurements of valence, gluon, and sea polarization suggest that the polarized glue term $-3(\alpha_s/2\pi)A g$ and the strange-quark contribution $\Delta s_{\text{partons}}$ in Eq. (22) are unable to resolve the small value of $g_A^{[0]}$. Two explanations are suggested within the theoretical and experimental uncertainties depending upon the magnitude of SU(3) breaking in the nucleon and hyperon axial charges. One is a value of $g_A^{[0]} \sim 0.5$ plus an axial U(1) topological effect at $x = 0$ associated with a finite subtraction constant in the $g_1$ dispersion relation. The second is a much larger pion-cloud reduction of $g_A^{[0]}$ to a value $\sim 0.4$. Combining the theoretical error on the pion-cloud chiral corrections embraces both possibilities. The proton spin puzzle seems to be telling us about the interplay of valence quarks with chiral dynamics and the complex vacuum structure of QCD.

OAM in relativistic quark models (for example, the MIT and cloudy bag models) without explicit gluon degrees of freedom has the usual interpretation of relativistic quantum mechanics. For QCD dynamics the definition of OAM is more subtle because of the gauge covariant derivative, meaning that quark orbital momentum is in principle sensitive to the gluon fields in the nucleon that the quarks interact with. Going beyond spin and helicity to consider also orbital and total angular momentum, several operator decompositions have been proposed. Starting from the relation between angular momentum and the energy-momentum tensor, Ji (1997b) takes

$$\vec{J}_g = \int d^3 x \vec{\gamma} \times \vec{\gamma}_g$$

$$= \int d^3 x \left[ \psi^\dagger \frac{\gamma}{2} \psi + \psi^\dagger \vec{\gamma} \times (-i\vec{\partial}) \psi \right], \quad (26)$$

$$\vec{J}_g = \int d^3 x \vec{\gamma} \times (\vec{E} \times B).$$

The gauge covariant derivative $D_\mu = \partial_\mu + igA_\mu$ with $A_\mu$ as the gluon field means that $L_q$ is a priori sensitive to gluonic degrees of freedom. The $J_g$ and $J_q$ quantities here are amenable to QCD lattice calculations and, in principle, measurable through deeply virtual Compton scattering. In an alternative approach, taking the plus light-cone component of the QCD angular-momentum tensor in $A_+ = 0$ gauge, Jaffe and Manohar (1990) proposed the operator decomposition

$$M_{\Delta J = 2} = \frac{1}{2} \gamma_5 q g_+ + q^2 (\vec{\pi} \times i\vec{\partial}) q_+$$

$$+ 2 \text{Tr} F^{\mu\nu}(\vec{\pi} \times i\vec{\partial}) A_\mu + \text{Tr} e^{\mu\nu}(F^{\mu\nu} A_\mu), \quad (27)$$

where the gluon term in the gauge covariant derivative is no longer present through the gauge fixing.

The connection between the quark and gluon total angular-momentum contributions $J^q$ and $J^g$ and the QCD energy momentum tensor allows us to write down their LO QCD evolution equations (Ji, Tang, and Hoodbhoy, 1996). The quark and gluon total angular momenta in the infinite scaling limit are given by $J_q(\infty) = \frac{1}{2} (3f/(16 + 3f))$ and $J_g(\infty) = \frac{1}{2} (16/(16 + 3f))$, with $f$ the number of active flavors, that is, the same scaling limit as the quark and gluon momentum contributions at infinite $Q^2$. The Ji and Jaffe-Manohar definitions of orbital angular momentum satisfy the same (LO) QCD evolution equation, and as such, at LO, are equal in a model calculation if the glue contribution can be set equal to zero at a low-energy input scale.

To obtain information about the quark “orbital angular momentum” $L_q$ we need to subtract the value of the “intrinsinc spin” $S_q = \frac{1}{2} \Delta q$ measured in polarized deep inelastic scattering from the total quark angular momentum $J_q$. This means that $L_q$ is scheme dependent with different schemes corresponding to different physics content depending on how the scheme handles information about the axial anomaly, large-$k_t$ physics and any possible “subtraction at infinity” in the dispersion relation for $g_1$. The quark total angular momentum $J_q$ is anomaly free in QCD so that axial-anomaly effects occur with equal magnitude and opposite sign in $L_q$ and $S_q$. When looking at physical observables that are sensitive to OAM and quark spin (with possible axial-anomaly contribution) it will be important to identify which OAM definition and which scheme quantity is most relevant to the observable; for example, in SIDIS the largest $k_t$ events are
included in the \( \overline{\text{MS}} \) version of \( \Delta g \) whereas they are omitted in the JET scheme version (Bass, 2003).

There are some theoretical subtleties when dealing with gluon angular momentum. In the parton model the gluon polarization \( \Delta g \) has a clean interpretation in light-cone gauge as the forward matrix element of the local Chern-Simons current \( K_\chi \) (appearing in the QCD axial anomaly) up to a surface term which has support only at \( x = 0 \) (Manohar, 1990; Bass, 2005). In the light-cone gauge \( K_\chi \) coincides with the gluon spin operator (Jaffe, 1996). In general, Ji’s \( J_g \) in Eq. (26) is not readily separable into spin and orbital components. New ideas have recently been investigated where one separates the gluon field into a “physical” transverse part and “pure” gauge part, with different conventions on how to deal with the gauge part (Chen et al., 2008; Wakamatsu, 2010; Hatta, 2012; Lorce, 2012). Discussion of total orbital angular momentum involving gluonic degrees of freedom should be labeled with respect to the scheme or convention used.

To connect quark model predictions with lattice calculations and fits to data it is necessary to use QCD evolution of the model results from the low-energy scale where the model applies up to the hard scale of deep inelastic scattering. Model calculations (and also lattice calculations without disconnected diagrams) of \( \Delta g \) are commonly understood to refer to the scale-invariant version of this quantity, e.g., the chiral/JET or AB scheme quark spin contributions in Eq. (22). One chooses a model ansatz for the gluon polarization and total angular momentum, typically \( \Delta g = J_g = 0 \) at the model input scale. For illustration, Fig. 13 shows the evolution of total and orbital angular-momentum contributions in the cloudy bag from the model scale up to \( Q^2 = 4 \text{ GeV}^2 \). Various phenomenological investigations (Mattingly and Stevenson, 1994; Steffens, Holtmann, and Thomas, 1995) found that by going beyond leading-order QCD (and including pions in the nucleon wave function), the optimal fit to high-energy scattering data involved taking the running coupling \( \alpha_s \) about 0.6–0.8 at the low-energy input scale. For this range of \( \alpha_s \), the scale dependence of \( \Delta \Sigma \) (in full QCD) through Eq. (15) converges well. (Going to higher orders in the model fits to data and putting in pions raises the model input scale \( \mu_0 \) needed for the calculations.) Table V compares the results of lattice calculations (Hägler et al., 2008) for up- and down-quark spin and total angular momentum with the cloudy bag model results and the values extracted from QCD fits to hard exclusive reaction data, GPDs (Goloskokov and Kroll, 2009), and transverse single-spin asymmetries, TMDs (Bacchetta and Radici, 2011). The lattice calculation involves connected diagrams only (no axial-anomaly contribution) plus chiral extrapolation. The QCD fit numbers are central values modulo (possibly large) systematic errors from the model functional forms of distributions used in the fits. There is good convergence of the different theoretical values with “data.” Here one has \( L_u \sim L_d \sim 15\% \) at the scale of typical deep inelastic measurements.

In an alternative approach, the proton spin puzzle has also been addressed in the Skyrme model, where baryons emerge as topological solitons in the meson fields at large number of colors \( N_c \), and in the ChQSM, where explicit quark degrees of freedom are also present in the model. The nucleon’s axial charges in these models are sensitive to which mesons are included in the model and the relative contribution of a quark source and pure meson component. In an early calculation Brodsky, Ellis, and Karliner (1988) found that \( g_A^{(0)} \) vanishes in a particular version of the Skyrme model with just pseudo-scalar mesons. Finite values of \( \Delta \Sigma \sim -0.08 \) close to the value obtained from inclusive \( g_1 \) measurements with good SU(3) assumed for \( s_A^{(0)} \) are found in the ChQSM (Wakamatsu, 2007).

