Mapping the internal structure of hadrons through color and spin effects

van Daal, Tom Adrianus Antonius

Mapping the internal structure of hadrons through color and spin effects

Tom van Daal

rijksuniversiteit
groningen

Mapping the internal structure of hadrons through color and spin effects

Proefschrift

ter verkrijging van de graad van doctor aan de
Rijksuniversiteit Groningen
op gezag van de
rector magnificus prof. dr. E. Sterken
en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op

donderdag 13 december 2018 om 12.45 uur

door

Tom Adrianus Antonius van Daal

geboren op 25 oktober 1989
te Nijmegen

Promotores:
Prof. dr. D. Boer
Prof. dr. P.J.G. Mulders

Beoordelingscommissie:
Prof. dr. A. Schäfer
Prof. dr. I. Scimemi
Prof. dr. R.G.E. Timmermans

The material presented in this thesis is partly based on the following publications:

- D. Boer, S. Cotogno, T. van Daal, P. J. Mulders, A. Signori and Y.-J. Zhou, *Gluon and Wilson loop TMDs for hadrons of spin ≤ 1*, *JHEP* **10** (2016) 013, [1607.01654]

- D. Boer, T. van Daal, J. R. Gaunt, T. Kasemets and P. J. Mulders, *Colour unwound – disentangling colours for azimuthal asymmetries in Drell-Yan scattering*, *SciPost Phys.* **3** (2017) 040, [1709.04935]

- S. Cotogno, T. van Daal and P. J. Mulders, *Positivity bounds on gluon TMDs for hadrons of spin ≤ 1*, *JHEP* **11** (2017) 185, [1709.07827]

- D. Boer, T. van Daal, P. J. Mulders and E. Petreska, *Directed flow from C-odd gluon correlations at small x*, *JHEP* **07** (2018) 140, [1805.05219]

A full up-to-date list of my publications, including contributions to conference proceedings, can be found at the high-energy physics literature database called INSPIRE-HEP.

CONTENTS

INTRODUCTION

One of the main goals in physics is to unravel the fundamental structure of the universe in terms of its most elementary constituents. In the 19th century, the elementary building blocks of nature were thought to be those appearing in Mendeleev's periodic table of elements. Today, however, we know that the chemical elements are not elementary but composite; they consist of protons and neutrons that make up the atomic nuclei, surrounded by electrons.

Protons and neutrons are examples of hadrons. In the 1960s, a plethora of different kinds of hadrons were identified in particle detectors, all of which were believed to be elementary particles, albeit with a finite size. This multitude of newly discovered particles with similar properties, dubbed the 'particle zoo', led Gell-Mann [5] and Zweig [6, 7] to propose the quark model in 1964. They postulated that hadrons are composite objects, composed of more elementary particles called quarks. To explain the physical properties of the hadrons known at that time, their model contained three flavors of quarks, called up, down, and strange. The existence of quarks was experimentally verified in 1968 at SLAC [8, 9]; through deep-inelastic scattering of electrons off protons and neutrons the up and down quark flavors were discovered. For various reasons, theorists predicted the existence of more flavors. A total of six quark flavors was proposed in 1973 by Kobayashi and Maskawa [10], which was not experimentally verified until the discovery of the much heavier top quark in 1995 at Fermilab [11]. Soon after the conception of the quark model, it was proposed that quarks are bound together inside hadrons through the strong force mediated by so-called gluons. The existence of gluons was confirmed in 1979 at DESY [12–15]. So far, there are no experimental indications that quarks and gluons have any substructure – they are considered to be elementary particles.

The dynamics of quarks, antiquarks, and gluons (collectively known as partons) is described by the theory of the strong interaction called quantum chromodynamics (QCD). QCD is a non-Abelian gauge theory (also called Yang-Mills theory) whose fields carry a quantum number called color. Unlike the situation in quantum electrodynamics where the gauge field is electrically neutral, the gluons carry color charge. This implies that they not only interact with quarks and antiquarks, but also with each other. A consequence of the non-Abelian nature of QCD is that the strong force between partons increases as they are being separated. When the distance between two partons grows, it becomes at some point energetically favorable to create a new $q\bar{q}$ pair. As a result, we do not observe individual partons in experiments – rather, we detect jets of colorless combinations of partons, the hadrons. This phenomenon is called (color) confinement. There is no mathematical proof yet of confinement.

On the other hand, at smaller length scales, QCD becomes weakly coupled and the partons essentially behave as free particles. This is called asymptotic freedom and was discovered in 1973 by Gross, Wilczek, and Politzer [16, 17] – an important finding for which they were awarded the 2004 Nobel Prize in Physics. While the weakly coupled regime can be described using perturbation theory, we rely on fits to data for the nonperturbative part where the theory is strongly coupled. In this thesis we will focus on the latter regime.

Especially at hadron colliders, QCD plays a vital role in describing the high-energy collisions. When there is at least one hadron involved in either the initial or final state, the calculation of a cross section entails the perturbative calculation of the hard scattering event ('hard' means that $Q \gg M$, with Q the hard scale of the process and M the hadron mass), combined with unknown functions that parametrize the hadron structure that is governed by soft processes ('soft' means that $Q \sim M$ or smaller). Those functions are called parton distribution functions (PDFs) for partons inside initial-state hadrons and fragmentation functions (FFs) for partons decaying into final-state hadrons. Both types of functions can be constrained by fits to data. The simplest and most well-known type of PDF is the so-called collinear PDF, introduced by Feynman in 1969 [18]. These PDFs depend only on the motion of partons *collinear* to their parent hadron (the other momentum components are integrated over). A collinear PDF $f_i(x)$ represents the probability density to find a parton of type i with longitudinal momentum fraction x inside a given hadron.

The experimental extraction of PDFs relies on the possibility to *factorize* cross sections into separate pieces that describe short- and long-distance physics. When factorization applies, a cross section can be expressed in terms of a hard scattering factor and one or more collinear PDFs. The latter are *universal*, i.e. independent of the hard scattering event, which implies that they can be extracted from one process and used in others – a key ingredient of QCD phenomenology. Besides being a useful tool to describe collisions at hadron colliders, PDFs can teach us more about the inner workings of hadrons. However, the one-dimensional collinear PDFs are not sensitive to various interesting physics phenomena that are at play inside hadrons. To that end, we must go beyond a collinear description of partons inside hadrons.

To obtain more insight into the internal dynamics of hadrons, one can study *multidimensional* PDFs instead. The phenomenology of multidimensional quark and gluon PDFs forms the main subject of this thesis. The type of PDFs that we will discuss most are the so-called transverse momentum dependent (TMD) PDFs, often simply referred to as TMDs. Besides x, these functions also depend on the partonic transverse momentum $\boldsymbol{k}$ (the transverse directions are perpendicular to the collision axis), making them three-dimensional. For example, TMDs can encode interesting correlations between $\boldsymbol{k}$ and the spin polarization of the parton or hadron.

Typically, TMD effects manifest themselves in cross sections through angular correlations between the produced particles. To understand how those correlations arise, let us consider the production of two jets in a hadron-hadron collision. If the

partons inside the hadrons would have no transverse momentum,[1] i.e. they move perfectly collinear to their parent hadrons, then the produced jets would appear back-to-back in the transverse plane due to momentum conservation. However, if at least one parton participating in the hard scattering process has a nonzero transverse component, then the jets are no longer produced back-to-back in the transverse plane. The latter is referred to as an azimuthal asymmetry and can be measured in experiment. Observables can be sensitive to TMDs (and thus the three-dimensional inner structure of hadrons) if they are differential in sufficiently many variables, particularly (directions of) transverse momenta of final-state particles.

As it turns out, unlike the collinear PDFs, TMDs are not universal – they are *process dependent*. The process dependence of TMDs originates from the color structure of the hard scattering event. Fortunately, as the color structure of the hard scattering event can be calculated, the process dependence of TMDs is calculable event-by-event, implying that not all universality (and hence predictability) is in fact lost. Compared to collinear PDFs, our current knowledge of TMDs is rather limited; this is because the extraction of TMDs requires measurements of more complex processes.

In general, greater knowledge on multidimensional PDFs teaches us more about the inner structure of hadrons. Through multidimensional PDFs we could deepen our understanding of the angular momentum structure of hadrons. For example, it is unclear how the proton spin can be decomposed into quark and gluon contributions. Furthermore, multidimensional PDFs could shed light on the origins of hadron mass. After all, the Higgs mechanism only explains quark masses that add up to at most tens of MeVs, which is a few percent of the proton mass. The actual proton mass is much larger due to QCD dynamics that leads to the hadronic scale of several hundreds of MeVs, which is also the order of magnitude of intrinsic partonic transverse momenta. On the other hand, many theoretical and experimental efforts on PDFs over the last decades have set the stage to test some important fundamental predictions of QCD, for example ones related to the process dependence of TMDs. Furthermore, an improved understanding of PDFs can lead to more accurate descriptions of certain processes that are sensitive to physics beyond the Standard Model. Thus, the study of PDFs not only enriches our knowledge on the dynamics of quarks and gluons inside hadrons, it can also serve to check the consistency of the QCD framework and possibly identify deviations from the Standard Model.

In this thesis we will discuss multidimensional PDFs from a theoretical perspective. To this end, we will first provide an introduction to PDFs in chapter 2. After discussing the basics of QCD, we will define various types of PDFs, both collinear and multidimensional ones. Since we will mainly focus on TMDs, we will set up the necessary formalism in detail for the TMD case.

---

[1] For simplicity we only consider here 'intrinsic' transverse momentum and not perturbatively generated transverse momentum.

The next two chapters comprise part I of this thesis and discuss factorization of Drell-Yan (DY) scattering in terms of a perturbatively calculable hard factor and two TMDs, called TMD factorization. The DY process describes lepton-antilepton production through $q\bar{q}$ annihilation in hadron-hadron collisions. For certain TMDs that give rise to particular azimuthal asymmetries in the DY cross section, it is unclear whether the (standard) factorization theorem holds due to seemingly nontrivial color structures. In chapter 3 we review how TMD factorization has been established for DY, and in chapter 4 we perform a factorization study largely along the lines of the original proof for the potentially problematic TMDs, investigating whether factorization could be violated.

Part II of this thesis is about multidimensional gluon PDFs. In chapter 5 we introduce TMDs that can be used to describe gluons inside polarized hadrons up to spin 1. We also derive bounds for those TMDs. Subsequently, in chapter 6, we study gluon TMDs and other multidimensional gluon PDFs at small x. In the small-x kinematic region the hadron content is dominated by gluons. In that regime, we show that the number of independent gluon PDFs significantly drops, thereby establishing a simplified picture of the gluonic structure of hadrons. We will also demonstrate that correlations between gluons inside nuclei can give rise to (odd-harmonic) angular correlations between two final-state hadrons in pA collisions.

In chapter 7 we summarize the main conclusions of this thesis. In appendices A–C we list the most important conventions, and in the remaining appendices we provide additional technical details. Finally, we provide a lay summary of this thesis (including a Dutch translation).

MAPPING THE INTERNAL STRUCTURE OF HADRONS

Since all research presented in this thesis is related to the strong interaction, we first review the most important aspects of QCD. Subsequently, in section 2.2 we discuss how the internal structure of hadrons can be described in terms of quark PDFs. After reviewing the collinear case, we discuss the more involved multidimensional PDFs that will play a central role in later chapters. Special attention is given to the three-dimensional TMDs in section 2.3. We will set up the TMD formalism for quarks inside polarized hadrons, needed for part I of this thesis. Fortunately, most of the discussion on quark distributions in this chapter can be easily extended to the multidimensional gluon distributions that are discussed in part II.

2.1 QUANTUM CHROMODYNAMICS IN A NUTSHELL

QCD is a sector of the Standard Model gauge theory. The underlying symmetry group is the non-Abelian Lie group $SU(N_c)$, with $N_c = 3$ (the number of colors). Invariance under local $SU(3)$ transformations gives rise to eight massless gauge boson fields called gluons. The gluons belong to the adjoint representation of the gauge group, while the fermions that couple to the gluons, the quarks, belong to the fundamental representation.

2.1.1 *The Lagrangian*

A quantum field theory is usually described by a Lagrangian density. The Lagrangian density of QCD is given by [19]

$$\mathcal{L}_{\text{QCD}} = -\frac{1}{4} F^a_{\mu\nu} F^{a\mu\nu} + \sum_q \overline{\psi}_q \left(i\slashed{D} - m_q I \right) \psi_q \,, \tag{2.1}$$

where the field ψ_q with mass m_q is a color triplet quark field pertaining to one of the six quark flavors labeled by q. In this thesis, I denotes the identity matrix in *any* space. The covariant derivative is given by

$$D_\mu \equiv I \, \partial_\mu + i g t^a A^a_\mu \,, \tag{2.2}$$

where A denotes the color octet gluon field and g is the gauge coupling. The gluon field strength tensor is defined as

$$F^a_{\mu\nu} \equiv \partial_\mu A^a_\nu - \partial_\nu A^a_\mu - g f^{abc} A^b_\mu A^c_\nu \,, \tag{2.3}$$

where the last term gives rise to three- and four-gluon vertices in the first term of eq. (2.1).

The color matrices t^a are the generators of the fundamental representation; they are traceless and Hermitian, and satisfy the following (anti)commutation relations:

$$[t^a, t^b] = if^{abc}t^c, \qquad \{t^a, t^b\} = \frac{1}{N_c}\delta^{ab} + d^{abc}t^c.$$ (2.4)

Furthermore, they are normalized as

$$\mathrm{Tr}(t^a t^b) = \tfrac{1}{2}\delta^{ab}.$$ (2.5)

The structure constants of SU(3) are given by

$$f^{abc} = -2i\,\mathrm{Tr}\left([t^a, t^b]t^c\right), \qquad d^{abc} = 2\,\mathrm{Tr}\left(\{t^a, t^b\}t^c\right).$$ (2.6)

Both f and d are real and they are, respectively, totally antisymmetric and symmetric under the interchange of any of their indices. The numerical values of these constants, as well as explicit representations for the color matrices, can be found in e.g. [20].

2.1.2 *The running coupling*

Despite its name, the QCD coupling constant $\alpha_s \equiv g^2/(4\pi)$ is *not* a constant. The value that should be used in calculations depends on the energy scale involved. Since α_s satisfies a first-order differential equation (a renormalization group equation), its value at any scale Q can be related to one particular reference scale. For this reference scale we could take the Landau pole, which is the energy scale at which the coupling constant becomes infinite, also referred to as the QCD scale (Λ_{QCD}). At leading order (i.e. at one-loop order), the running coupling is given by[1]

$$\alpha_s(Q) = \frac{2\pi}{b_0 \ln(Q/\Lambda_{\mathrm{QCD}})}, \qquad b_0 \equiv 11 - \frac{2}{3}n_f,$$ (2.7)

with n_f the number of active flavors at the scale Q. The value of the QCD scale is not uniquely determined; besides a dependence on n_f and the perturbative accuracy, this scale also depends on the renormalization scheme (this dependence starts at three loops) – typically $\Lambda_{\mathrm{QCD}} \sim 0.2$ GeV [23]. Note that for $n_f \leq 16$ we have $b_0 > 0$, which implies that the coupling decreases with increasing energy. As a consequence, the strong interaction becomes asymptotically weaker at decreasing length scales, a property of QCD called asymptotic freedom. This phenomenon is a feature of Yang-Mills theories in general [24, 25].

Whenever the energy scale is somewhat larger than Λ_{QCD}, perturbative calculations are allowed; typically one takes for this scale $Q = 1$ GeV (which is approximately the mass of the proton), where $\alpha_s(Q) \approx 0.4$. At lower energies, however, the theory becomes strongly coupled, leading to the *confinement* of quarks and gluons

---

1 Today, the running of α_s is known up to five loops [21, 22].

inside hadrons – this is the nonperturbative regime. When quarks are produced in particle accelerators, they cannot be isolated in the detectors; we instead observe jets that contain many color-neutral mesons and baryons. Hence, the low- and high-energy particle degrees of freedom of QCD are constituted by hadrons and partons, respectively.

2.1.3 *Deep-inelastic scattering*

QCD is essential in providing descriptions of collisions that feature the strong interaction. For a given process that involves hadrons, the hard scattering event can be calculated using perturbation theory, while the hadronic structure can be parametrized in terms of PDFs or FFs that need to be fitted to data (in this thesis we will only consider PDFs). The simplest and most well-known type of PDF is the collinear one. As already mentioned in the introduction, a collinear PDF $f_i(x)$ represents the probability density for finding a parton of type i with longitudinal momentum fraction x inside a given hadron.

To illustrate how a collinear PDF enters a cross section calculation, we need to consider a process that involves at least one hadron in the initial state, e.g. deep-inelastic scattering (DIS) given by $e^- H \to e^- X$. DIS is the simplest and most-studied process in QCD. By H we denote a hadron and X represents all unobserved final-state particles; *only* the electron is measured after the scattering. For that reason, this process is also referred to as 'inclusive' DIS. In contrast, the process $e^- H \to e^- H' X$, where also one hadron H' is measured in the final state, is called semi-inclusive DIS (SIDIS).

The leading-order (also referred to as tree-level) diagram for DIS is given in figure 2.1. The hadron and quark momenta are denoted by P and k respectively, and the photon momentum is denoted by q. The hard scale of the process is set by the photon virtuality $Q^2 \equiv -q^2 > 0$. The graph in figure 2.1 is split by a dotted line, called the final-state cut, into an amplitude (all that is to the left of the cut) and a conjugate amplitude (all that is to the right of the cut). The green 'blob', denoted by Φ, represents the quark-quark correlation function, also referred to as the correlator (the operator definition of this quantity will be given in the next section). Since Φ encodes the internal dynamics of the hadron, it cannot be calculated using perturbation theory; rather, it is parametrized in terms of a collinear PDF called f_1 (see eq. (2.19) in the next section).

The cross section for DIS has the following form:

$$\sigma = \sum_i \int_0^1 dx \, f_{1,i}(x) \, \hat{\sigma}_i(x), \tag{2.8}$$

where the sum only includes contributions from quarks and antiquarks. The hard-scattering coefficient $\hat{\sigma}_i$ describes the partonic process $e^- q_i \to e^- q_i$ and can be calculated using perturbation theory. Since $f_{1,i}$ provides the probability for finding
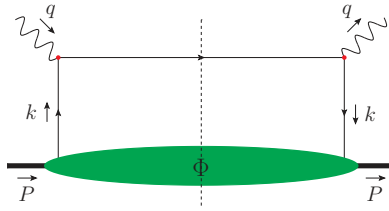
Figure 2.1: The tree-level (amplitude squared) graph for DIS: an electron emits a photon that subsequently scatters off a quark inside a hadron (the electrons are not shown). The green 'blob' represents the quark-quark correlator, and the dotted line in the middle represents the final-state cut.

a parton i inside the hadron, it must be normalized in a way such that the valence quark content of the hadron is respected (this is called a number sum rule). The range of allowed values for x ($0 \leq x \leq 1$) will be commented on in the next section.

The result in eq. (2.8) is an approximation – it is given at leading order in α_s and at leading power in M/Q. This is based on the fact that the accuracy of observables in QCD is determined by *two different* expansions. On the one hand, there is an expansion in the coupling constant α_s, and, on the other hand, there is an expansion in the ratio M/Q. The latter is referred to as the twist expansion and higher order terms in this expansion are called power corrections.[2] Whereas the coupling expansion quantifies the accuracy of the perturbative structure of the theory, the twist expansion quantifies the accuracy of its nonperturbative structure. In contrast to the expansion in α_s, the full twist expansion is always limited to a certain power. In this thesis we will only consider quantities at leading twist, which is justified as long as $Q \gg M$. Power corrections to eq. (2.8) modify the hard-scattering coefficient and include PDFs of subleading twist.

Let us also briefly comment on higher-order α_s corrections to eq. (2.8). Perturbative QCD corrections modify both the hard-scattering coefficient and the PDF. Already at the one-loop level, several types of divergences appear. The ultraviolet divergences can be taken care of using renormalization, which implies the introduction of a renormalization scale μ. Since a physical observable cannot depend on μ, this leads to renormalization group equations. For the collinear PDFs they are called the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations [27–29]; these coupled differential equations can be used to evolve PDFs from one scale to another (as evolution will not play a role in this thesis, the scale dependence of PDFs will always be implicit). Other divergences that appear are infrared divergences, reflecting the sensitivity of the cross section to long-distance effects. They can arise when a gluon is emitted collinear to a quark or when a gluon has a vanishing momentum. For DIS, the latter type of divergence cancels upon summing over all real and virtual

---

2 The precise definition of 'twist' is not relevant for our purposes. In fact, there exist various definitions of twist in the literature, see e.g. the review in [26].

diagrams [30]. The former type, however, also referred to as the collinear divergence, does not cancel and can be absorbed by the nonperturbative PDF.

The DIS cross section in eq. (2.8) shows a very important feature, namely the *factorization* of the short-distance physics (captured by $\hat\sigma$) and the long-distance physics (captured by f_1). Factorization is a key ingredient of QCD phenomenology. As already mentioned in the introduction, a consequence of factorization is that the collinear PDFs are *universal* – the internal dynamics of hadrons is the same for all processes. As it turns out, proving factorization is highly nontrivial due to for instance gluon exchanges between the different parts of a process. Only for a limited number of processes proofs are available up to all orders in α_s. Factorization will be discussed in detail in chapter 3.

2.2 THE QUARK-QUARK CORRELATION FUNCTION

In this section we provide the precise field-theoretical definition of the quark-quark correlator Φ which we encountered already in the DIS graph given in figure 2.1. In subsection 2.2.2 we describe the collinear correlator that can be parametrized in terms of the collinear PDF f_1, and in subsection 2.2.3 we discuss more complex correlators that can be parametrized in terms of multidimensional PDFs. First, however, let us discuss some important conventions.

Throughout this thesis we make extensive use of light-cone coordinates, which are convenient for describing high-energy scattering processes. In terms of those coordinates, a four-vector a can be represented as $a = (a^+, a^-, \boldsymbol{a})$, where $a^{\pm} \equiv (a^0 \pm a^3)/\sqrt{2}$ and $\boldsymbol{a} \equiv (a^1, a^2)$. Adopting the 'mostly minus' convention for the Minkowski metric g, a dot product between two four-vectors a and b is given by $a \cdot b \equiv g^{\mu\nu} a_\mu b_\nu = a^+ b^- + a^- b^+ - \boldsymbol{a} \cdot \boldsymbol{b}$. Furthermore, we employ the lightlike vectors $\bar{n} \equiv (1, 0, \boldsymbol{0})$ and $n \equiv (0, 1, \boldsymbol{0})$, such that a can be decomposed as $a^\mu = a^+ \bar{n}^\mu + a^- n^\mu + a_T^\mu$. We have defined the quantity $a_T \equiv (0, 0, \boldsymbol{a})$, so that $a_T^2 = -\boldsymbol{a}^2$. An overview of relevant conventions is provided in appendix A. In terms of $\bar{n}$ and n, the hadron and quark momenta P and k are given by:

$$P^\mu = P^+ \bar{n}^\mu + \frac{M^2}{2P^+} n^\mu, \tag{2.9}$$

$$k^\mu = x P^+ \bar{n}^\mu + k^- n^\mu + k_T^\mu. \tag{2.10}$$

Note that we have chosen the hadron to move along the z-axis, satisfying the on-shell condition $P^2 = M^2$. Furthermore, we have defined the longitudinal momentum fraction as $x \equiv k^+ / P^+$. Typically, $P^+ \sim Q$ in cross section calculations, where Q is the relevant hard scale in the process. In the high-energy limit, the minus component of P becomes negligible compared to its plus component. Hence, in this limit, P becomes lightlike and the hadron momentum is directed along the plus direction.

2.2.1 *The unintegrated correlator*

The field-theoretical definition of the quark-quark correlation matrix Φ is given by

$$\Phi_{\alpha\beta}(k) \equiv \sum_X \int \frac{d^3\vec{P}_X}{(2\pi)^3 \, 2P_X^0} \, \langle P| \, \overline{\psi}_\beta(0) \, |X\rangle\langle X| \, \psi_\alpha(0) \, |P\rangle \, \delta^{(4)}(P - k - P_X), \qquad (2.11)$$

where α, β are Dirac indices. A summation over color triplet indices is implicitly understood. The dependence of Φ on P is implicit (as this is not the dependence we are ultimately interested in), and, to lighten the notation, we suppress flavor labels for the most part in this thesis. Any dependence on hadron polarization will not be discussed until section 2.3, so here we consider unpolarized hadrons only. The product of matrix elements in eq. (2.11) gives the probability density for the extraction of a quark with momentum k from the hadron, yielding a remnant X with momentum P_X. As X is a final state we integrate over its phase space, and because it is not measured experimentally we also perform a sum over all possibilities. The delta function reflects momentum conservation; its Fourier transform to the position variable z can be used to perform a translation from the point 0 to z of the field operator ψ. Then, upon using the completeness relation for the unobserved final states, we obtain

$$\Phi_{\alpha\beta}(k) = \int \frac{d^4z}{(2\pi)^4} \, e^{ik\cdot z} \, \langle P| \, \overline{\psi}_\beta(0)\psi_\alpha(z) \, |P\rangle \, . \qquad (2.12)$$

Since this correlator depends on the full quark four-momentum k, it is also referred to as the *unintegrated* correlator. The pictorial definition of Φ is given by the graph on the left in figure 2.2.
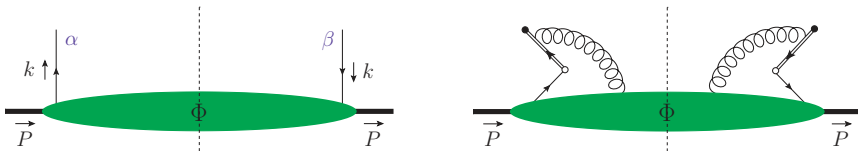
Figure 2.2: The definition of the quark-quark correlator Φ. The graph on the left is not gauge invariant; the necessary Wilson line is only included in the graph on the right (infinitely many gluon emissions are implied). See appendix C for the notation regarding eikonal lines and vertices.