### VII. TRANSVERSE NUCLEON STRUCTURE AND ORBITAL ANGULAR MOMENTUM

Confinement induces transverse hadronic scales in the nucleon with accompanying finite quark orbital angular momentum and finite spin-orbit couplings which can be probed in experiments. The search for orbital angular momentum has motivated new theoretical and experimental investigations of the three-dimensional structure of the nucleon. Key observables in DVCS and transverse single-spin asymmetries in lepton-nucleon and proton-proton scattering, and also the large \( x \) limit of the down-quark helicity distribution, are sensitive to orbital angular momentum in the nucleon.

Hard exclusive reactions such as DVCS are described theoretically using the formalism of GPDs and probe the three-dimensional spatial structure of the nucleon, as reviewed by Ji (1998), Goeke, Polyakov, and Vanderhaeghen (2001), and Diehl (2003). Ji (1997b) derived a sum rule connecting the forward limit of GPDs to information about the quark and gluon total angular momentum in the proton.

![FIG. 13](color online). Calculation of the NLO QCD evolution of \( J_u, L_d, J_d, \) and \( L_u \) in the cloudy bag with model input scale \( Q_0 = 0.4 \text{ GeV} \). From Thomas, Casey, and Matevosyan, 2010.

### TABLE V. Model, lattice, and fit extractions of angular-momentum contributions in the proton, quoted for 4 GeV\(^2\) (except GPD at 2 GeV\(^2\)).

<table>
<thead>
<tr>
<th>Source</th>
<th>Cloudy bag</th>
<th>Lattice</th>
<th>GPD</th>
<th>TMD</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Delta \mu )</td>
<td>0.85 ± 0.06</td>
<td>0.82 ± 0.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \Delta d )</td>
<td>-0.42 ± 0.06</td>
<td>-0.41 ± 0.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( J_u )</td>
<td>0.30</td>
<td>0.24 ± 0.05</td>
<td>0.24</td>
<td>0.24</td>
</tr>
<tr>
<td>( J_d )</td>
<td>-0.04</td>
<td>0.00 ± 0.05</td>
<td>0.02</td>
<td>0.02</td>
</tr>
</tbody>
</table>
Considerable experimental and theoretical effort has and continues to be invested aimed at accessing this information. Studies of single-spin asymmetries for semi-inclusive meson production in high-energy lepton-nucleon and proton-proton collisions are sensitive to possible spin-orbit coupling both in the nucleon and in the final-state hadronization process; for a recent review, see Barone, Bradamante, and Martin (2010). One studies correlations between the transverse momentum (orbital motion) of partons, their spin, and the spin polarization of the nucleon. The theoretical tools are TMD distributions and fragmentation functions. TMDs probe the three-dimensional transverse-momentum structure of the nucleon and are associated, in part, with finite orbital angular momentum.

Experimental studies of three-dimensional nucleon structures have been pioneered at HERMES and JLab for GPDs and at COMPASS, HERMES, and RHIC for TMDs in single-spin asymmetry measurements. There has also been considerable theoretical effort aimed at model and lattice calculations of these observables.

In the remainder of this section we present the theory and present status of these new GPDs and TMD distributions plus spin-orbit coupling in fragmentation and the prospects for future experiments including key observables that will be studied. The aim for experiments should be to focus on observables that have the cleanest theoretical interpretation with minimal model dependence.

Quark orbital angular momentum in the nucleon may also be manifest in future measurements of the large-\(x\) behavior of the polarized down-quark distribution \(\Delta d/d\) and in the ratio of the proton’s spin-flip Pauli form factor to the Dirac form factor \(Q^2\). These observables can be studied with the 12 GeV upgrade of JLab. Valence Fock states with nonzero orbital angular momentum induce a logarithmic correction to the polarized down-quark distribution \(d=\overline{d}\) should cross zero and become 

\[ F_2/F_1 \approx (\log^2 Q^2/A^2)/Q^2 \tag{28} \]

for the ratio of Pauli-to-Dirac form factors at large \(Q^2\) (Belitsky, Ji, and Yuan, 2003). Form-factor measurements at JLab (Jones et al., 2000; Gayou et al., 2002) are consistent with this behavior and also with \(F_2/F_1 \approx 1/\sqrt{Q^2}\) for \(Q^2\) between 4 and 6 GeV\(^2\), in contrast to the counting rules prediction without orbital angular momentum \(F_2/F_1 \approx 1/Q^2\) (Lepage and Brodsky, 1980). One also finds a logarithmic correction to the leading large-\(x\) behavior of the negative-helicity spin-dependent quark distributions \(-1-x)^3 \log(1-x)\) (Avakian et al., 2007). An interesting prediction here is that \(\Delta d/d\) should cross zero and become positive at a value \(x \approx 0.75\) when this term is included, in contrast to the model expectation that crossing occurs at \(x \approx 0.5\) when this orbital angular-momentum effect is neglected. An accurate measurement of \(\Delta d/d\) at \(x\) close to unity would be very interesting if this quantity can be extracted free of uncertainties from nuclear effects (Kulagin and Melnitchouk, 2008a, 2008b) in the neutron structure functions measured from deuterons or \(^3\)He targets.

### A. Generalized parton distributions

Observables in deeply virtual Compton scattering and deeply virtual meson production are sensitive to information about total angular momentum in the nucleon. In these hard exclusive reactions a deeply virtual photon impacts on a nucleon target and a real photon or a meson is liberated from the struck nucleon into the final state, leaving the target nucleon intact. These processes can be described using the formalism of GPDs, involving the Fourier transforms of off-diagonal nucleon matrix elements (Mueller et al., 1994; Ji, 1997a, 1997b; Radyushkin, 1997).

The important kinematic variables are the virtuality of the hard photon \(Q^2\), the momenta \(p - \Delta/2\) of the incident proton and \(p + \Delta/2\) of the outgoing proton, the invariant four-momentum transferred to the target \(t = \Delta^2\), the average nucleon momentum \(P\), the generalized Bjorken variable \(k^+ = xp^+\), and the light-cone momentum transferred to the target proton \(\xi = -\Delta^+ / 2p^+\). In the Bjorken limit, \(\xi\) is related to Bjorken \(x_B\) via \(\xi = x_B/(2 - x_B)\). The generalized parton distributions are defined as the light-cone Fourier transform of the point-split matrix element

\[
\frac{P_+}{2\pi} \int dy^- e^{-ixP^y} \langle p|\bar{\psi}_\alpha(y)\gamma^\mu\psi_\beta(0)|p\rangle^+ \gamma_5 = 0
\]

\[
= \frac{1}{4} \gamma_{\alpha\beta} \left[ H(x, \xi, t) \bar{u}(p) \gamma^\mu u(p) + E(x, \xi, t) \bar{u}(p) \sigma^\mu u(p) \right] \frac{\Delta^+}{2M} u(p) + \frac{1}{4} \gamma_5 \gamma^\mu u(p)
\]

\[
\times \left[ \tilde{H}(x, \xi, t) \bar{u}(p) \gamma^\mu \gamma_5 u(p) + \tilde{E}(x, \xi, t) \bar{u}(p) \gamma^\mu \gamma_5 u(p) \right].
\tag{29}
\]

The physical interpretation of the generalized parton distributions (before worrying about possible renormalization effects and higher order corrections) is the following. Expanding out the quark field operators in Eq. (29) in terms of light-cone quantized creation and annihilation operators one finds that for \(x > \xi\) (\(x < \xi\)) the GPD is the amplitude to take a quark (antiquark) of momentum \(k - \Delta/2\) out of the proton and reinsert a quark (antiquark) of momentum \(k + \Delta/2\) into the proton some distance along the light cone to reform the recoiling proton. In this region the GPD is a simple generalization of the usual parton distributions studied in inclusive and semi-inclusive scattering which are formally defined via light-cone correlation functions; see, e.g., Bass (2005). In the remaining region \(-\xi < x < \xi\) the GPD involves taking out (or inserting) a \(q\bar{q}\) pair with momentum \(k - \Delta/2\) and \(-k + \Delta/2\) (or \(k + \Delta/2\) and \(-k + \Delta/2\)), respectively. Note that the GPDs are interpreted as probability amplitudes rather than densities. The nonforward matrix elements give access to transverse degrees of freedom in the nucleon.

In the forward limit the GPDs \(H\) and \(\tilde{H}\) are related to the parton distributions studied in deep inelastic scattering.

---

We work in the light-cone gauge \(A_+ = 0\) (so the path-ordered gauge link needed for gauge invariance in the correlation function becomes trivial and set equal to 1).