The expression for Φ in eq. (2.12) is not completely correct – one important ingredient is missing. Since the two quark fields reside at two different spacetime positions (the matrix element is said to be bilocal), the correlator is *not* gauge invariant. After all, under local SU(3) transformations V, the quark fields transform as

$$\psi_\alpha(z) \;\rightarrow\; V(z)\,\psi_\alpha(z), \qquad \overline{\psi}_\beta(0) \;\rightarrow\; \overline{\psi}_\beta(0)\,V^\dagger(0), \qquad (2.13)$$

so that the operator combination $\overline{\psi}_\beta(0)\psi_\alpha(z)$ changes. The gauge dependence of the correlator is concerning. Since the PDFs appearing in the parametrization of Φ can be extracted from experiments, they should be independent of arbitrary gauge transformations. This problem can be fixed as follows: if we could insert in between the quark fields a (color) matrix-valued quantity $U_{[0,z]}$ that transforms as

$$U_{[0,z]} \rightarrow V(0)\, U_{[0,z]} V^\dagger(z), \tag{2.14}$$

then the correlator would be gauge invariant. A quantity with these properties exists – it is called a Wilson line (or gauge link), given by [31]

$$U_{[0,z]} = \mathcal{P} \exp\left(ig \int_0^z d\eta^\mu A_\mu(\eta) \right), \tag{2.15}$$

where $A_\mu \equiv A_\mu^a t^a$. The symbol $\mathcal{P}$ denotes path ordering; for two fields A and B at different positions, path ordering is defined as

$$\mathcal{P} A(x)B(y) \equiv \theta(x-y)\, A(x)B(y) + \theta(y-x)\, B(y)A(x), \tag{2.16}$$

where θ is the Heaviside step function. Eq. (2.16) can be easily generalized to include more fields.

For a correlator to be gauge invariant, one needs to include a Wilson line. The gauge-invariant correlator corresponds to the graph on the right in figure 2.2. Physically, a Wilson line reflects a sum over infinitely many gluon emissions which can be calculated explicitly [32]. The latter is discussed in detail for the TMD case in e.g. [33] (in section 4.4 we calculate the $\mathcal{O}(g)$ contribution to a Wilson line).

One might wonder whether there are any restrictions on the integration path of the Wilson line in eq. (2.15). In principle, any possible path running between 0 and z guarantees gauge invariance of Φ. However, since in QCD a closed path does not vanish, different paths correspond to physically different situations. As it turns out, for certain types of correlators different paths correspond to different processes in which the correlator appears. The latter has far-reaching consequences and forms one of the main themes of this thesis. We will discuss this matter in detail in section 2.3.

2.2.2 *The collinear correlator*

Since the unintegrated correlator in eq. (2.12) does not appear in any known process, we are led to study certain projections. The most well-known projection is the collinear correlator, obtained after integration over k^- and $\mathbf{k}$; it is given by [34]

$$\Phi_{\alpha\beta}(x) \equiv \int \frac{dz^-}{2\pi}\, e^{ik\cdot z}\, \langle P|\, \overline{\psi}_\beta(0)\, U^n_{[0,z^-;\mathbf{0}]}\, \psi_\alpha(z)\, |P\rangle \bigg|_{z^2=0}. \tag{2.17}$$

As this correlator is evaluated at $z^+ = 0$, $\boldsymbol{z} = \mathbf{0}$ (implying $z^2 = 0$), it is also sometimes called the light-cone correlator. The gauge link that bridges the nonlocality runs along the minus direction (indicated by the superscript n) and is given by

$$U^n_{[0,z^-;\mathbf{0}]} \equiv \mathcal{P} \exp \left(ig \int_0^{z^-} d\eta^- A^+ (\eta^+ = 0, \eta^-, \boldsymbol{\eta} = \mathbf{0}) \right). \qquad (2.18)$$

Note that the integration path is fixed, as illustrated in figure 2.3. This implies that, in the collinear case, the gauge link is uniquely determined. It might seem strange at this point to indicate the transverse position of the link ($\boldsymbol{\eta} = \mathbf{0}$) on the left-hand side of eq. (2.18), but this notation will be useful later when we will also consider links that are not unique, involving integration paths at $\boldsymbol{\eta} \neq \mathbf{0}$. In the $A^+ = 0$ light-cone gauge, the collinear Wilson line reduces to unity and can hence be omitted. By contrast, this is *not* possible in the more complex situations that will be considered later.
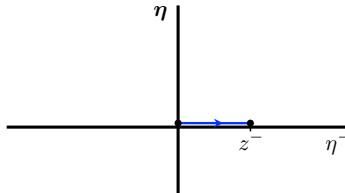
Figure 2.3: The integration path of the gauge link $U^n_{[0,z^-;\mathbf{0}]}$. The path runs along the light-front $\eta^+ = 0$ at transverse position $\boldsymbol{\eta} = \mathbf{0}$.

The correlator in eq. (2.17) is precisely the one that we encountered in the DIS example in subsection 2.1.3. As mentioned earlier, it can be parametrized in terms of one collinear PDF at leading twist (this is indicated by the subscript 1):

$$\Phi(x) = \frac{1}{2} \not{n} f_1(x). \qquad (2.19)$$

At subleading twist, i.e. at $\mathcal{O}(M/P^+)$ or beyond, there are more functions, each appearing with a different Dirac structure [35, 36].

Let us comment on the range of allowed values for x. The intermediate state in the correlator in figure 2.2 carries momentum $P - k$, and, since it is cut, its energy must be positive. Thus, it follows that $P^+ - k^+ \geq 0$, which translates into the condition $x \leq 1$. Similarly, for negative values of x it follows that $x \geq -1$. For negative x, though, a quark PDF can be related to an antiquark PDF (indicated by a bar) at positive x: $f_1(x) = -\bar{f}_1(-x)$. Conventionally, only positive values of x are considered, both for quarks and antiquarks, and so we always take $0 \leq x \leq 1$.

Besides number sum rules, PDFs also satisfy a momentum sum rule. This rule says that the total fractional momentum carried by partons is unity:

$$\sum_i \int_0^1 dx\, x f_{1,i}(x) = 1, \qquad (2.20)$$
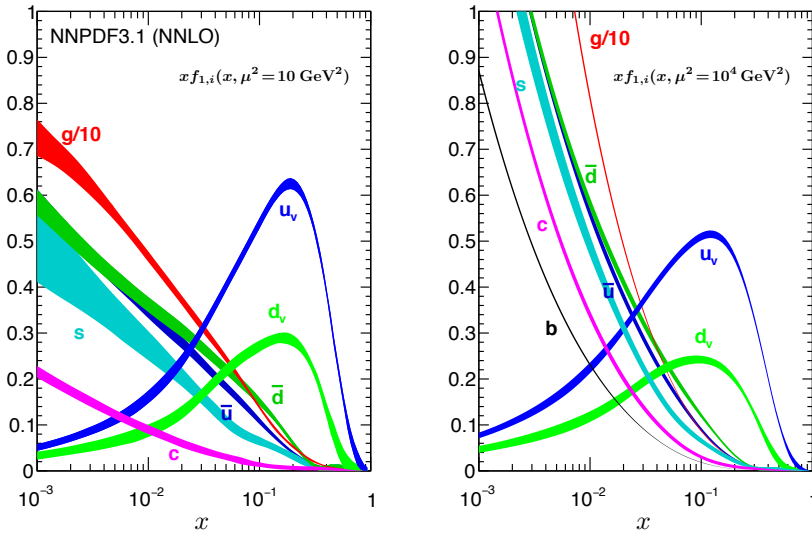
Figure 2.4: The collinear PDFs $x f_{1,i}(x, \mu^2)$ for the proton at two different energy scales μ; in the left plot $\mu^2 = 10$ GeV^2 and in the right plot $\mu^2 = 10^4$ GeV^2 [37]. The subscript i labels all valence quarks, sea quarks, and gluons. Note that the gluon contributions have been divided by a factor of 10.

the total momentum of the parent hadron in the plus direction. For the case of a proton, the momentum contributions from each parton are given in the plot in figure 2.4. As it turns out, when $Q \gg M$, quarks and antiquarks only account for about half of the total momentum – the other half is carried by gluons. Furthermore, note that at small x the proton content is dominated by gluons. For that reason, we will study gluon distributions at small x in chapter 6.

2.2.3 *Multidimensional correlators*

So far we have described the hadron structure only in terms of the one-dimensional collinear PDFs that appear in the parametrization of the collinear correlator defined in eq. (2.17). The collinear correlator can be obtained through certain projections or limits from more general correlators, see figure 2.5. Generally speaking, correlators that depend on more kinematical variables contain more information on the internal structure of hadrons. Similarly to the one-dimensional case in eq. (2.19), those correlators can be parametrized in terms of *multidimensional* PDFs. In this subsection we introduce the various types of PDFs that will be discussed in later chapters. Recent reviews on multidimensional distribution functions include [38–40].

Let us start our discussion on the various distribution functions from the so-called generalized TMDs (GTMDs). GTMDs depend on six momentum variables, and of all functions we will consider here, they encode most information about the distributions
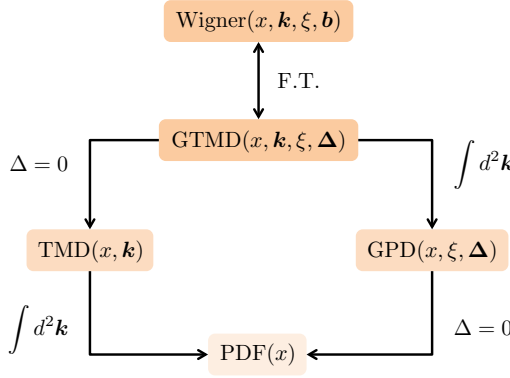
Figure 2.5: Relations between Wigner distributions, GTMDs, GPDs, TMDs, and (collinear) PDFs. In between brackets it is indicated which kinematic variables the functions depend on.

of partons inside hadrons. These functions appear in the parametrization of the GTMD correlator. The quark-quark GTMD correlator is defined as[3]

$$
G_{\alpha\beta}(x, \boldsymbol{k}, \xi, \boldsymbol{\Delta}) \equiv \int \frac{dz^- d^2\boldsymbol{z}}{(2\pi)^3} \, e^{ik \cdot z} \, \langle p' | \, \overline{\psi}_\beta \left(-\tfrac{z}{2}\right) U_{\left[-\frac{z}{2}, \frac{z}{2}\right]} \psi_\alpha \left(\tfrac{z}{2}\right) | p \rangle \bigg|_{z^+=0}. \qquad (2.21)
$$

The term 'generalized' in the name refers to the fact that this is the Fourier transform of an *off-forward* matrix element. The off-forwardness is quantified by the momentum transfer $\Delta \equiv p' - p$ (the momenta p and p' can be taken to have large plus components). The average hadron and parton momenta are respectively given by $P \equiv (p' + p)/2$ and k, see also eqs. (2.9) and (2.10). The kinematics for off-forward distributions is clarified in figure 2.6. Note that there is no final-state cut – an off-forward correlator represents an amplitude rather than a squared amplitude. The momentum transfer Δ can be decomposed as

$$
\Delta^\mu = -2\xi P^+ \bar{n}^\mu + \frac{\xi M^2}{P^+} n^\mu + \Delta_T^\mu, \qquad (2.22)
$$

where ξ denotes the skewness parameter, given by $2\xi \equiv -\Delta^+/P^+$ (its physical region corresponds to $-1 \leq \xi \leq 1$, see e.g. [42]).

For a spin-$\frac{1}{2}$ hadron, there are sixteen quark GTMDs at leading twist [43]. In contrast to all other distributions appearing in figure 2.5, GTMDs are complex-valued functions. Very recently it was shown for the first time that observables can be sensitive to quark GTMDs. It was found that they could be accessed (in principle) through exclusive double DY scattering [44].

---

3 It was recently shown that a proper definition of the GTMD correlator includes additional dependence on soft radiation described by a so-called soft factor [41]. However, as this will not play any role in this thesis, it is not discussed here (a similar dependence on soft radiation arises in the TMD case; this will be briefly discussed in section 3.5).

Figure 2.6: The definition of the off-forward quark-quark correlator G (the contributions to the necessary gauge link are not shown).

The so-called Wigner distributions can be obtained from GTMDs by a Fourier transformation from $\boldsymbol{\Delta}$ to $\boldsymbol{b}$:

$$W_{\alpha\beta}(x, \boldsymbol{k}, \xi, \boldsymbol{b}) \equiv \int d^2\boldsymbol{\Delta}\, e^{-i\boldsymbol{\Delta}\cdot\boldsymbol{b}}\, G_{\alpha\beta}(x, \boldsymbol{k}, \xi, \boldsymbol{\Delta}). \qquad (2.23)$$

The variable $\boldsymbol{b}$ is called the impact parameter and represents the average partonic transverse position. The six-dimensional Wigner functions for the parton-hadron system were introduced in [45, 46] and are analogous to classical phase-space distributions. Whereas the latter are probability densities, the former are *quasiprobability* densities – Wigner distributions are for example not positive definite, but *can* be used to compute expectation values of physical observables.

Quark GTMDs and Wigner distributions have not been measured yet in experiment and for that reason one currently has to rely on model studies, see e.g. [47–55]. In section 6.5 we will study an observable that is sensitive to gluon Wigner distributions.