---

\[ H(x, \xi, t)|_{x'=0} = q(x), \quad \bar{H}(x, \xi, t)|_{x'-0} = \Delta q(x), \]  

(30)

whereas the GPDs \( E \) and \( \bar{E} \) have no such analog. Integrating over \( x \) the first moments of the GPDs are related to the nucleon form factors

\[
\int_{-1}^{+1} dx H(x, \xi, t) = F_1(t), \\
\int_{-1}^{+1} dx E(x, \xi, t) = F_2(t), \\
\int_{-1}^{+1} dx \bar{H}(x, \xi, t) = G_A(t), \\
\int_{-1}^{+1} dx \bar{E}(x, \xi, t) = G_P(t). 
\]  

(31)

Here \( F_1 \) and \( F_2 \) are the Dirac and Pauli form factors of the nucleon, and \( G_A \) and \( G_P \) are the axial and induced-pseudoscalar form factors, respectively. (The dependence on \( \xi \) drops out after integration over \( x \).)

GPDs contain vital information about quark total angular momentum in the nucleon. Ji’s sum rule (Ji, 1997b) relates \( J_q \) to the forward limit of the second moment in \( x \) of the spin-independent quark GPDs

\[
J_q = \frac{1}{2} \int_{-1}^{+1} dx \{ H^q(x, \xi, t = 0) + E^q(x, \xi, t = 0) \}. 
\]  

(32)

The gluon “total angular momentum” could then be obtained through

\[
\sum_q J_q + J_g = \frac{1}{2}. 
\]  

(33)

In principle, it could be extracted from precision measurements of the \( Q^2 \) dependence of DVCS at next-to-leading-order accuracy where the quark GPDs mix with glue under QCD evolution or via

\[
J_g = \frac{1}{2} \int_{-1}^{+1} dx \{ H^g(x, \xi, t = 0) + E^g(x, \xi, t = 0) \} 
\]

if the gluon GPD can be accurately measured in more direct experiments. In these equations \( J_q \) and \( J_g \) are defined through the proton matrix elements of the angular-momentum operators in Eq. (26). If information about \( J_g \) can be extracted from experiments, then the corresponding quark orbital angular momentum can be deduced by subtracting the value of the quark spin content \( \Delta q \) extracted from deep inelastic scattering and polarized proton-proton collisions.4

Experimental attempts to access \( J_g \) via Eq. (32) require accurate determination of the two unpolarized GPDs \( H \) and \( E \). Measurements from a proton target are more sensitive to \( J_g \), the total angular momentum carried by up quarks, while the neutron (via a deuteron or \(^3\)He target) is most sensitive to \( J_d \). The experiments require high luminosity to measure the small exclusive cross section, plus measurements over a wide range of kinematics in \( Q^2, x, \) and \( t \) (since sum rule tests and evaluations depend on making reliable extrapolations into unmeasured kinematics). In particular, one has to extrapolate the GPDs to \( t = 0 \). One also needs reliable theoretical technology to extract the GPDs from the measured cross sections.

GPDs appear in the amplitudes for DVCS and hard exclusive meson production as convolutions with the hard-scattering coefficient and only these so-called Compton form factors (CFF) are experimentally accessible. Measuring photon and also meson production in the final state gives access to different flavor combinations of GPDs, such as in semi-inclusive DIS. However, meson production is more sensitive to QCD radiative corrections and power corrections in \( 1/Q \), and reliable theoretical description requires larger values of \( Q^2 \) compared to DVCS. Channels particularly sensitive to gluons in the proton are hard exclusive vector meson production where both quark and gluon GPDs appear at lowest order in the strong coupling constant. There is a challenging program to disentangle the GPDs from the formalism and to undo the convolution integrals which relate the GPDs to measured cross sections. In practice, the approach used is to constrain models of GPDs against experimental data in measured kinematics. These models are then integrated to obtain the Ji moments of \( J_q \) and \( J_g \), which may then be compared to the predictions of QCD inspired model plus lattice calculations; see Table V.

For the remainder of this discussion we focus on deeply virtual Compton scattering.

### 1. Deeply virtual Compton scattering

Measurements of hard exclusive processes are much more challenging than traditional inclusive and semi-inclusive scattering experiments. These exclusive processes require a difficult full reconstruction of final-state particles and their cross sections are usually small, demanding high luminosity machines.

DVCS experiments have to be careful to choose the kinematics so as not to be saturated by a large Bethe-Heitler (BH) background where the emitted real photon is radiated from the incident lepton rather than from the proton target; see Fig. 14.

Most of the DVCS program so far has focused on the DVCS-BH interference term. Use of different combinations of beam and target polarization plus changing the electric charge of the incident lepton beam gives maximum access to most combinations of DVCS observables. Measurement of the DVCS-BH interference term, see Eq. (34), allows one to measure not only the size of the DVCS amplitude

\[ x \times C_f(x) \times C_a(x) \times F_L(x) \times F_A(x) \times F_P(x) \times F_L(x) \times F_A(x) \times F_P(x) \]

\[ \Gamma(x, \xi, t) \times \Gamma(x, \xi, t) \times \Gamma(x, \xi, t) \times \Gamma(x, \xi, t) \times \Gamma(x, \xi, t) \times \Gamma(x, \xi, t) \times \Gamma(x, \xi, t) \times \Gamma(x, \xi, t) \]

\[ \Delta(x) \times \Delta(x) \times \Delta(x) \times \Delta(x) \times \Delta(x) \times \Delta(x) \times \Delta(x) \times \Delta(x) \]

4We note recent discussion of a \( J = 0 \) fixed pole contribution to DVCS (Brodsky, Llanes-Estrada, Londergan, and Szczepaniak, 2009; Brodsky, Llanes-Estrada, and Szczepaniak, 2009), which corresponds to a \( x \delta(x) \) term in the GPD \( H \) and affects the 1/\( x \) moment of this GPD although not the sum rules in Eqs. (31) and (32). The same fixed pole also contributes to the Schwinger term sum rule for the 1/\( x \) moment of the longitudinal structure function \( F_L \) (Broadhurst, Gunion, and Jaffe, 1973).
but also its phase; that is, it gives separate information about the real and imaginary parts of the Compton form factors.

Pioneering measurements of DVCS have been performed at DESY (HERMES, H1, and ZEUS) and JLab (Hall A and Hall B), which complement each other in the covered kinematic phase space and the extracted observables.

The experiments use different measurement techniques to access exclusive reactions. The HERA collider experiments at DESY, H1, and ZEUS as well as CLAS (Hall B at JLab) have the advantage of nearly hermetic spectrometers, whereas the fixed-target experiments HERMES and JLab Hall A had to deal with the restrictions caused by incomplete event reconstruction due to their forward spectrometers. Hall A and HERMES successfully employed the so-called missing mass technique together with careful background subtraction (Camacho et al., 2006; Airapetian et al., 2008b). For Hall A the low beam energy and high resolution spectrometer allowed one to resolve pure elastic scattering from associated production with an excited nucleon in the final state. The latter contribution was treated as part of the signal in HERMES results. Very recently, beam-spin asymmetries for a pure DVCS sample have also been reported by HERMES (Airapetian et al., 2012c). In the fixed-target experiments the spin-dependent DVCS cross sections have been explored using longitudinally polarized lepton beams with longitudinally (JLab and HERMES) and transversely (HERMES) polarized targets. HERMES also took advantage of the available different beam charges.

JLab experiments focus on kinematics dominated by valence quarks. Data from Hall A suggest leading twist-2 dominance of DVCS even at the relatively low $Q^2$ of $1.5\text{–}2.3\text{ GeV}^2$ (Camacho et al., 2006).

The HERA collider experiments H1 and ZEUS measured the DVCS cross section close to the forward direction with $\xi < 10^{-2}$, integrated over its azimuthal dependence, in an $x_B$ range where two-gluon exchange plays a major role in addition to the leading-order quark-photon scattering process. Figure 15 shows the cross-section differential in $t$ for different ranges in $Q^2$ measured by H1 (Aaron et al., 2008).

\[
\frac{d\sigma}{dt}(p \rightarrow \ell \gamma p) = d\sigma_{UU}^{HH} + e_\ell d\sigma_{UL}^{D} + d\sigma_{UL}^{DVCS} + P_T S_L d\sigma_{LL}^{HH} + e_\ell P_T S_L d\sigma_{LL}^{D} + P_T d\sigma_{LL}^{DVCS} + P_T d\sigma_{LT}^{D} \\
+ e_\ell P_T S_T d\sigma_{LT}^{D} + P_T S_T d\sigma_{LT}^{DVCS} + e_\ell P_T d\sigma_{LT}^{DVCS} + e_\ell S_L d\sigma_{UL}^{DVCS} + e_\ell S_T d\sigma_{UT}^{DVCS}.
\]

(34)

Here the first subscript $U$, $L$ on $d\sigma$ indicates an unpolarized or longitudinally polarized lepton beam and the second subscript $U$, $L$, $T$ denotes an unpolarized, longitudinally, or transversely polarized proton target; $P_T$ is the lepton beam polarization; and $S_L$ and $S_T$ denote longitudinal and transverse proton polarization. Of particular interest is also the dependence on the sign of the charge of the beam lepton $e_\ell$, which allows one to disentangle contributions from the pure interference term and the DVCS term as pioneered by Airapetian et al. (2007b, 2008b). The various cross-section terms depend on the azimuthal angle $\phi$ between the lepton scattering plane and the photon production plane, and, in the case of a transversely polarized proton target, also on the azimuthal angle $\phi_\gamma$ between the lepton plane and the transverse target spin vector. Equation (34) indicates the large variety of observables accessible with polarized beams and/or targets.