Upon integrating the GTMD correlator over $\boldsymbol{k}$, we obtain the so-called generalized parton distribution (GPD) correlator that depends on x, ξ, and $\boldsymbol{\Delta}$. At leading twist, there are eight quark GPDs for a spin-$\frac{1}{2}$ hadron [56]. Interest in quark GPDs is for example related to the observation that they could be used to obtain more information about the total angular momentum carried by quarks [57]. GPDs can be accessed through various exclusive processes, albeit very challenging. We will not discuss GPDs further in this thesis; for more information, see e.g. [42].

When we set $\Delta = 0$ in the GTMD correlator (this is called the 'forward' limit), we obtain the TMD correlator. As mentioned already in the introduction, TMDs depend on x as well as on the partonic transverse momentum $\boldsymbol{k}$. Since the phenomenology of quark and gluon TMDs forms one of the main topics of this thesis, we provide an extensive introduction to the TMD formalism in the next section. Among other things, we will see that there are eight leading-twist quark TMDs for a spin-$\frac{1}{2}$ hadron [36].

Since GPDs and TMDs can be obtained from GTMDs (or, equivalently, from Wigner distributions) through certain projections or limits, the latter are sometimes referred to as the 'mother' distributions. It is important to note that GTMDs contain *more information* on hadron structure than can be acquired from the combined knowledge of GPDs and TMDs – several GTMDs vanish for the GPD and TMD limits of the correlator [43]. The reason for this is that GTMDs can contain information on correlations between $\boldsymbol{k}$ and $\boldsymbol{\Delta}$, which cannot be described by GPDs or TMDs alone.

Both the GPD and TMD correlators reduce to the collinear correlator in eq. (2.17) upon setting $\Delta = 0$ and integrating over $\boldsymbol{k}$, respectively. For a spin-$\frac{1}{2}$ hadron, there are three collinear PDFs at leading twist [58], all of which have been extracted from experiment.

2.3 THE TMD FORMALISM

In this section we set up the TMD formalism. We discuss in detail the definition of the quark-quark TMD correlator and its parametrization in terms of quark TMDs (the gluon case is very similar and is extensively discussed in chapter 5). Since hadron polarization forms a key component of TMD phenomenology, we start this section with a description of hadron spin. Although we will only consider spin-1 hadrons explicitly, spin-$\frac{1}{2}$ and spin-0 systems are trivially covered as well. In subsection 2.3.3 we explain the origin of the process dependence of TMDs. Most of the discussions in this section straightforwardly apply to GTMDs as well, which we encounter in chapter 6 for the case of gluons. Introductions to the TMD formalism are also found in e.g. [26, 33, 59].

2.3.1 *Inclusion of hadron spin*

So far we have ignored any dependence of the correlator on the spin polarization of the hadron. Whereas spin-0 hadrons are unpolarized (by definition), an ensemble of spin-$\frac{1}{2}$ hadrons can be unpolarized or vector polarized. To describe the degree of vector polarization, we make use of a spacelike spin vector S. For spin-1 hadrons, we need besides S also a symmetric traceless spin tensor T to describe tensor polarized states. The simplest and hence most-studied spin-1 system is the deuteron (this is a deuterium nucleus which consists of one proton and one neutron). Satisfying $P \cdot S = 0$ and $P_\mu T^{\mu\nu} = 0$, the spin vector and spin tensor can be parametrized as follows [60, 61]:

$$S^\mu = S_L \frac{P^+}{M} \bar{n}^\mu - S_L \frac{M}{2P^+} n^\mu + S_T^\mu, \tag{2.24}$$

$$T^{\mu\nu} = \frac{1}{2} \left[\frac{4}{3} S_{LL} \frac{(P^+)^2}{M^2} \bar{n}^\mu \bar{n}^\nu + \frac{P^+}{M} \bar{n}^{\{\mu} S_{LT}^{\nu\}} - \frac{2}{3} S_{LL} \left(\bar{n}^{\{\mu} n^{\nu\}} - g_T^{\mu\nu} \right) \right.$$
$$\left. + S_{TT}^{\mu\nu} - \frac{M}{2P^+} n^{\{\mu} S_{LT}^{\nu\}} + \frac{1}{3} S_{LL} \frac{M^2}{(P^+)^2} n^\mu n^\nu \right], \tag{2.25}$$

where curly brackets denote symmetrization of the indices, and $g_T^{\mu\nu} \equiv g^{\mu\nu} - \bar{n}^{\{\mu} n^{\nu\}}$ denotes the metric tensor in transverse space with nonzero components $g_T^{11} = g_T^{22} = -1$. The spin vector has three independent parameters, namely S_L and the two transverse components of S_T; they characterize longitudinal and transverse polarization, respectively. The spin tensor, on the other hand, contains five independent

parameters, namely S_{LL}, the two transverse components of S_{LT}, as well as the two independent components of the symmetric traceless transverse tensor S_{TT}. The physical interpretations of the various components of the spin tensor are discussed in e.g. [60,61].

From the unintegrated correlator in eq. (2.12) we can obtain the TMD correlator upon integration over k^-. The quark-quark TMD correlator for spin-1 hadrons is given by[4]

$$\Phi_{\alpha\beta}^{[U]}(x, \boldsymbol{k}, S, T) \equiv \int \frac{dz^- d^2\boldsymbol{z}}{(2\pi)^3} \, e^{ik\cdot z} \, \langle P; S, T | \, \overline{\psi}_\beta(0) \, U_{[0,z]} \, \psi_\alpha(z) \, | P; S, T \rangle \bigg|_{z^+=0}. \quad (2.26)$$

Since this correlator is evaluated at $z^+ = 0$, it is also referred to as the light-front correlator. The gauge link that is required for gauge invariance is discussed in detail in subsection 2.3.3 (the dependence on the gauge link will be implicit for the most part). As will be briefly discussed in section 3.5, eq. (2.26) is actually not the full definition of the TMD correlator – there is also a dependence on soft radiation. More specifically, the bilocal matrix element needs to be accompanied by a soft factor that removes so-called rapidity divergences [30,62].[5] With regards to hadron spin, the correlator in eq. (2.26) is defined as

$$\Phi_{\alpha\beta}(x, \boldsymbol{k}, S, T) \equiv \text{Tr} \left(\rho(S, T) \, M_{\alpha\beta}(x, \boldsymbol{k}) \right), \quad (2.27)$$

where the information on the spin states of the hadron is encoded in the 3×3 density matrix ρ and the combined information on the hadron and quark spins is contained in M. The latter quantity can be regarded as a 6×6 matrix in parton $\otimes$ hadron spin space, which is discussed in detail for the case of gluons in section 5.2. The density matrix can be parametrized in terms of S and T as follows:

$$\rho(S, T) = \frac{1}{3} \left(I + \frac{3}{2} S^i \Sigma^i + 3 T^{ij} \Sigma^{ij} \right), \quad (2.28)$$

where we have chosen a Cartesian basis of 3×3 matrices consisting of the identity matrix I, the three matrices Σ^i (these are a generalization of the Pauli matrices to the three-dimensional case), and the five matrices $\Sigma^{ij} \equiv \frac{1}{2}(\Sigma^i\Sigma^j + \Sigma^j\Sigma^i) - \frac{2}{3}I \delta^{ij}$. Making use of eqs. (2.24) and (2.25), the density matrix takes the explicit form [60,61]:

$$\rho(S, T) = \begin{pmatrix} \frac{1}{3} + \frac{S_L}{2} + \frac{S_{LL}}{3} & \frac{S_T^1 - iS_T^2}{2\sqrt{2}} + \frac{S_{LT}^1 - iS_{LT}^2}{2\sqrt{2}} & \frac{S_{TT}^{11} - iS_{TT}^{12}}{2} \\ \frac{S_T^1 + iS_T^2}{2\sqrt{2}} + \frac{S_{LT}^1 + iS_{LT}^2}{2\sqrt{2}} & \frac{1}{3} - \frac{2S_{LL}}{3} & \frac{S_T^1 - iS_T^2}{2\sqrt{2}} - \frac{S_{LT}^1 - iS_{LT}^2}{2\sqrt{2}} \\ \frac{S_{TT}^{11} + iS_{TT}^{12}}{2} & \frac{S_T^1 + iS_T^2}{2\sqrt{2}} - \frac{S_{LT}^1 + iS_{LT}^2}{2\sqrt{2}} & \frac{1}{3} - \frac{S_L}{2} + \frac{S_{LL}}{3} \end{pmatrix}. \quad (2.29)$$

---

[4] To obtain the correlator for a spin-0 or spin-$\frac{1}{2}$ hadron, one simply sets $S = T = 0$ or $T = 0$, respectively.
[5] The inclusion of a soft factor brings in a dependence on a rapidity scale. The dependence of TMDs on both renormalization and rapidity scales will be implicit throughout this thesis.

As we will see in the next subsection, the inclusion of hadron spin in the correlator gives rise to a wealth of different TMDs. This is because TMDs can encode correlations between spin components and partonic transverse momentum.

2.3.2 *Parametrization of the TMD correlator*

Since a hadronic correlator cannot be calculated using perturbative methods, it is parametrized in terms of PDFs. The correlator in eq. (2.26) can be parametrized in terms of quark TMDs. For unpolarized and vector polarized hadrons, the first parametrizations were given in the 1990s in [35,63], and in 2000 also tensor polarized hadrons were considered [60] (the collinear case was studied already in 1988 in [64]). For the quark TMDs we will adopt the notation of [60], which coincides with the notation in [35,63] for the unpolarized and vector polarized cases. This is often referred to as the 'Amsterdam' notation.

It is easiest to start the parametrization at the unintegrated level and then perform the integration over k^-. The correlator is a matrix in Dirac space and hence it can be parametrized in terms of Dirac matrices. A convenient basis of Dirac structures is given by the set

$$\{I, \gamma^\mu, \sigma^{\mu\nu}, \gamma^\mu\gamma^5, i\gamma^5\}, \tag{2.30}$$

where $\sigma^{\mu\nu} \equiv \frac{i}{2}[\gamma^\mu, \gamma^\nu]$. The various Dirac structures can be contracted with the available vectors and tensors k, P, S, T, and n.[6] The correlator is constrained by requirements that arise from Hermiticity and parity. For the unintegrated correlator, they are respectively given by [60]:

$$\Phi(k, P, S, T, n) = \gamma^0\, \Phi^\dagger(k, P, S, T, n)\, \gamma^0, \tag{2.31}$$

$$\Phi(k, P, S, T, n) = \gamma^0\, \Phi(\bar{k}, \bar{P}, -\bar{S}, \bar{T}, \bar{n})\, \gamma^0, \tag{2.32}$$

where we have introduced the notation $\bar{a}^\mu \equiv \delta^{\mu\nu} a_\nu$ and $\bar{b}^{\mu\nu} \equiv \delta^{\mu\rho}\delta^{\nu\sigma} b_{\rho\sigma}$. By contrast, time reversal (T) does *not* constrain the form of the parametrization; this will be explained in the next subsection.

Parametrizations of hadronic correlators in terms of TMDs are usually given in momentum space (or k_T-space). However, for certain applications such as the implementation of TMD evolution, correlators are studied in coordinate space instead (or z_T-space, where z_T is Fourier conjugate to k_T). To ensure a one-to-one correspondence between TMDs in k_T-space and z_T-space, it is essential that a correlator is parametrized in terms of TMDs of *definite rank*, which has not been done consistently in existing parametrizations. To this end, a correlator in k_T-space needs to be parametrized using symmetric traceless tensors in k_T (see appendix B for the definitions of these tensors up to rank 4). Here we will present the parametrization

---

6 The four-vector n was not included in the original parametrizations in [35,60,63] but, as was argued in [65], it should be included as it is ultimately connected to the direction of the Wilson line. However, as it turns out, it will not give rise to any new functions at leading twist [66].