As an example for the azimuthal dependence of the cross section we give the expression for the interference term for the case of an unpolarized target and polarized beam (Belitsky, Mueller, and Kirchner, 2002).

FIG. 15 (color online). The $t$ dependence of the DVCS cross section for several values of $Q^2$ as measured by H1 and ZEUS. The curves are results of fits of the form $e^{-b|t|}$ with $b$ being related to the transverse extension of partons in the proton at a given $x$ and $Q^2$ (see the text). The inner error bar represents the statistical uncertainty; the full bar represents the quadratic sum of statistical and systematic uncertainties.

and ZEUS (Chekanov et al., 2009). The data are well described by the exponential behavior $d\sigma/dt \propto e^{-b|t|}$. The distribution of partons in the transverse plane is then obtained from this dependence by a Fourier transform with respect to $\Delta_T$ [the transverse-momentum shift in Eq. (29)] $F(b, x, Q^2) \propto \int d^2\Delta_T \exp(-ib\Delta_T)q|d\sigma/dt|$ (Diehl, 2002; Burkardt, 2003). The impact parameter provides an estimate of the transverse extension of the partons probed during the hard process. While DVCS data provide information about the transverse distribution of quarks in the proton, data on exclusive heavy vector meson production ($J/\Psi$ or $\Upsilon$) describe the transverse distribution of glue at specific values of $x$. The full DVCS cross section reads (Diehl and Sapeta, 2005)
The proportionality involves a kinematic factor and the lepton propagators of the BH process; \( \lambda \) is the helicity of the incoming lepton. The Fourier coefficients \( c_n^l \) provide an experimental constraint on the real part of the Compton form factor and \( s_n^k \) on the imaginary part. Their relation to linear combinations of Compton form factors and hence to the respective GPDs is listed in Table VI. A specific Fourier coefficient can be accessed experimentally by weighting the cross section with the respective azimuthal modulation.

\[
I \propto -e_i \left( \sum_{n=0}^{3} c_n^l \cos(n \phi) + \lambda \sum_{n=1}^{2} s_n^k \sin(n \phi) \right) \quad (35)
\]

The DVCS-BH interference term was extracted by varying the electric charge of the incident lepton (HERMES) and studying polarization observables, varying the beam or target helicity (JLab and HERMES). JLab experiments focused on studying their accessible observables fully differentially. HERMES explored the advantages of using simultaneously polarization and charge observables to cleanly isolate the interference term and obtained the most complete set of DVCS observables measured so far providing access to all interference terms listed in Eq. (34).

Figure 16 shows a summary of the HERMES DVCS measurements with polarized proton and deuterium targets at their average kinematics (Airapetian et al., 2008b, 2009c, 2010b, 2010d, 2011a, 2011b, 2012b, 2012c). Here \( A_C \) is the charge asymmetry and \( A_{XY} \) are the polarization-dependent asymmetries with \( X \) and \( Y \) indicating the beam and target polarization, respectively, which could be longitudinal (L) or transverse (T). The subscript \( I \) indicates an extraction of the pure interference term. The measured asymmetries are subjected to a harmonic expansion with respect to the azimuthal angle(s) as given by the superscript of \( A_{XY} \) in the figure. These data denoted by squares in Fig. 16 show results extracted from a DVCS sample with kinematically complete event reconstruction (Airapetian et al., 2012c). The dependence on the kinematic variables \( t, Q^2 \), and \( x_B \) was explored for each observable.

An example of the high statistics data from JLab is shown in Fig. 17 for the beam-spin asymmetry \( A_{LU} \) measured fully differentially by CLAS (Girod et al., 2008). The presented data contain an admixture of the \( A_{LU}^{\text{int}} \) and \( A_{LU}^{\text{DVCS}} \) contributions from the interference and pure DVCS terms, which

<table>
<thead>
<tr>
<th>HERMES DVCS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hydrogen</td>
</tr>
<tr>
<td>Deuterium</td>
</tr>
<tr>
<td>Hydrogen Pure</td>
</tr>
</tbody>
</table>

**FIG. 16** (color online). Overview of all DVCS azimuthal asymmetry amplitudes measured at HERMES with proton and deuterium targets, given at the average kinematics. The inner error bar represents the statistical uncertainty; the full bar represents the quadratic sum of statistical and systematic uncertainties.

**FIG. 17** (color online). The leading beam-spin asymmetry amplitude \( a(t) = A_{LU}^{(t)} \) differential in \( t, x, \) and \( Q^2 \) as measured by CLAS, from Girod et al. (2008). An earlier CLAS measurement (Stepanyan et al., 2001) is indicated by the square. The open triangles represent the cross-section data from Hall A (Camacho et al., 2006). Error bars are statistical errors only.
cannot be separated here. CLAS also provides measurements of \(A_{UL}^{\sin^6}\) and \(A_{UL}^{\sin^2\theta}\) (Chen et al., 2006).

2. The quest for orbital angular momentum and GPD parametrizations

Of the two GPDs \(H\) and \(E\) entering Ji’s sum rule, Eq. (32), measurements with unpolarized targets but longitudinally polarized beams and also beam-charge asymmetries are mainly sensitive to \(H\). As indicated in Table VI, transverse target polarization provides kinematically suppressed access to \(E\). The GPD \(E\) is essentially unknown. In contrast to \(H\), it is not related to a deep inelastic parton distribution in the forward limit; \(E\) describes helicity flip at the proton vertex and requires finite orbital angular momentum in the nucleon. Contributions from \(E\) to most DVCS observables are damped by kinematic factors \(~[t]/M_p^2\), with the average \([t]\) value generally much smaller than 1 GeV\(^2\) in the experiments. To access \(E\) requires DVCS and/or vector meson production asymmetry measurements with transversely polarized nucleon targets. It may also be accessed through the beam polarization dependence of DVCS with a neutron target because of the different size of the form factors for the neutron (Belitsky, Mueller, and Kirchner, 2002). Measurements have been performed already for all channels (Mazouz et al., 2007; Airapetian et al., 2008b, 2009a; Adolph et al., 2012a). Despite the lack of precision for these observables, attempts to extract information about quark total angular momentum have been performed by fitting theoretical models of GPDs to the DVCS measurements (Mazouz et al., 2007; Airapetian et al., 2008b). Although this analysis is very model dependent, the results agree well with model and lattice expectations, e.g., the calculations reported in Table V. For example, within the model of Vanderhaeghen, Guichon, and Guidal (1999), JLab Hall A DVCS measurements from the neutron were interpreted to give \(J_u + J_d/5 = 0.18 \pm 0.14\) (expt) (Mazouz et al., 2007), whereas HERMES results from the proton gave \(J_u + J_d/2 = 0.49 \pm 0.17\) (expt) (Airapetian et al., 2008b) in the same model.

To go further and perform global fits of GPDs to hard exclusive observables one faces several challenging theoretical issues. Parametrizations of GPDs have to deal with two longitudinal variables instead of one plus the \(t\) dependence of DVCS. It is also not yet known whether relatively simple and smooth functions such as those used in QCD fits to deep inelastic data are sufficient to describe GPDs. A reliable parametrization of GPDs might therefore require a larger number of moments than employed in usual QCD parton descriptions. In addition, the dependence of the functions on the variable \(x\) is not directly accessible as \(x\) represents a mute variable which is integrated over. In the interpretation of DVCS observables one has to deal with complex amplitudes; the GPDs are embedded in the Compton form factors which relate to the measured cross sections. Despite these complications and the early stage of global fitting for GPDs, many results have been obtained in recent years fitting to different hard exclusive scattering data. Interested readers are referred to the original literature by Vanderhaeghen, Guichon, and Guidal (1999), Goloskokov and Kroll (2008), Guidal (2010), Kumericki and Mueller (2010), and Goldstein, Hernandez, and Liuti (2011). This phenomenology is complemented by progress in lattice QCD calculations of GPD moments (Brommel et al., 2007; Göckeler et al., 2007; Hägler et al., 2008).

B. Transversity, transverse-momentum-dependent distributions, and fragmentation functions

Striking single-spin asymmetries associated with spin-momentum correlations (expected with parton orbital angular momentum) were first observed in the 1970s. Using a 12 GeV polarized proton beam from the Argonne National Laboratory Zero Gradient Synchrotron on a fixed target, up to 40% more positive pions were produced left of the beam when the beam was polarized up, and up to 20% more negative pions were produced to the right of the beam (Klem et al., 1976). These measurements were confirmed by similar experiments (Dragoset et al., 1978; Antille et al., 1980; Apokin et al., 1990; Saroff et al., 1990), but it was not until the 1990s that a theoretical framework was developed to attempt interpreting them.

Single-spin asymmetries have now also been observed in proton-proton collisions at RHIC, where they reach up to ~40%, and in lepton-nucleon collisions at COMPASS, HERMES, and JLab, where they are typically 5%-10%. Single-spin asymmetries for hadron production from transversely polarized targets tell us about spin-orbit coupling in the nucleon and/or in the fragmentation process. Transverse-momentum-dependent distributions simultaneously describe the dependence on longitudinal momentum fraction of the parton within the parent hadron as well as the parton’s transverse momentum. Similarly, transverse-momentum-dependent fragmentation functions describe the dependence on longitudinal momentum fraction of the produced hadron with respect to the scattering parton as well as the hadronic transverse momentum with respect to the jet axis.

We next introduce these distributions and fragmentation functions and discuss their phenomenology. In QCD there are eight leading-twist quark TMDs. These are listed in Table VII and discussed by Mulders et al. (1996) and Bacchetta et al. (2007). The three distributions highlighted in boldface survive integration over transverse momentum \(k_t\). The yield the unpolarized parton distribution \(f_{i\perp}(x, k_t)\), the spin-dependent parton distribution \(g_i(x, k_t)\), and the transversity distribution \(h_i(x, k_t)\). The other five distributions do not survive integration over \(k_t\). They describe correlations between the quark transverse momentum with the spin of the quark and/or the spin of the parent nucleon, viz. spin-orbit correlations. The three TMDs denoted by \(h\) describe the distribution of transversely polarized partons. They are

<table>
<thead>
<tr>
<th>(N/q)</th>
<th>(U)</th>
<th>(L)</th>
<th>(T)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(U)</td>
<td>(f_{i\perp})</td>
<td>(g_{i\perp})</td>
<td>(h_{i\perp})</td>
</tr>
<tr>
<td>(L)</td>
<td></td>
<td>(g_{i\perp})</td>
<td>(h_{i\perp})</td>
</tr>
<tr>
<td>(T)</td>
<td>(f_{i\perp})</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
TABLE VIII. Experimental access to the leading-twist TMD distributions in SIDIS with unpolarized \((U)\), longitudinally \((L)\), or transversely polarized \((T)\) beam (modulation first subscript) and/or target (modulation second subscript).

<table>
<thead>
<tr>
<th>Modulation</th>
<th>Combination Distribution name</th>
<th>(\sqrt{s}) (GeV)</th>
<th>Target type</th>
<th>Observed hadron types</th>
<th>Measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\sin(\phi + \phi_S)_{UT})</td>
<td>(h_1^+ \otimes H_1^\perp) Transversity</td>
<td>3.5</td>
<td>7.4</td>
<td>(p)</td>
<td>(\pi^\pm, K^\pm)</td>
</tr>
<tr>
<td>(\sin(\phi - \phi_S)_{UT})</td>
<td>(f_{1T}^\perp \otimes D) Sivers</td>
<td>18</td>
<td>(d)</td>
<td>(h^\pm, \pi^\pm, K^\pm, K^0)</td>
<td>Prelim. Kotzinian (2007)</td>
</tr>
<tr>
<td>(\cos(2\phi)_{UU})</td>
<td>(h_1^+ \otimes H_1^\perp) Boer-Mulders</td>
<td>3.5</td>
<td>7.4</td>
<td>(p)</td>
<td>(\pi^\pm, K^\pm)</td>
</tr>
<tr>
<td>(\sin(3\phi - \phi_S)_{UT})</td>
<td>(h_1^+ \otimes H_1^\perp) Pretzelosity</td>
<td>18</td>
<td>(d)</td>
<td>(h^\pm)</td>
<td>Prelim. Parsamyan (2011)</td>
</tr>
<tr>
<td>(\sin(2\phi)_{UL})</td>
<td>(h_1^+ \otimes H_1^\perp) Worm-gear 1</td>
<td>3.5</td>
<td>7.4</td>
<td>(p)</td>
<td>(\pi^\pm, K^\pm)</td>
</tr>
<tr>
<td>(\cos(\phi - \phi_S)_{LT})</td>
<td>(f_{1T}^\perp \otimes D) Worm-gear 2</td>
<td>18</td>
<td>(d)</td>
<td>(h^\pm)</td>
<td>Prelim. Kotzinian (2007)</td>
</tr>
</tbody>
</table>

Chiral-odd distributions and appear only in observables involving two chiral-odd partners, such as Drell-Yan processes (two chiral-odd parton distributions) or SIDIS (chiral-odd parton distribution and the Collins fragmentation function discussed later). The three distributions \(f_{1T}^\perp\) (the Sivers distribution), \(h_1^+\) (the Boer-Mulders distribution), and \(h_1^\perp\) (pretzelosity) require orbital angular momentum in the nucleon since they involve a transition between initial and final nucleon states whose orbital angular momentum differs by \(\Delta L^\perp = \pm 1\) (Sivers and Boer-Mulders) or \(\Delta L^\perp = \pm 2\) (pretzelosity). The “worm-gear” functions \(h_1^\perp\) and \(g_{1T}^\perp\) link two perpendicular spin directions and are also connected to quark orbital motion inside nucleons.

Transverse-momentum distributions have been studied most in semi-inclusive DIS experiments where they appear in combination with the usual unpolarized fragmentation function \(D(z, p_T)\) or, in the case of the chiral-odd TMD distributions, with a chiral-odd Collins fragmentation function \(H_1^\perp(z, p_T)\) discussed in Sec. VII.B.2. One measures the azimuthal distribution of the produced final-state hadron with respect to the virtual-photon axis. Each species of TMD comes with a different angular modulation in the semi-inclusive cross section allowing it to be projected out to yield information about the different spin-momentum correlations (Bacic, 2004). All these leading-twist TMDs have been measured in semi-inclusive DIS over the last decade. However, several have been the focus of more intense studies and we focus on those here.

The different modulation combinations are listed in Table VIII together with present experimental measurements. Results quoted at \(\sqrt{s} = 18\) GeV are from COMPASS, 7.4 GeV from HERMES, and 3.5 GeV from JLab. Here \(\phi\) is the angle between the lepton direction and the plane spanned by the exchanged photon and tagged final-state hadron, e.g., a high-energy meson; \(\phi_S\) is the angle between the lepton direction and the transverse nucleon target spin. The convolution is taken over the involved transverse momenta of the quark and the hadron produced in the fragmentation process.

When one projects out the terms with different azimuthal angular dependence summarized in Table VIII, the COMPASS, HERMES, and JLab data suggest that the Sivers, Collins, and Boer-Mulders effects are all present in the proton target data; see Secs. VII.B.1 and VII.B.2. JLab data from CLAS reveal a clear signal for the worm-gear-1 distribution; there is some hint for a nonzero worm-gear-2 distribution (with low significance) and so far no significant signal for pretzelosity in the proton. For the deuteron target, there is evidence for a Boer-Mulders effect from COMPASS and HERMES. The Sivers, Collins, worm-gear, and pretzelosity effects are all consistent with zero in the deuteron target data. The Collins and Sivers effects observed in the proton data therefore contain a predominant isovector contribution.

We next focus on the Sivers, Boer-Mulders, and Collins effects.