of the TMD correlator in terms of symmetric traceless tensors given in our work in [67], which is based on the original parametrization in [60]. Separating the various possible hadronic polarization states, the correlator in eq. (2.26) can be parametrized in terms of leading-twist quark TMDs of definite rank as follows:[7]

$$\Phi(x,k) = \Phi_U(x,k) + \Phi_L(x,k) + \Phi_T(x,k)$$
$$+ \Phi_{LL}(x,k) + \Phi_{LT}(x,k) + \Phi_{TT}(x,k), \qquad (2.33)$$

with

$$\Phi_U(x,k) = \frac{1}{2}\left[\slashed{n} \, f_1(x,k^2) + \frac{\sigma_{\mu\nu}k_T^\mu \bar{n}^\nu}{M} \, h_1^\perp(x,k^2) \right], \qquad (2.34)$$

$$\Phi_L(x,k) = \frac{1}{2}\left[\gamma^5 \slashed{n} \, S_L \, g_1(x,k^2) + \frac{i\sigma_{\mu\nu}\gamma^5 \bar{n}^\mu k_T^\nu S_L}{M} \, h_{1L}^\perp(x,k^2) \right], \qquad (2.35)$$

$$\Phi_T(x,k) = \frac{1}{2}\left[\frac{\slashed{n} \, \epsilon_T^{S_T k_T}}{M} \, f_{1T}^\perp(x,k^2) + \frac{\gamma^5 \slashed{n} \, k\cdot S_T}{M} \, g_{1T}(x,k^2) \right.$$
$$\left. + i\sigma_{\mu\nu}\gamma^5 \bar{n}^\mu S_T^\nu \, h_1(x,k^2) - \frac{i\sigma_{\mu\nu}\gamma^5 \bar{n}^\mu k_T^{\nu\rho} S_{T\rho}}{M^2} \, h_{1T}^\perp(x,k^2) \right], \qquad (2.36)$$

$$\Phi_{LL}(x,k) = \frac{1}{2}\left[\slashed{n} \, S_{LL} \, f_{1LL}(x,k^2) + \frac{\sigma_{\mu\nu}k_T^\mu \bar{n}^\nu S_{LL}}{M} \, h_{1LL}^\perp(x,k^2) \right], \qquad (2.37)$$

$$\Phi_{LT}(x,k) = \frac{1}{2}\left[\frac{\slashed{n} \, k\cdot S_{LT}}{M} \, f_{1LT}(x,k^2) + \frac{\gamma^5 \slashed{n} \, \epsilon_T^{S_{LT} k_T}}{M} \, g_{1LT}(x,k^2) \right.$$
$$\left. + \sigma_{\mu\nu}\bar{n}^\nu S_{LT}^\mu \, h_{1LT}(x,k^2) - \frac{\sigma_{\mu\nu}\bar{n}^\nu k_T^{\mu\rho} S_{LT\rho}}{M^2} \, h_{1LT}^\perp(x,k^2) \right], \qquad (2.38)$$

$$\Phi_{TT}(x,k) = \frac{1}{2}\left[\frac{\slashed{n} \, k_T^{\mu\nu} S_{TT\mu\nu}}{M^2} \, f_{1TT}(x,k^2) - \frac{\gamma^5 \slashed{n} \, \epsilon_{T\mu\nu}k_T^{\mu\rho} S_{TT\rho}^\nu}{M^2} \, g_{1TT}(x,k^2) \right.$$
$$\left. - \frac{\sigma_{\mu\nu}\bar{n}^\nu k_T^\rho S_{TT\rho}^\mu}{M} \, h_{1TT}(x,k^2) + \frac{\sigma_{\mu\nu}\bar{n}^\nu k_T^{\mu\rho\sigma} S_{TT\rho\sigma}}{M^3} \, h_{1TT}^\perp(x,k^2) \right], \qquad (2.39)$$

where we have employed the notation $\epsilon_T^{ab} \equiv \epsilon_T^{\mu\nu} a_\mu b_\nu$, with $\epsilon_T^{\mu\nu} \equiv \epsilon^{\mu\nu-+}$ (its nonzero components are $\epsilon_T^{12} = -\epsilon_T^{21} = 1$). To avoid clutter, we suppress in the names of the functions a reference to quarks as it should be clear from the context whether we are discussing quark or gluon TMDs. As we will motivate in the next subsection, TMDs depend on the precise form of the gauge link that is used. In appendix E we provide the equivalent parametrization of the TMD correlator in coordinate space.

---

[7] In most of this thesis, the dependence of correlators on P, S, T, or n is implicit.

In table 2.1 we have organized the quark TMDs appearing in eq. (2.33) by hadron and quark polarizations. We have also indicated the behavior under time reversal (the difference between T-even and T-odd functions will be explained in subsection 2.3.3) and which TMDs have a collinear counterpart. The polarization of the quarks is projected out by the Dirac structures that come with the functions; we use the letters f, g, and h to indicate unpolarized, longitudinally polarized, and transversely polarized quarks, respectively.

	γ^+	$\gamma^+\gamma^5$	$i\sigma^{j+}\gamma^5$
U	$\mathbf{f_1}$		$h_1^\perp$
L		$\mathbf{g_1}$	$h_{1L}^\perp$
T	$f_{1T}^\perp$	g_{1T}	$\mathbf{h_1}, h_{1T}^\perp$
LL	$\mathbf{f_{1LL}}$		$h_{1LL}^\perp$
LT	f_{1LT}	g_{1LT}	$h_{1LT}, h_{1LT}^\perp$
TT	f_{1TT}	g_{1TT}	$h_{1TT}, h_{1TT}^\perp$

Table 2.1: An overview of the leading-twist quark TMDs for unpolarized (U), vector polarized (L or T), and tensor polarized (LL, LT, or TT) hadrons. The functions indicated in boldface also occur as collinear PDFs, and the ones in red are T-odd. The Dirac structures γ^+, $\gamma^+\gamma^5$, and $i\sigma^{j+}\gamma^5$ correspond to unpolarized, longitudinally polarized, and transversely polarized quarks, respectively.

As with respect to existing literature our parametrization has been written completely in terms of symmetric traceless tensor structures, the functions that were not of definite rank have been replaced by functions that are of definite rank. Employing the common shorthand notation

$$f^{(n)}(x, \mathbf{k}^2) \equiv \left(\frac{\mathbf{k}^2}{2M^2}\right)^n f(x, \mathbf{k}^2), \tag{2.40}$$

the functions h_{1T}, h'_{1LT}, and h'_{1TT} that appear in the original parametrizations in [35, 60] have been replaced by the following definite-rank TMDs:[8]

$$h_1 \equiv h_{1T} + h_{1T}^{\perp(1)}, \qquad h_{1LT} \equiv h'_{1LT} + h_{1LT}^{\perp(1)}, \qquad h_{1TT} \equiv h'_{1TT} + h_{1TT}^{\perp(1)}, \tag{2.41}$$

where, as we will do often, we suppressed the arguments of the functions to lighten the notation. An additional benefit of using definite-rank TMDs is that the functions also appearing in the collinear case, i.e. those that survive integration over $\mathbf{k}$, are

---

8 These definitions were already proposed in [60], so there is no conflict of notation here.

simply the rank-0 functions (the function h_{1LT} is an exception to this, though; this will be commented on in the next subsection). Furthermore, we have defined $g_1 \equiv g_{1L}$.

Some of the TMDs also go by a name, especially those appearing in the description of a spin-$\frac{1}{2}$ hadron. The most well-known T-even TMDs are f_1, g_1, and h_1, referred to as the unpolarized, helicity, and transversity functions, respectively. The unpolarized distribution represents the number density of unpolarized quarks with momentum fraction x and transverse momentum $\mathbf{k}$ inside an unpolarized hadron. The helicity distribution, on the other hand, represents the number density of quarks with momentum fraction x and transverse momentum $\mathbf{k}$ and polarization parallel to that of the hadron *minus* the number density of quarks with the same momentum but antiparallel polarization. The interpretation of h_1 is similar to that of g_1, but involves transversely polarized quarks instead. These three functions also have collinear counterparts [58] that bear the same names. We will discuss T-odd functions in the next subsection.

As will be explained in detail in chapter 5.2 for the case of gluons, in general TMDs satisfy certain bounds (also called positivity bounds) that follow from the operator structure of the correlator. In the quark case they arise from the positive semidefiniteness of the matrix M appearing in eq. (2.27). Bounds on quark TMDs can be found in [68] for spin-$\frac{1}{2}$ hadrons and in [69] for spin-1 hadrons. For example, the two TMDs appearing in the parametrization for unpolarized hadrons in eq. (2.34) satisfy the following inequality:

$$\frac{|\mathbf{k}|}{M} \, |h_1^{\perp}| \leq f_1 \, , \tag{2.42}$$

which provides an upper bound for the magnitude of the function $h_1^{\perp}$.

2.3.3 *Process dependence*

As explained in subsection 2.2.1, ensuring gauge invariance of the correlator requires the inclusion of a gauge link. In the collinear case, discussed in subsection 2.2.2, we saw that the path of the gauge link is fixed; it runs along the minus direction at transverse position $\eta = 0$. However, now the situation is different. In the TMD case, the gauge link is no longer unique as the quark fields are not only separated in the minus direction but *also* in the transverse directions. This means that the gauge link necessarily includes a transverse piece [70,71].

The two most important gauge links that will be discussed in this thesis are the so-called future- and past-pointing gauge links (also referred to as the 'staple-like' gauge links) denoted by $U_{[0,z]}^{[+]}$ and $U_{[0,z]}^{[-]}$, respectively. They are defined as

$$U_{[0,z]}^{[\pm]} \equiv U_{[0,\pm\infty;0]}^{n} \, U_{[\pm\infty;0,z]}^{T} \, U_{[\pm\infty,z^-;z]}^{n} \, . \tag{2.43}$$

The links along the minus direction are defined in eq. (2.18) and the link along the transverse directions (indicated by the superscript T) is given by

$$U^T_{[\pm\infty;0,z]} \equiv \mathcal{P} \exp\left(ig \int_0^z d\boldsymbol{\eta} \cdot \boldsymbol{A}(\eta^+ = 0, \eta^- = \pm\infty, \boldsymbol{\eta})\right). \qquad (2.44)$$

The integration paths of the staple-like gauge links are illustrated in figure 2.7.
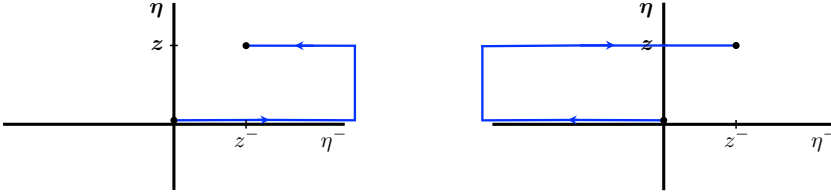
Figure 2.7: The integration paths of the future-pointing gauge link $U^{[+]}_{[0,z]}$ (on the left) and of the past-pointing gauge link $U^{[-]}_{[0,z]}$ (on the right). The paths run along the light-front $\eta^+ = 0$ via plus or minus light-cone infinity, respectively.

Any path connecting the fields ensures gauge invariance, however, as mentioned already in subsection 2.2.1, different paths reflect different physics and turn out to correspond to different processes. The process dependence of the gauge link (and thus of the TMD correlator) originates from the color structure of the hard scattering process. In general, the more partons participate in the hard scattering, the more complicated the resulting gauge links [72]. The future- and past-pointing gauge links of eq. (2.43) are obtained in the SIDIS and DY processes, respectively. In DY the color flows into the initial state (the gluon interactions happen *before* the quarks are annihilated), whereas in SIDIS the color flows into the final state (the gluon interactions happen *after* the quark is struck by the photon). For this reason, the associated gauge links are also referred to as initial- and final-state Wilson lines, respectively.

In the previous subsection we discussed the constraints on the correlator that arise from Hermiticity and parity, see eqs. (2.31) and (2.32). Now, let us discuss the effects of time reversal transformations on the correlator (we will only consider staple-like gauge links here). Time reversal leads to the following constraint [60]:

$$\Phi^{[\pm]*}(k, P, S, T, n) = i\gamma^1\gamma^3 \, \Phi^{[\mp]}(\bar{k}, \bar{P}, \bar{S}, \bar{T}, \bar{n}) \, i\gamma^1\gamma^3. \qquad (2.45)$$

Note that under time reversal a future-pointing gauge link becomes a past-pointing one and vice versa. This means that time reversal does not constrain a single correlator – it relates *two different* correlators with different gauge links. Hence, unlike Hermiticity and parity, time reversal does not actually constrain the parametrization of the TMD correlator in terms of TMDs.

From eq. (2.45) it follows that we can define T-even and T-odd contributions to the correlator according to the behavior

$$\Phi^{(\text{T-even})*}(k, P, S, T, n) = i\gamma^1\gamma^3\,\Phi^{(\text{T-even})}(\bar{k}, \bar{P}, \bar{S}, \bar{T}, \bar{n})\,i\gamma^1\gamma^3, \qquad (2.46)$$

$$\Phi^{(\text{T-odd})*}(k, P, S, T, n) = -i\gamma^1\gamma^3\,\Phi^{(\text{T-odd})}(\bar{k}, \bar{P}, \bar{S}, \bar{T}, \bar{n})\,i\gamma^1\gamma^3, \qquad (2.47)$$

where the T-even and T-odd parts of Φ are given by

$$\Phi^{(\text{T-even})}(k, P, S, T, n) = \frac{1}{2}\left[\Phi^{[+]}(k, P, S, T, n) + \Phi^{[-]}(k, P, S, T, n)\right], \qquad (2.48)$$

$$\Phi^{(\text{T-odd})}(k, P, S, T, n) = \frac{1}{2}\left[\Phi^{[+]}(k, P, S, T, n) - \Phi^{[-]}(k, P, S, T, n)\right]. \qquad (2.49)$$

This means that we can classify TMDs as being either T-even or T-odd, see table 2.1. The most well-known T-odd functions are $h_1^\perp$, called the Boer-Mulders (BM) function [63], and $f_{1T}^\perp$, called the Sivers function [73,74]. The former describes a transversely polarized quark inside an unpolarized hadron, whereas the latter describes an unpolarized quark inside a transversely polarized hadron.

The existence of T-odd functions completely relies on the gauge link structure – after all, if we were to naively set the gauge links to unity,[9] these functions would be *forbidden* by time reversal symmetry. For that reason, T-odd functions were initially thought to vanish (for more than a decade), until a model calculation by Brodsky, Hwang, and Schmidt (BHS) in 2002 showed that the Sivers function could be nonzero [75]. A few years later, in 2004, a nonzero quark Sivers function was measured for the first time by the HERMES experiment at DESY [76]. The quark BM and Sivers functions will play an important role in chapter 4.