1. The Sivers and Boer-Mulders TMD distributions

The Sivers distribution was first proposed by Sivers (1990) in an attempt to explain the large transverse single-spin asymmetries observed in the 1970s and 1980s. It describes the correlation between the transverse momentum \(k_T\) of the struck quark and the spin \(S\) and momentum \(p\) of its parent nucleon.
The $k_t$ dependence means that the Sivers distribution is sensitive to nonzero parton orbital angular momentum in the nucleon, although the mapping from Sivers observables to quark (and gluon) orbital angular momentum is so far model dependent with present theoretical technology.

The Sivers distribution has the interesting property that it is odd under time reversal. Because of this feature, such a correlation was believed to be forbidden for more than a decade. Then Brodsky, Hwang, and Schmidt (2002a, 2002b) showed that, with initial- or final-state interactions, the Sivers effect could be nonzero in QCD processes. Final-state interactions in SIDIS can generate the azimuthal asymmetry before the quark fragments into hadrons. Shortly afterward, Collins (2002) realized that initial-state color interactions in the case of Drell-Yan and final-state interactions in the case of SIDIS would lead to a process-dependent correlation. Collins (2002) showed that, with initial- or final-state interactions, the Sivers effect could be nonzero in QCD processes. Final-state interactions in SIDIS can generate the azimuthal asymmetry before the quark fragments into hadrons. Shortly afterward, Collins (2002) realized that initial-state color interactions in the case of Drell-Yan and final-state interactions in the case of SIDIS would lead to a process-dependent sign difference in the Sivers distribution. SIDIS measurements (Airapetian et al., 2005b, 2009b; Alekseev et al., 2010b; Pesaro, 2011; Qian et al., 2011; Adolph et al., 2012c) suggest sizable asymmetries at the level of about 5%–10% for a proton and a neutron target, while Drell-Yan measurements are planned for the future.

A qualitative picture of the Sivers distribution can already be deduced from SIDIS measurements. The nonzero amplitudes shown in Figs. 18 and 19 were obtained with a proton target. For HERMES the amplitude includes a kinematic factor depending on the ratio of transverse-to-longitudinal photon flux, which in the COMPASS data is divided out. Since scattering off $u$ quarks dominates these data due to the quark charge factor, the positive Sivers amplitudes for $\pi^+$ (and $h^+$, which is dominated by the pion yield) suggest a large and negative Sivers function for up quarks. The vanishing amplitudes for $\pi^-$ ($h^-$) require cancellation effects, e.g., from a $d$ quark Sivers distribution opposite in sign to the $u$-quark Sivers distribution. These cancellation effects between Sivers distributions for up and down quarks are supported by the vanishing Sivers amplitudes extracted from deuteron data by the COMPASS Collaboration (Ageev et al., 2007;Aleksseev et al., 2009a). An interesting facet of the HERMES data is the magnitude of the $K^+$ amplitude, which is nearly twice as large as that of $\pi^+$ (Airapetian et al., 2009b). Again, on the basis of $u$ quark dominance, one might naively expect that the $\pi^+$ and $K^+$ amplitudes should be similar. Their difference in size may thus point to a significant role of other quark flavors, e.g., sea quarks. A sizable Sivers amplitude for $\pi^+$ was also recently reported by JLab Hall A (Qian et al., 2011) for measurements with a $^3$He (neutron) target. In that data a negative Sivers amplitude for $\pi^+$ was found which independently supports a $d$-quark Sivers distribution opposite in sign to the $u$-quark one.

The Boer-Mulders distribution (Boer and Mulders, 1998) describes the correlation between transversely polarized quarks in an unpolarized nucleon and the quarks’ transverse momentum $s_q \cdot (k \times \hat{p})$, where $s_q$ denotes the spin of the quark and might hence yield unexpected spin effects even in an unpolarized nucleon. It is similar to the Sivers distribution in that it is $T$ odd. However, it is also chiral odd and hence must be probed in conjunction with a second chiral-odd function. For Drell-Yan production, the second function is a Boer-Mulders distribution in the second incident hadron. For SIDIS, the Collins fragmentation function described below is involved. Like the $T$-odd Sivers distribution, the Boer-Mulders distribution is also expected to change sign between Drell-Yan production and SIDIS. Future experimental effort is planned to test this QCD prediction.

Azimuthal distributions sensitive to the Boer-Mulders distribution were originally measured in Drell-Yan experiments (Falciano et al., 1986; Guanziroli et al., 1988; Conway et al., 1989; Zhu et al., 2007, 2009). The SeaQuest fixed-target Drell-Yan experiment currently underway at Fermilab (Reimer, 2007) expects to be sensitive to the Boer-Mulders distribution at high $x$. In SIDIS the distinctive pattern of Boer-Mulders modulations for oppositely charged pions and for pions and kaons was recently reported by HERMES (Airapetian et al., 2012a). The amplitudes for kaons are larger in magnitude than the amplitudes for pions. The amplitudes for the negative pions have the opposite sign to the amplitudes for negative kaons. This hints at a significant contribution from sea quarks, in particular, from strange quarks. Measurements of the Boer-Mulders amplitudes were also reported by COMPASS for unidentified hadrons (Sbrizzai, 2011) and by CLAS for pions (Ospienko et al., 2009). The interpretation of the SIDIS amplitudes for the Boer-Mulders distribution is, however, complicated by contributions from the twist-4 Cahn effect (Cahn, 1978, 1989) which have been estimated to be sizable even at
COMPASS kinematics (Anselmino, Boglione et al., 2007). The Cahn effect accounts for the parton intrinsic transverse momenta in the target nucleon and the fact that produced hadrons might acquire transverse momenta during the fragmentation process. Theoretical estimates of the Boer-Mulders effect are still plagued by large uncertainties, mainly related to the insufficient knowledge of the transverse-momentum dependence of the unpolarized distribution \( f_1(x, k_t) \) and fragmentation function \( D(z, p_t) \).

2. The Collins TMD fragmentation function

The Collins TMD fragmentation function describes a spin-momentum correlation in the hadronization process \( s_q \cdot (k_q \times p_t) \) with a hadron produced in fragmentation having some transverse momentum \( p_t \) with respect to the momentum direction \( k \) of a transversely polarized fragmenting quark with spin \( s_q \) (Collins, 1993; Collins, Heppelmann, and Ladinsky, 1994). The Collins fragmentation function has been investigated in semi-inclusive lepton-nucleon scattering experiments having some transverse momentum \( q \) (Seidl et al., 2008). The HERMES asymmetries. The magnitude of the effect is approximately 5\%-10\%, similar to that found for the Sivers asymmetry.

For \( e^+ e^- \) annihilation the chiral-odd Collins fragmentation function enters with a second Collins function in the production but comes with opposite sign. This hints at an underlying factorization of the unpolarized distribution \( f_1(x, k_t) \) and the transversity distribution \( h_T(x, \theta, t) \) for the transverse polarization of quarks within a transversely polarized nucleon. Along with the unpolarized and helicity distributions, it survives integration over partonic transverse momentum and is thus a collinear distribution.

3. Probing transversity

The transversity distribution introduced in Sec. II describes the transverse polarization of quarks within a transversely polarized nucleon. Along with the unpolarized and helicity distributions, it survives integration over partonic transverse momentum and is thus a collinear distribution.

The first moment of the transversity distribution is proportional to the nucleon’s \( C \)-odd tensor charge, viz.

\[
\delta q = \int_0^1 dx h_T^c(x) \quad \text{with} \quad \langle p, s | \bar{q} i \gamma_5 q | p, s \rangle = (1/M)(s_{\mu} p_{\nu} - s_{\nu} p_{\mu}) \delta q. \quad (37)
\]

FIG. 21 (color online). Collins asymmetries for the double ratios of like-sign (L), unlike-sign (U), and any charged (C) pion pairs from Belle (Seidl et al., 2008). \( A^{UL} \) and \( A^{UC} \) are sensitive to different combinations of the favored and unfavored Collins fragmentation functions. The bands indicate the systematic uncertainties.

FIG. 22 (color online). Collins asymmetries for unidentified charged hadrons measured by COMPASS with a proton target (Adolph et al., 2012b). The hadron yield is dominated by pions. Note that a different definition of the Collins angle results in amplitudes with the opposite sign compared to other measurements. The bands indicate the systematic uncertainties.

The difference between the transversity and helicity spin distributions reflects the relativistic character of quark motion in the nucleon. In bag models this effect is manifest as follows. The lower component of the Dirac spinor enters the relativistic spin depolarization factor with the opposite sign to $\Delta q$ because of the extra factor of $\gamma_{\mu}$ in the tensor charge (Jaffe and Ji, 1992). The relativistic bag depolarization factor mentioned in Sec. VI becomes 0.83 for transversity in contrast to 0.65 for helicity and the nucleon’s axial charges. In leading-order QCD the transversity distributions are bound by Soffer’s inequality $|h_{1}^q(x, Q^2)| \leq \frac{1}{2}[(q + \bar{q})(x, Q^2 + \Delta q(x, Q^2))]$ (Soffer, 1995). QCD-motivated fits to transversity observables are reported by Anselmino et al. (2009b), which also review the comparison to model predictions.

Transversity is measured through the Collins effect and also in dihadron production, where the chiral-odd partner of $h_{1}^q$ is given by the dihadron fragmentation function $H_{1}^{gq}$ (Collins, Heppelmann, and Lads经贸, 1994; Bianconi et al., 2000; Bacchetta, Courtoy, and Radici, 2011). This describes how the transverse spin of the fragmenting quark is transferred to the relative orbital angular momentum of the hadron pair. Consequently, this mechanism does not require transverse momentum of the produced hadron pair. Standard collinear factorization applies allowing one to study the transversity distribution without having to worry about solving convolution integrals over transverse momentum or issues of TMD factorization and evolution.

Pioneering measurements of two-pion production in polarized semi-inclusive DIS by HERMES (Airapetian et al., 2008a) and COMPASS (Adolph et al., 2012f) reveal a sizable effect and have already been employed for an extraction of transversity (Courtoy, Bacchetta, and Radici, 2012). First measurements of azimuthal correlations of two pion pairs in back-to-back jets in $e^+e^-$ annihilation related to the dihadron fragmentation function have just become available from Belle (Vossen et al., 2011) and a first extraction of the dihadron fragmentation function was performed by Courtoy, Bacchetta, Radici, and Bianconi (2012).

4. Current status and recent progress with TMD distributions

There has been considerable progress in the understanding of intrinsic transverse momentum and spin-momentum correlations in QCD over the past decade, motivated by the theoretical breakthroughs regarding $T$-odd TMD distributions (Brodsky, Hwang, and Schmidt, 2002a, 2002b; Collins, 2002) and by a vast program of theoretical and experimental activity.

Recently Collins (2011) gave definitions of TMD distributions which allow QCD evolution to be applied rigorously for the first time with separately identifiable TMD distributions and fragmentation functions. Building upon this progress, the evolution of previously unevolved models and fits has now been published for unpolarized TMD distributions and fragmentation functions (Aybat and Rogers, 2011b) and the Sivers distribution (Aybat et al., 2012). QCD evolution is just starting to be applied to phenomenological studies (Aybat, Prokudin, and Rogers, 2012), which will be a major step forward in interpreting and comparing results from different experiments. The new definitions of TMD distributions also recently made possible a determination of the hard parts for SIDIS and Drell-Yan at next-to-leading order (Aybat and Rogers, 2011a), which should lead to improved phenomenology.

Much effort has been dedicated to phenomenological extractions of TMDs and parametrizations of the Sivers distribution from SIDIS data; see, e.g., Anselmino et al. (2009a, 2011), with QCD evolution now starting to be considered (Anselmino, Boglione, and Melis, 2012). One parametrization of the Sivers function includes both semi-inclusive deep inelastic and proton-proton data (Kang and Prokudin, 2012), modulo issues related to factorization breaking (Rogers and Mulders, 2010) discussed below. Fits to the Collins TMD fragmentation function have been performed using both $e^+e^-$ and SIDIS data as input (Efremov, Goeke, and Schweitzer, 2006; Anselmino et al., 2007). The Boer-Mulders distribution has been extracted based on Drell-Yan (Zhang et al., 2008; Lu and Schmidt, 2010) as well as SIDIS data (Barone, Melis, and Prokudin, 2010).

These first phenomenological fits to TMD observables have been performed using a simple Gaussian ansatz for the transverse-momentum dependence of quarks in the nucleon and fragmentation functions. For example, the Sivers function in Eq. (36) was parametrized in the fits by the product of the unpolarized distribution $f_{q/p}(x, k_t)$ with an $x$-dependent factor and an $x$-independent Gaussian $\sim (k_t/M_1)e^{-k_t^2/M_1^2}$ containing all the $k_t$ dependence. While the Gaussian ansatz is unstable with respect to QCD evolution with increasing $Q^2$, the method does provide a reasonable fit to present data with values $\langle k_t^2 \rangle = 0.25 \text{ GeV}^2$ and $\langle p_t^2 \rangle = 0.20 \text{ GeV}^2$ taken from fits to the Cahn effect in unpolarized scattering (Anselmino et al., 2005). A longer-term goal for TMD experiments is to observe deviation from Gaussian behavior for transverse-momentum dependence. With extensive unpolarized Drell-Yan and weak boson production data available over scales from $-4 \text{ GeV}^2$ to $M_2^2$, new fits of unpolarized TMD distributions are quite promising as a means to test the $Q^2$ evolution of TMD distributions as well as to learn more about the shape of the distributions in $k_t$.

Lattice calculations of the Sivers and Boer-Mulders distributions have been performed (Göckeler et al., 2007; Hägler et al., 2009; Musch et al., 2012). There have also been efforts in recent years to implement TMDs in Monte Carlo event generators (Bianconi, 2011; Hautmann, Hentschinski, and Jung, 2012). Models can provide helpful insight into TMD distributions, and a wealth of different model calculations have been explored and published. See Avakian et al. (2009), Lorce and Pasquini (2011), Pasquini and Schweitzer (2011), and Bacchetta (2012) for recent discussions of models related to TMD distributions, including attempts to address the relationship between TMD distributions and orbital angular momentum in the nucleon.

5. Proton-proton asymmetries and TMD-factorization breaking

Despite the fact that the large transverse single-spin asymmetries observed in hadronic scattering originally inspired the development of TMDs, inclusive hadron production in $p + p$ scattering cannot be cleanly separated into Sivers, Collins, or other contributions as is possible in SIDIS. In
recent work Rogers and Mulders (2010) argued that the TMD framework is not valid in the case of hadroproduction of hadrons, as factorization does not hold. While the short-distance (perturbative) components are still believed to factorize from the long-distance (nonperturbative) ones, the long-distance components become entangled and no longer factorize from one another into independent TMD distributions and/or fragmentation functions. What is particularly interesting is that the factorization breaking effects are relevant in precisely the kinematic regime where a parton description is generally expected to apply. It will be exciting to see this experimentally tested in the upcoming years, exploring long-distance quantum entanglement effects in QCD. In the longer-term future, it may be possible to develop well-defined functions within the framework of pQCD which describe the correlations between the partons in the incoming and/or outgoing hadrons.

In the meantime, single-spin asymmetries for forward meson production in $p + p$ collisions have been shown to remain large across a very wide range of center-of-mass energies (Adams et al., 1991, 1996; Allgower et al., 2002; Abelev et al., 2008a; Arsene et al., 2008) and up to the highest measured $p_T$ of $\sim 5$ GeV (Koster, 2012). As shown in Fig. 23, the transverse single-spin asymmetries in charged pion production as a function of Feynman $x$ are remarkably similar from $\sqrt{s} = 4.9$ GeV all the way up to 62.4 GeV measured by the BRAHMS experiment at RHIC.

At higher energies and, in particular, at $p_T$ values large enough to serve as a hard scale, one can try to interpret these phenomena utilizing the tools of pQCD. With no explicitly measured scale sensitive to the partonic transverse momentum in inclusive single-spin asymmetries, a more appropriate framework than TMD distributions in which to interpret the asymmetries may be a collinear, twist-3 picture (Efremov and Teryaev, 1982, 1985; Qiu and Sterman, 1998). A relationship between the TMD and the collinear, twist-3 frameworks was laid out by Ji et al. (2006).

Surprises continue to emerge from these kinds of measurements, with large asymmetries for negative kaons as well as antiprotons from BRAHMS (Arsene et al., 2008) suggesting that the pion asymmetries are not a valence quark effect as previously believed, and a recent hint from STAR (Adamczyk et al., 2012c) that the asymmetry for $\eta$ mesons may be larger than that of neutral pions.

**VIII. FUTURE PROJECTS**

A new program of dedicated experiments is planned to investigate key open questions in QCD spin physics. We briefly outline these experiments and their prime physics objectives.

Since May 2012 CEBAF is undergoing a major upgrade that will bring the maximum available energy of the electron beam to 12 GeV. The experimental equipment in all three halls will be upgraded (Halls A and C) or completely renewed (Hall B), in order to better match the increased energy and luminosity. A new experimental Hall D is being built. Commissioning of the new accelerator and of the experimental halls is expected for 2014. The future physics program focuses on dedicated studies of large $x$ phenomena, hard exclusive reactions, and TMD effects in kinematics where valence quarks dominate the physics (Dudek et al., 2012).