Since T-odd TMDs change sign under time reversal (which amounts to swapping future- and past-pointing gauge links), their signs actually depend on whether they are probed in SIDIS or in DY. This was first pointed out by Collins in a seminal paper [77] a few months after the BHS calculation came out. The sign flip of T-odd TMDs between SIDIS and DY is one of the most striking predictions of the TMD formalism. By contrast, T-even quark TMDs are the same in both processes. For example, the unpolarized and BM functions that appear in the parametrization for unpolarized hadrons in eq. (2.34) are related in SIDIS and DY as

$$f_1^{[+]} = f_1^{[-]}, \qquad h_1^{\perp[+]} = -h_1^{\perp[-]}. \qquad (2.50)$$

These relations follow from the parity and time reversal constraints given in eqs. (2.32) and (2.45), respectively. The process dependence of TMDs (both T-even and T-odd ones) becomes more involved in processes with more complicated gauge links [78–80]. Although for a given process the gauge link can always be calculated, it is not always possible to relate TMDs between different processes through a constant factor as in

---

9 In a light-cone gauge, the transverse links do *not* reduce to unity [70].

eq. (2.50) [72]. We will consider an example of the latter for the case of gluons in section 5.3.

One might wonder whether the described nonuniversality of TMDs can spoil factorization. As it turns out, the factorized structure of a cross section can remain valid as long as TMDs are defined with the appropriate process-dependent gauge links (this is sometimes referred to as *generalized* factorization). This for example means that we take future-pointing links for SIDIS and past-pointing ones for DY. In section 3.6 we mention a process for which this procedure potentially fails.

A different way in which PDFs in general and TMDs in particular depend on the hard process under consideration is through the energy scale at which the hard scattering occurs. The energy scale dependence (as well as the dependence on the rapidity scale) can be calculated perturbatively, though. This means that TMDs that are extracted in one process at a particular scale can be used in other processes by evolving them to the appropriate scale. Hence, the violations of universality, both due to the dependence on the color structure and the energy scale of the hard process, are calculable – they are, in a sense, 'under control', enabling global extractions of TMDs.

Upon integrating the TMD correlator over k, we recover the collinear situation, in which case the staple-like gauge links reduce to the unique *process-independent* link $U^n_{[0,z^-;\mathbf{0}]}$ given in eq. (2.18). It then follows from eq. (2.49) that all T-odd functions vanish. Consequently, all collinear PDFs must be T-even. Even though the T-odd, rank-0 Bacchetta function h_{1LT} survives integration over k, it in fact vanishes due to the combined effects of Hermiticity and time reversal constraints [26], as it should.

Finally, let us briefly comment on the gauge link (in)dependence of the other multidimensional distributions that appear in the scheme in figure 2.5. Not only TMDs suffer from complications regarding universality – in fact, *all* distributions that have a dependence on k are in principle process dependent. Thus, besides TMDs, also GTMDs and Wigner distributions come with a process-dependent gauge link. Similarly to the collinear correlator, the GPD correlator features a Wilson line with a unique straight path directed along the minus direction at $\eta = 0$. Hence, just as collinear PDFs, also GPDs are universal.

2.3.4 *Status of extractions*

So far, our discussion on TMDs has been rather theoretical. In this subsection we briefly comment on the experimental side of TMD phenomenology. We will focus here on quark rather than gluon TMDs (the latter will be covered in section 5.4).

As mentioned already in the introduction, TMD effects typically give rise to angular correlations between the produced particles. An example of an angular correlation is the well-known $\cos(2\phi)$ azimuthal asymmetry in both the SIDIS and DY cross sections as induced by the quark BM function $h_1^\perp$ (see section 3.2 for more details). The BM function can be extracted through a measurement of the azimuthal

angular distribution of the final-state particles. Since $h_1^\perp$ encodes a correlation between partonic transverse momentum and partonic transverse spin, its extraction provides access to the internal (transverse) structure of hadrons. The experimental status of the BM function will be discussed in section 4.1.

One of the vital tests of the TMD formalism is to experimentally verify the predicted sign change of T-odd functions between SIDIS and DY. A nonzero quark Sivers function $f_{1T}^\perp$ has been extracted from SIDIS measurements by various experiments, see e.g. [81–87]. Recently, $f_{1T}^\perp$ was also extracted from W boson production and DY measurements (both processes come with a future-pointing gauge link): both the STAR experiment at BNL [88] and the COMPASS experiment at CERN [89] show results that favor a sign change, but more data is needed to provide conclusive evidence. Due to its profound origin (gauge invariance and factorization theorems), a falsification of the expected sign flip would imply that there is something very general about QCD that we do not understand, or (probably more likely) that something has been overlooked in the factorization proofs.

Of all TMDs, the most well-known one is the unpolarized quark TMD f_1. In fact, recently f_1 was extracted from a fit to available SIDIS and DY data [90] – the first *global* extraction of a TMD. To this end, TMD evolution at next-to-leading logarithmic accuracy was implemented to connect the data at different scales.

Most TMDs are (virtually) unknown, though. The future Electron-Ion Collider (EIC), which is currently considered in the United States, might change this and revolutionize the field of TMD phenomenology [91,92]. At high luminosity and high energy, it could scatter electrons off polarized protons and light nuclei, including deuterons. Since it will be able to probe small values of x, also sea quark and gluon TMDs could be extracted with much higher precision than what has been achieved so far. More information on the experimental status of TMDs can be found in e.g. [26,40,93–96].

Part 1

FACTORIZATION OF DRELL-YAN SCATTERING

3

ELEMENTS OF CSS FACTORIZATION

Part I of this thesis is centered around the concept of factorization, which forms the cornerstone of QCD phenomenology. In particular, we discuss factorization of the DY cross section differential in the transverse momentum of the lepton pair, involving one TMD for each colliding hadron – this will be referred to as TMD factorization. The DY cross section integrated over transverse momentum was shown to factorize in terms of a hard scattering factor and collinear PDFs (called collinear factorization) to all orders in perturbation theory in the 1980s by Collins, Soper, and Sterman (CSS) [97, 98]. This proof of factorization is highly nontrivial and contains many ingredients – in [98] it was remarked: "The complexity of such a proof will be comparable to that of proofs of renormalizability". In the TMD case, the factorization proof proceeds largely along the same lines and is extensively covered in [30]. The latter is *also* referred to as the CSS proof and forms the main topic of this chapter. Other approaches to TMD factorization, particularly in the context of soft-collinear effective theory (SCET) [62, 99, 100], will not be discussed here.

It is unclear whether TMD factorization of DY also holds for T-odd TMDs that are associated to azimuthal asymmetries; no explicit proof for this has been provided yet. In [101] it was suggested that in this case the factorization theorem, if it holds, would come with a modified color factor compared to the naive prediction based on the CSS works. In chapter 4 we will investigate this issue in detail by performing a factorization calculation largely along the lines of the CSS proof. To facilitate that calculation, we provide in this chapter the necessary ingredients. In particular, we discuss how leading-power regions of loop momenta can be identified and how contributions from those regions need to be combined; the latter procedure is given by the Collins subtraction scheme. Furthermore, we review the complications that arise from so-called Glauber modes. A significant part of the CSS proof is devoted to the unitarity cancellation of Glauber exchanges, which turns out to be an important aspect in our factorization study in chapter 4.

Because of its importance, factorization is a vast topic for which a lot of literature is available. The material presented in this chapter is largely based on treatments in Collins's book "Foundations of perturbative QCD" [30], our review in [2], as well as on reviews provided in [102–104]; other useful resources include [105–109].

3.1 CONCEPT OF FACTORIZATION

Factorization is a consequence of the approximation that long-distance dynamics describing the hadron structure decouples from short-distance dynamics which

describes a hard scattering event. The most well-known and well-established form of factorization is collinear factorization. This type of factorization applies to scattering processes in which all involved kinematic scales are large and of the same order; other energy scales, such as Λ_{QCD} and M, characterize the long-distance physics. An example of such a process is inclusive DY scattering in which the virtuality of the intermediate gauge boson sets the hard scale. The transverse momentum of the boson should either be of the same order as the hard scale or integrated over. Collinear factorization of DY means that its cross section factorizes into two collinear PDFs (one for each hadron), and a perturbatively calculable hard factor that describes the partonic scattering.

The idea of factorization arises very naturally. In DY the colliding hadrons and associated soft gluon fields are Lorentz contracted in the center-of-mass (CM) frame of the collision to very thin objects such that they essentially do not 'feel' each other prior to the collision. Furthermore, a large virtuality of the gauge boson implies that a hard scattering event happens on a very short time scale, while the soft physics that governs the internal structure of hadrons involves much longer time scales. Thus, intuitively, the hadrons appear 'frozen' with respect to the hard scattering event and therefore should remain unaffected. Hence, the short-distance dynamics is decoupled from the long-distance dynamics, with any interference effects suppressed by inverse powers of the hard scale. This picture also illustrates why the PDFs are expected to be universal – they pertain to hadrons and are independent of any hard scattering event. This makes factorization the cornerstone of QCD phenomenology; the PDFs that are extracted from one process can be used in others. Without factorization, QCD loses its predictive power in scattering processes.

Since factorization is such a vital component of QCD, a solid proof of this is essential. As it turns out, proving factorization is very challenging, for instance due to the fact that arbitrarily many soft gluons can be exchanged between the two hadrons. However, in the 1980s CSS managed to establish collinear factorization of DY scattering to all orders in perturbation theory [97,98]. The CSS proof holds up to leading power in the inverse hard scale. A few years later, the proof was extended to the first subleading power [110,111].

In most of this thesis and in particular in part I, we focus on another type of factorization called TMD factorization. This type of factorization, unlike collinear factorization, takes into account transverse momentum effects. TMD factorization is different due to the presence of two kinematic scales in the process that are of different order. One scale is provided by the virtuality of the gauge boson, Q^2, and the other by its noncollinearity, Q_T^2. If the latter is substantially smaller than the former, one could expand in their ratio and only consider the leading term (also logarithms of the form $\alpha_s \ln(Q/Q_T)$ need to be resummed). At leading power and to all orders in perturbation theory, TMD factorization for DY was established in [30]. This proof proceeds largely along the same lines as the one for the collinear case. Besides DY, TMD factorization has also been proven to hold for e^+e^- annihilation into two hadrons [112] and SIDIS [113], and has largely been completed for the

double DY process [102, 103, 114]. Note that all these processes involve two hadrons only.

3.2 THE DRELL-YAN PROCESS

In part I of this thesis we focus on inclusive DY scattering. The DY process was proposed in 1970 to describe the production of lepton-antilepton pairs in high-energy hadron-hadron collisions [115]. In the collision of two hadrons with momenta p_1 and p_2, a quark and antiquark with momenta k_1 and k_2 interact to produce a virtual photon (or Z boson) with momentum q. This gauge boson in turn decays into a charged lepton-antilepton pair with momenta l and l'. The leading-order diagram for this process is schematically illustrated in figure 3.1. Although for definiteness we will consider $p\bar{p}$ collisions in this part of the thesis, most discussions trivially apply to other hadrons as well.
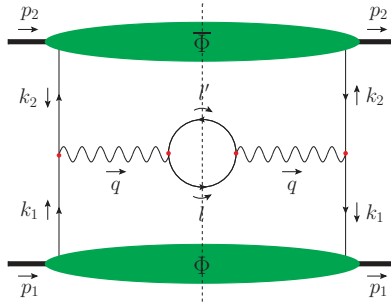
Figure 3.1: The tree-level graph for the DY process: a quark and antiquark from the colliding hadrons produce a virtual photon that subsequently decays into a charged lepton-antilepton pair. The green 'blobs' represent the quark and antiquark correlators.

The distribution of quarks and antiquarks inside the colliding protons and antiprotons is encoded by the quark and antiquark correlators Φ and $\overline{\Phi}$, respectively. As discussed in detail in subsection 2.3.2, these correlators can be parametrized in terms of TMDs. For unpolarized hadrons, the quark-quark TMD correlator can be parametrized in terms of two leading-twist quark TMDs, namely the T-even unpolarized function f_1 and the T-odd BM function $h_1^\perp$, see eq. (2.34) (we will denote the antiquark analogues with a bar). These TMDs depend on the longitudinal momentum fraction $x_1 \equiv k_1^+/p_1^+$, as well as on the transverse momentum k_1^2.

The TMD factorization theorem by CSS holds up to leading power in Λ/Q, where $Q^2 \equiv q^2 > 0$ represents the hard scale of the process and Λ represents the transverse momentum of the electroweak boson, $Q_T \equiv \sqrt{q^2}$, or the scale of nonperturbative

QCD interactions. For unpolarized hadrons, the factorization formula at leading order in the hard scattering takes the following form [116, 117]:

$$\frac{d\sigma}{d\Omega\, dx_1 dx_2\, d^2\boldsymbol{q}} = \frac{\alpha^2}{N_c\, q^2} \sum_q e_q^2 \Big\{ A(\theta)\, \mathcal{F}\left[f_1 \bar{f}_1\right]$$

$$+ B(\theta) \cos(2\phi)\, \mathcal{F}\left[w(\boldsymbol{k}_1, \boldsymbol{k}_2)\, h_1^\perp \bar{h}_1^\perp\right] \Big\}, \tag{3.1}$$

with the convolution of TMDs defined as:

$$\mathcal{F}\left[f_1 \bar{f}_1\right] \equiv \int d^2\boldsymbol{k}_1 \int d^2\boldsymbol{k}_2\, \delta^{(2)}(\boldsymbol{k}_1 + \boldsymbol{k}_2 - \boldsymbol{q})\, f_{1,q}(x_1, \boldsymbol{k}_1^2)\, \bar{f}_{1,q}(x_2, \boldsymbol{k}_2^2). \tag{3.2}$$

The functions A and B are given by

$$A(\theta) = \frac{1}{4}(1 + \cos^2\theta), \qquad B(\theta) = \frac{1}{4}\sin^2\theta, \tag{3.3}$$

and the BM weight function reads

$$w(\boldsymbol{k}_1, \boldsymbol{k}_2) = \frac{2(\hat{h}\cdot\boldsymbol{k}_1)(\hat{h}\cdot\boldsymbol{k}_2) - \boldsymbol{k}_1\cdot\boldsymbol{k}_2}{M^2}. \tag{3.4}$$

The cross section is given in terms of the Collins-Soper (CS) angles θ and ϕ [118], see figure 3.2. The unit vector $\hat{h}$ is defined in the hadron CM frame as $\hat{h} \equiv \boldsymbol{q}/|\boldsymbol{q}|$, and $p_1^2 = p_2^2 = M^2$, where M is the mass of the hadron, as before. The sum in eq. (3.1) runs over the different quark flavors labeled by q. Furthermore, the electrical charge e_q is given in units of the elementary charge e, and α denotes the fine-structure constant.
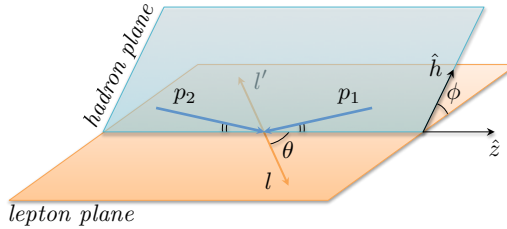
Figure 3.2: The definition of the angles θ and ϕ in the CS frame. This frame is a lepton CM frame in which the hadrons come in symmetrically with respect to the z-axis.

We will refer to the first term in eq. (3.1) as the 'double unpolarized' contribution, because it involves unpolarized quarks, and the second term will be referred to as the 'double BM' (dBM) contribution and involves transversely polarized quarks. The dBM term comes with a $\cos(2\phi)$ dependence which can be used to isolate this contribution in the experimental data. The BM function $h_1^\perp$, as opposed to the unpolarized TMD f_1, is called chiral-odd as it flips the quark helicity between

amplitude and conjugate amplitude. Since helicity is conserved (in the massless limit), the BM function needs to be coupled to another chiral-odd function in the cross section. Hence, for the unpolarized DY process, $h_1^\perp$ can only appear together with $\bar{h}_1^\perp$.

The quark TMDs f_1 and $h_1^\perp$ appearing in eq. (3.1) have the following operator definitions:

$$f_{1,q}(x_1, \boldsymbol{k}_1^2) \equiv \int \frac{dz^- d^2\boldsymbol{z}}{(2\pi)^3} e^{ik_1 \cdot z} \langle p_1 | \bar{\psi}_q(0) U_{[0,z]}^{[-]} \Gamma_U \psi_q(z) | p_1 \rangle \bigg|_{z^+=0}, \qquad (3.5)$$

$$\frac{\tilde{k}_{1T}^j}{M} h_{1,q}^\perp(x_1, \boldsymbol{k}_1^2) \equiv \int \frac{dz^- d^2\boldsymbol{z}}{(2\pi)^3} e^{ik_1 \cdot z} \langle p_1 | \bar{\psi}_q(0) U_{[0,z]}^{[-]} \Gamma_T^j \psi_q(z) | p_1 \rangle \bigg|_{z^+=0}, \qquad (3.6)$$

where we have employed the notation $\tilde{a}_T^\nu \equiv \epsilon_T^{\mu\nu} a_{T\mu}$. The necessary soft factors are not considered explicitly as they will not play a role in our factorization calculation in chapter 4. As appropriate for the DY process (see section 2.3.3), both TMDs come with the past-pointing gauge link $U_{[0,z]}^{[-]}$, defined in eq. (2.43). Furthermore, Γ_U and Γ_T^j are Dirac projectors that, respectively, select unpolarized and transversely polarized quarks:

$$\Gamma_U \equiv \tfrac{1}{2}\gamma^+, \qquad \Gamma_T^j \equiv \tfrac{1}{2}i\,\sigma^{j+}\gamma^5, \qquad (3.7)$$

where j is a transverse index. In the TMD definitions (3.5) and (3.6) a summation over color is implicitly assumed. Since only quarks and antiquarks of the same color can annihilate, the factorization formula, eq. (3.1), appears with a $1/N_c$ color factor. The definitions of the unpolarized and BM functions for the antiquark are analogous to the quark case.

In the next sections we discuss various ingredients of the CSS factorization proof that are essential for our factorization calculation in chapter 4. Although we do not discuss different quark or hadron polarizations explicitly, the CSS proof is intended to apply to all polarized cases. In section 3.6 we will come back to this statement.

3.3 MOMENTUM REGIONS ANALYSIS

The TMD factorization approach is based on an expansion in powers of the small parameter Λ/Q and, today, factorization of the DY cross section has only been established up to leading power in this parameter. In this section we discuss how leading graph contributions are found. For each graph that contributes to the cross section, the Libby-Sterman method can be used to identify the leading regions of loop momentum [119, 120]. Since the details of this method are not relevant for our factorization study in chapter 4, we will only summarize its main results here; a detailed review of this method can be found in [30]. Subsequently, in subsection 3.3.2, we will discuss the Collins subtraction scheme that describes how the various leading region contributions should be combined to reconstruct the full graph.

3.3.1 *Identifying leading regions*

Naively, one expects dominant contributions to loop integrations for a given graph to originate from momentum regions around points at which certain lines go on shell, as in that case the propagator denominators vanish.[1] However, if the pole structure of the propagators is such that the Feynman integrals (that run along real axes) can avoid these singularities, then, according to Cauchy's theorem, the integration contours can be deformed into the complex plane away from the poles, giving finite integrals. Hence the following condition arises: the leading momentum regions must correspond to situations where the integration contours cannot be deformed away from the on-shell poles. This is the case when two poles on opposite sides of the contour are merged into a single point upon setting to zero all quantities of order Λ in the diagram; the contour is then 'pinched', prohibiting a deformation away from the singularity. The conditions on the loop momenta that lead to a pinch define a surface in multidimensional momentum space, called the pinch surface. Leading momentum regions are each associated to such a surface. For a given diagram, the pinch surfaces can be found with the Landau equations [121], aided by the Coleman-Norton method [122]. Since we will not make use of these techniques in chapter 4, we will not discuss them further here.

Subsequently, a power counting analysis is used to determine which momentum scalings give leading contributions to the DY graphs. We choose a coordinate system in the hadron CM frame where both incoming hadrons have zero transverse momentum, with one hadron moving fast in the $+\hat{z}$ direction (the 'right-moving' hadron) and the other moving fast in the $-\hat{z}$ direction (the 'left-moving' hadron). For the DY process, it follows from the Libby-Sterman method that the potentially relevant regions of loop momentum ℓ are given by [30, 119, 120]:[2]

$$\text{hard } (H): \qquad \ell \sim (1,1,1)Q, \qquad\qquad (3.8)$$

$$\text{right-moving collinear } (C_1): \quad \ell \sim (1,\lambda^2,\lambda)Q, \qquad (3.9)$$

$$\text{left-moving collinear } (C_2): \quad \ell \sim (\lambda^2,1,\lambda)Q, \qquad (3.10)$$

$$\text{central soft } (S): \qquad \ell \sim (\lambda,\lambda,\lambda)Q, \qquad\qquad (3.11)$$

$$\text{central ultrasoft } (U): \qquad \ell \sim (\lambda^2,\lambda^2,\lambda^2)Q, \qquad (3.12)$$

$$\text{Glauber}: \qquad |\ell^+\ell^-| \ll \boldsymbol{\ell}^2 \ll Q^2, \qquad (3.13)$$

where λ is a small parameter of order Λ/Q. In the CSS approach, the soft and ultrasoft regions are treated together and in this chapter we often use 'soft' to refer to both the central soft and central ultrasoft regions simultaneously; in chapter 4, however, we distinguish between these two modes. The Glauber condition (3.13)

---

1 An exception to this is the hard region defined in (3.8), which can contribute at leading power due to the volume of phase space [30].

2 The term 'central' refers to the central rapidity region. The rapidity y of a four-vector ℓ is defined as $y \equiv \frac{1}{2}\ln(|\ell^+/\ell^-|)$.

allows for various possible scalings that are in fact all treated together in the CSS methodology. The Glauber scalings that are important in the analysis of chapter 4 are:

right-moving Glauber (G_1): $\quad \ell \sim (\lambda, \lambda^2, \lambda)Q,$ $\qquad\qquad$ (3.14)

left-moving Glauber (G_2): $\quad \ell \sim (\lambda^2, \lambda, \lambda)Q,$ $\qquad\qquad$ (3.15)

central Glauber (G): $\quad \ell \sim (\lambda^2, \lambda^2, \lambda)Q.$ $\qquad\qquad$ (3.16)

The DY graphs can be organized into subgraphs that each characterize a particular momentum scaling. From the Libby-Sterman method it follows which connections between subgraphs give a leading contribution. In the TMD case, the dominant graphs for DY have the structure shown in figure 3.3. There are two collinear subgraphs: the one corresponding to the right-moving hadron is denoted by A and the other one by B. Furthermore, there is a hard subgraph denoted by H that is connected to both A and B by one fermion line and an arbitrary number of gluons. Lastly, there is a subgraph S that initially contains both soft and Glauber modes, which connect via soft and Glauber gluon attachments to either of the collinear subgraphs.
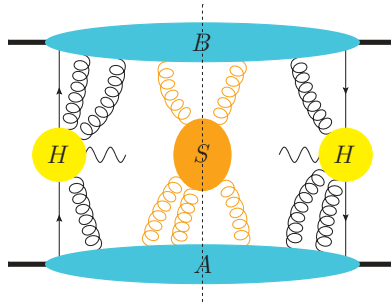
Figure 3.3: The partitioning of the leading-power DY graphs in the TMD case into various subgraphs (represented by 'blobs') that are each characterized by a particular momentum scaling. The right- and left-moving collinear subgraphs are respectively denoted by A and B, the soft (plus Glauber) subgraph by S, and the hard subgraph by H. For convenience we have suppressed the final-state leptons.

To get a feel for how the power counting works in practice for a simple DY graph, we now consider an example. To this end, we use the model that is discussed in section 14.3 of [30]: the hadrons are treated as being composed of a quark and an antiquark (the details are not relevant here). Let us consider a single-gluon exchange graph where the gluon is connected between an 'active' and a 'spectator' line (an active line is involved in the hard scattering while a spectator line is not).[3] The proton

---

3 A more general definition of active and spectator lines is given in [123].

and active quark carry C_1 momenta, whereas the antiproton and active antiquark carry C_2 momenta; to be precise,

$$p_1 \sim (1, \lambda^2, 0)Q, \qquad p_2 \sim (\lambda^2, 1, 0)Q, \tag{3.17}$$

$$k_1 \sim (1, \lambda^2, \lambda)Q, \qquad k_2 \sim (\lambda^2, 1, \lambda)Q. \tag{3.18}$$

To determine which gluon momentum scalings give a leading contribution to the graph, one first needs to define what is meant by 'leading'. By definition, the leading power is set by the tree-level graph.[4] Both the tree-level and single-gluon exchange graphs are given in figure 3.4. With respect to the tree-level graph there are three additional propagators in the single-gluon exchange graph, namely those pertaining to the momenta ℓ, $k_2 + \ell$, and $p_1 - k_1 + \ell$. Moreover, one propagator is modified in the amplitude: $k_1 \to k_1 - \ell$, which means that its scaling might change with respect to the tree-level graph (depending on the scaling of ℓ). In fact, it is easy to see that the scaling of this line is unchanged if ℓ^- is at most of order $\lambda^2 Q$ – then there is no 'penalty' factor. In that case, the exchanged gluon could give a leading contribution if

$$\int \frac{d^4\ell}{(2\pi)^4} \frac{i}{(p_1 - k_1 + \ell)^2 + i\epsilon} \frac{i}{\ell^2 + i\epsilon} \frac{i}{(k_2 + \ell)^2 + i\epsilon} \tag{3.19}$$

is of order 1 (for convenience we will set $Q = 1$ in this example).
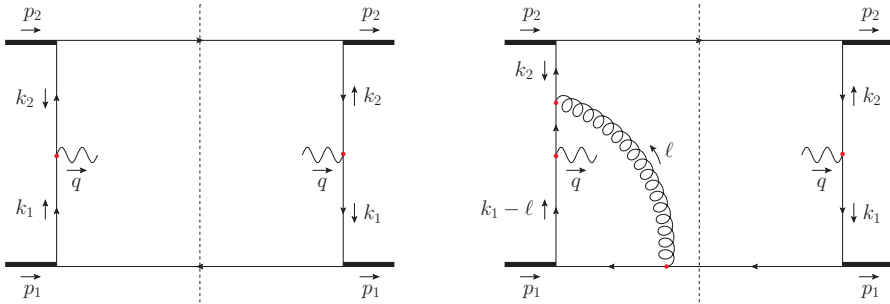
Figure 3.4: The diagram on the right features an active-spectator gluon exchange. Various gluon momentum regions contribute at leading power, which is set by the tree-level graph given on the left.