At CERN a proposal by the COMPASS Collaboration (Gautheron et al., 2010) to study TMDs and GPDs in the period 2014–2017 has been approved. The COMPASS data will provide a link between the kinematic domains of HERA, on the one hand, and of HERMES and JLAB, on the other hand. The program will start with the first ever polarized Drell-Yan experiment using a transversely polarized ammonia (proton) target and a negative pion beam. Because of the underlying annihilation of the anti-up-quark from the pion and the target up-quark, the process is dominated by the up-quark distribution in the valence region. An important goal is to check the QCD prediction of a sign change in the naive $T$-odd TMDs with respect to the DIS case. A study of GPDs in DVCS and hard exclusive meson production with a polarized muon beam will follow in 2015 using a liquid hydrogen target, a dedicated target recoil detector, and an additional large-angle electromagnetic calorimeter. An important measurement is the beam charge-and-spin asymmetry, which uses the property of the muon beam in which polarization changes sign when going from positive to negative muons. The first result of the correlation of transverse size and longitudinal momentum fraction might already be expected from a 2012 pilot run. In parallel semi-inclusive DIS data will be taken on the pure hydrogen target.

There are proposals to create a polarized fixed-target Drell-Yan program at Fermilab following the SeaQuest experiment, scheduled to complete data taking in 2014. Research and
development has begun for a suitable polarized target, and a
formal proposal to polarize the 120 GeV proton beam in the
Main Injector has been submitted to Fermilab management
(Courant et al., 2011). One of the primary physics motivations
for such a program would be to explore in detail the
QCD spin-momentum correlations described by TMD distrib-
utions such as the Sivers distribution, in particular, the role
of color flow in Drell-Yan versus semi-inclusive DIS
interactions.

A variety of possibilities for the medium-term future of
RHIC is currently under discussion. Research and develop-
ment is ongoing for a polarized $^3\text{He}$ source for RHIC
(Zelenksi et al., 2008), which would allow the neutron spin
structure to be studied in collider kinematics for the first time.
There are also proposals to significantly extend the detector
capabilities at RHIC; see, e.g., Aidala et al. (2012). Of
particular interest to nucleon structure studies are potential
upgraded forward spectrometers capable of reconstructing
jets, with hadronic particle identification and Drell-Yan mes-
urement capabilities up to pseudorapidities of $\sim 4$. The
ability to perform full jet reconstruction in the forward region
where large transverse single-spin asymmetries are observed,
and in addition to measure and identify hadrons within the jet,
would allow separation of effects due to distribution versus
fragmentation functions and shed light on the origin of these
significant spin-momentum correlations. An integrated de-
sign process for new detectors is underway such that they
would be able to take full advantage of electron-proton and
electron-ion collisions in the longer-term future should an
electron beam be added to RHIC.

Ideas for future polarization measurements are also dis-
cussed and investigated in more detail at FAIR (Germany),
J-PARC (Japan), and NICA (Russia).

A possible electron-ion collider (EIC) is being discussed in
connection with the future of RHIC and JLab. The goal is to
achieve highly polarized (greater than 70%) electron and
light-nucleus beams at center-of-mass energies ranging from
about 20 to 150 GeV at maximum collision luminosities
typically $\sim 10^{34}$ cm$^{-2}$ s$^{-1}$. Significant research and develop-
ment is ongoing to realize the technical challenges for reach-
ing this luminosity frontier for colliders and for achieving and
maintaining polarization of light nuclei (D and $^3\text{He}$) in a
storage ring. An EIC with the above performance offers
unique access to the small-$x$ region where gluons dominate as
well as to the intermediate and high-$x$ regions at unprece-
dented high $Q^2$. One could then study gluon polarization
down to $x$ values of about $10^{-4}$ (Aschenauer, Sassot,
and Stratmann, 2012). The high luminosity enables us to measure
and map GPDs over a broad range of the kinematic variables
and study the QCD evolution of the DVCS process plus TMD
distributions in kinematics where sea and glue effects are
expected to be important. In addition to being the first $ep$
collider exploring the structure of polarized protons, an EIC
would also be the first electron-nucleus collider allowing
precision studies of the gluon and sea-quark structure of
nuclei. Unpolarized ion beams from deuterium to the heaviest
nuclei (uranium or lead) would also be accelerated.
Knowledge about the spatial distribution of quarks and gluons
in nuclei is needed, for example, in the interpretation of heavy
ion collision data and the search for quark-gluon plasma. A
comprehensive review of EIC physics opportunities is given
by Boer et al. (2011).

IX. CONCLUSIONS AND OUTLOOK

The challenge to understand the internal spin structure of
the nucleon has inspired a global program of enormous
experimental and theoretical work in QCD during the last 25
years.

For longitudinal spin structure, there is a good convergence
of spin measurements from CERN, DESY, JLab, RHIC, and
SLAC taking into account the $Q^2$ dependence of the data and
kinematics of the different experiments. There is also good
convergence of theoretical understanding with the data, in-
cluding QCD-inspired models of the nucleon and lattice
calculations with disconnected diagrams included. Semi-
inclusive measurements in polarized lepton-nucleon and
proton-proton collisions have yielded much information
about the size of the separate valence, sea, and gluon spin
contributions to the nucleon’s spin. The small value of the
nucleon’s flavor-singlet axial charge, about 0.35, extracted
from polarized deep inelastic scattering seems to be a valence
quark effect. No significant sea-quark polarization is ob-
served in semi-inclusive deep inelastic scattering experi-
ments; the sum of valence spin contributions is in close
agreement with the measured total spin contribution
$g_A^{[0]}_{\text{pDIS}}$. While gluon polarization $\Delta g$ at the scale of the
experiments may be as much as 50% of the nucleon’s spin
at the scale of the experiments, the QCD anomaly correction
$-\frac{3}{2}(\alpha_s/2\pi)\Delta g$ is too small to resolve the difference between
$g_A^{[0]}_{\text{pDIS}}$ and the early quark model predictions, about 0.6.
Prime theory candidates to explain the small quark spin
content include transfer of valence quark spin to quark orbital
angular momentum through the pion cloud and a possible
topological effect whereby some fraction of the valence
quarks’ “spin” resides at Bjorken $x = 0$, where it is missed
by polarized deep inelastic scattering experiments. The
proton spin puzzle seems to be telling us about the interplay of
valence quarks with chiral dynamics and the complex vac-
uum structure of QCD. Ongoing and planned experimental
activity will improve the precision on the size of gluon and
strangeness polarization in the nucleon.

Finite orbital angular momentum of the valence quarks is
expected, also induced by confinement which introduces a
transverse scale in the physics. Quark orbital angular mo-
mentum through spin-orbit coupling is a prime candidate to
explain the large transverse single-spin asymmetries observed
in proton-proton collisions and lepton-nucleon scattering.
The desire to understand and measure QCD orbital angular-
momentum effects in the nucleon has spawned a new pro-
gram to explore and map the three-dimensional structure of
the nucleon, in both spatial coordinates (generalized parton
distributions) and transverse-momentum dependence.

Studies of transverse nucleon structure will drive the ex-
perimental program in the near future, with dedicated running
or approved programs at COMPASS, the 12 GeV upgrade
of JLab, FNAL, and RHIC. These experiments will test our
understanding of initial- and final-state interactions in QCD
(through comparison of Sivers and Boer-Mulders observables
in Drell-Yan and semi-inclusive deep inelastic scattering).

Precise measurements of GPDs and TMDs will test QCD evolution in a regime where transverse structure becomes important. The aim for precise information about quark (and gluon) total and orbital angular momentum in the nucleon is also a driving force for much theoretical work. Highlights include models of transverse spin phenomena, lattice calculations, and development of QCD fitting technology to extract GPDs and TMDs from the newly measurable observables.

ACKNOWLEDGMENTS

The research of S.D.B. is supported by the Austrian Science Fund, FWF, through Grants No. P20436 and No. P23753. We thank M. Diehl, R. Fatemi, A. Korzenev, S. Kuhn, W. Melnitchouk, B. Pasquini, T. C. Rogers, M. Stratmann, S. Tanega, A. W. Thomas, F. Videbaek, R. Windmolders, A. Zelenski, and E. Zemlyanichkina for helpful discussions.

REFERENCES

Aidala et al.: The spin structure of the nucleon

688