Let us now systematically consider all relevant regions defined in (3.8)–(3.12) for the gluon loop momentum ℓ, including the various Glauber scalings given in (3.14)–(3.16), to investigate which ones give a leading-power contribution to the graph:

- **The hard region (H).** Since $\ell \sim (1, 1, 1)$, we have $d^4\ell \sim 1$ and $\ell^2 \sim 1$. Furthermore, $(p_1 - k_1 + \ell)^2 \sim 1$ and $(k_2 + \ell)^2 \sim 1$ (note that in this example there

4 Sometimes a region gives a super-leading contribution to an individual graph. However, as it turns out, those contributions cancel generally upon summing over all graphs [30].

36

are no extra factors of λ coming from propagator numerators), and there is a penalty factor of order λ^2 coming from the $k_1 - \ell$ line. This means that the H region is suppressed with respect to the tree-level graph by a factor of order λ^2.

- **The collinear regions (C_1 and C_2).** For C_1, $\ell \sim (1, \lambda^2, \lambda)$, so that $d^4\ell \sim \lambda^4$ and $\ell^2 \sim \lambda^2$. Furthermore, $(p_1 - k_1 + \ell)^2 \sim \lambda^2$ and $(k_2 + \ell)^2 \sim 1$, which means that the expression in (3.19) is of order 1. Hence, the C_1 region gives a leading contribution. For the C_2 region, however, there is a penalty factor of order λ^2 coming from the $k_1 - \ell$ line, leading to an overall λ^2 suppression.

- **The central soft region (S).** Since $\ell \sim (\lambda, \lambda, \lambda)$, we have $d^4\ell \sim \lambda^4$ and $\ell^2 \sim \lambda^2$. Furthermore, $(p_1 - k_1 + \ell)^2 \sim \lambda$ and $(k_2 + \ell)^2 \sim \lambda$, and there is a penalty factor of order λ coming from the $k_1 - \ell$ line. This means that the S region is suppressed with respect to the tree-level graph by a factor of order λ.

- **The central ultrasoft region (U).** Since $\ell \sim (\lambda^2, \lambda^2, \lambda^2)$, we have $d^4\ell \sim \lambda^8$ and $\ell^2 \sim \lambda^4$. Furthermore, $(p_1 - k_1 + \ell)^2 \sim \lambda^2$ and $(k_2 + \ell)^2 \sim \lambda^2$, which means that the expression in (3.19) is of order 1. Hence, the U region gives a leading contribution.

- **The noncentral Glauber regions (G_1 and G_2).** For G_1, $\ell \sim (\lambda, \lambda^2, \lambda)$, so that $d^4\ell \sim \lambda^5$ and $\ell^2 \sim \lambda^2$. Furthermore, $(p_1 - k_1 + \ell)^2 \sim \lambda^2$ and $(k_2 + \ell)^2 \sim \lambda$, which means that the expression in (3.19) is of order 1. Hence, the G_1 region gives a leading contribution. For the G_2 region, however, there is a penalty factor of order λ coming from the $k_1 - \ell$ line, leading to an overall λ suppression.

- **The central Glauber region (G).** Since $\ell \sim (\lambda^2, \lambda^2, \lambda)$, we have $d^4\ell \sim \lambda^6$ and $\ell^2 \sim \lambda^2$. Furthermore, $(p_1 - k_1 + \ell)^2 \sim \lambda^2$ and $(k_2 + \ell)^2 \sim \lambda^2$, which means that the expression in (3.19) is of order 1. Hence, the G region gives a leading contribution.

From this power counting analysis it follows that the C_1, U, G_1, and G regions give rise to a leading-power contribution, whereas the other region contributions are power suppressed. However, unlike C_1 and U, the G_1 and G regions do *not* correspond to pinch surfaces. As will be discussed in detail in subsection 3.4.2, the ℓ^+ contours are not pinched in the Glauber regions, which allows for contour deformations into the C_1 region. This implies that there is no need to consider the Glauber regions separately from C_1.

To sum up, the (nontrivial) leading contributions to the single-gluon exchange diagram come from the U and C_1 regions. They are associated to the pinch surfaces defined by $\ell = 0$ and $\ell \propto k_1$, respectively. Furthermore, we note that the single-gluon exchange graph in figure 3.4 in fact contributes to *all* subgraphs defined in figure 3.3: the C_1 gluon (being a part of A) connects the A and H subgraphs (the $k_2 + \ell$ line is

hard), whereas the U gluon (being a part of S) connects the A and B subgraphs. The left-moving lines constitute the B subgraph.

Knowing the leading regions, we can approximate the complicated integral over the loop momentum ℓ. Denoting the single-gluon exchange graph by Γ (only the relevant part is considered here), we can write

$$\Gamma \equiv \int \frac{d^4\ell}{(2\pi)^4} \, J(\ell),$$

(3.20)

with

$$J(\ell) \equiv \frac{i}{(k_1 - \ell)^2 + i\epsilon} \, \frac{i}{(p_1 - k_1 + \ell)^2 + i\epsilon} \, \frac{i}{\ell^2 + i\epsilon} \, \frac{i}{(k_2 + \ell)^2 + i\epsilon} \, .$$

(3.21)

Up to power corrections, Γ is given by the sum of the two leading region contributions:

$$\Gamma = \int_U \frac{d^4\ell}{(2\pi)^4} \, J_U(\ell) + \int_{C_1} \frac{d^4\ell}{(2\pi)^4} \, J_{C_1}(\ell).$$

(3.22)

The subscript U on the integral denotes the restriction of the integration to the central ultrasoft region, and the same subscript on the integrand denotes the application of the leading-power approximation as appropriate for the central ultrasoft region (and likewise for the C_1 region). The restriction of the integration domain to a certain region turns out to be inconvenient in practice; we will come back to this in the next subsection. If we were instead to integrate $J_U(\ell)$ over *all* values of ℓ, then also contributions are picked up from regions where $J_U(\ell)$ fails to be a good approximation (and similarly for the C_1 region). In the next subsection we discuss how this problem can be solved by introducing appropriate subtraction terms to avoid any double counting.

3.3.2 *Collins subtraction method*

To obtain the leading-power contribution to the DY cross section, one needs to sum over all leading momentum regions for a given graph and subsequently sum over all graphs. In this subsection we focus on the former; we discuss how a graph can be reconstructed from its leading region contributions. This procedure is described by the Collins subtraction method that was introduced in [30] and reviewed in [103]; we follow the discussion of the latter.

Let us denote by Γ a particular Feynman graph (integrated over all its loop momenta) and by R a leading momentum region of Γ, specified by the set of scalings for all loop momenta. A region R can contain any of the hard, collinear, soft, and Glauber scalings as defined in (3.8)–(3.13). With each region R we associate an approximator T_R that approximates the integrand of Γ as appropriate for that region. We will employ an approximator that drops all terms in numerators and denominators that are power suppressed compared to the terms that are leading

in the considered region. In general, approximators must satisfy $T_R \Gamma_R = \Gamma_R$ (up to power corrections), where the subscript R on Γ denotes the restriction of the loop integrals to the 'design' region R. This restriction, however, complicates the analysis as the introduction of momentum cut-offs breaks Lorentz invariance. Hence, simply adding up the various contributions Γ_R to approximate the full graph is not what we are after.

The problem of restricting the loop integrals to the appropriate design regions can be overcome by instead integrating the approximated integrands over the *full* range of loop momenta. Consequently, the 'unwanted' contributions that are included in this way need to be properly subtracted. This is achieved by the Collins subtraction method [30].[5] For a given region R, contributions from so-called smaller regions need to be subtracted to avoid double counting. The hierarchy for the various possible regions is determined by the dimensions of the associated pinch surfaces: we say that R' is smaller than R (or $R' < R$) if R' is associated to a lower-dimensional pinch surface. The contribution from a particular region R to a graph Γ, including subtractions to avoid double counting with smaller regions, is denoted by $C_R \Gamma$ and given by

$$C_R \Gamma \equiv T_R \left(\Gamma - \sum_{R' < R} C_{R'} \Gamma \right). \qquad (3.23)$$

This is a recursive definition of $C_R \Gamma$, starting from the smallest leading region (note that for the smallest region $R = R_0$ one simply has $C_{R_0} \Gamma = T_{R_0} \Gamma$). We refer to the first and second terms on the right-hand side of eq. (3.23) by 'naive graph' and 'subtraction' terms, respectively. Finally, the correct leading-power approximation to the graph Γ can be obtained by summing over all leading region contributions $C_R \Gamma$:

$$\Gamma = \sum_R C_R \Gamma. \qquad (3.24)$$

We refer to [30] for a proof of this result.

Now let us apply the Collins subtraction method to the example discussed in the previous subsection. We found that the loop momentum ℓ in the single-gluon exchange graph in figure 3.4 receives leading-power contributions from two (non-trivial) momentum regions, namely from the U and C_1 regions. Since the associated pinch surfaces are, respectively, zero- and one-dimensional, we have $U < C_1$. Hence, it follows from eqs. (3.23) and (3.24) that we can approximate Γ at leading power as

$$\Gamma = C_U \Gamma + C_{C_1} \Gamma = T_U \Gamma + T_{C_1} (1 - T_U) \Gamma. \qquad (3.25)$$

The naive graph terms are given by

$$T_U \Gamma = \int \frac{d^4 \ell}{(2\pi)^4} \frac{i}{2k_2^- \ell^+ + i\epsilon} \frac{i}{2k_1^+ \ell^- + i\epsilon} \frac{i}{2(p_1^+ - k_1^+)\ell^- + i\epsilon} \frac{i}{2\ell^+ \ell^- - \ell^2 + i\epsilon}, \qquad (3.26)$$

---

5 A similar method exists in the context of SCET called zero-bin subtraction [124].

$$T_{C_1}\Gamma = \int \frac{d^4\ell}{(2\pi)^4} \frac{i}{2k_2^-\ell^+ + i\epsilon} \frac{i}{2(k_1^+ - \ell^+)(k_1^- - \ell^-) - (\mathbf{k}_1 - \boldsymbol{\ell})^2 + i\epsilon}$$
$$\times \frac{i}{2(p_1^+ - k_1^+ + \ell^+)(p_1^- - k_1^- + \ell^-) - (\mathbf{k}_1 - \boldsymbol{\ell})^2 + i\epsilon} \frac{i}{2\ell^+\ell^- - \boldsymbol{\ell}^2 + i\epsilon}. \qquad (3.27)$$

It is easy to see that in this simple example the subtraction term $T_{C_1} T_U \Gamma$ in fact equals the naive graph term $T_U \Gamma$. Hence, the graph can be approximated as $\Gamma = T_{C_1}\Gamma$; we say that the U region can be 'absorbed' into the C_1 region.

Finally, let us make an important remark about the subtraction procedure. Due to the presence of the subtraction terms in eq. (3.23) it is sufficient to consider only the design region of momentum for a particular region contribution to a given graph. We will make use of this fact in the factorization calculation in chapter 4. For example, the C_1 contribution to a graph Γ that receives leading contributions from both the S and C_1 regions is given by $C_{C_1}\Gamma = T_{C_1}(1 - T_S)\Gamma$. However, in the region S this contribution vanishes, $C_{C_1}\Gamma_S = 0$, as T_S then reduces to the unit operator. To establish factorization it is necessary to verify that the various steps described in this chapter hold for both the naive graph and subtraction terms.

3.3.3 Rapidity divergences

A complication of the momentum regions analysis is that it can come with so-called rapidity divergences, which need to be properly regulated [124]. Rapidity divergences can arise if the observable under consideration receives contributions from different momentum regions that are of the same virtuality but have distinct rapidities. For example, both the central soft and collinear regions have the same number of powers of λ in their phase space $\int d^4\ell$ (and are thus of the same virtuality), but their rapidities are parametrically different – soft modes have small rapidities as $\ell^+/\ell^- \sim 1$, while collinear modes have very large rapidities as $\ell^\pm/\ell^\mp \gg 1$. When both of these regions contribute to an observable, rapidity divergences can appear; this is illustrated below.

To illustrate how rapidity divergences arise, let us consider the following (finite) integral:

$$I \equiv \int_{Q_T}^{Q} \frac{d\ell^+}{\ell^+} = \int \frac{d\ell^+}{\ell^+} \theta\left(\ell^+ - Q_T\right) \theta\left(Q - \ell^+\right). \qquad (3.28)$$

The momentum component ℓ^+ scales differently in different momentum regions; for example, $\ell^+ \sim Q_T$ in S and $\ell^+ \sim Q$ in C_1. Hence, the integral range of I covers both the S and C_1 momentum regions. According to eqs. (3.23) and (3.24), I can be approximated at leading power as

$$I = C_S I + C_{C_1} I = T_S I + T_{C_1}\left(1 - T_S\right) I. \qquad (3.29)$$

With the naive graph and subtraction terms given by

$$T_S I = \int \frac{d\ell^+}{\ell^+} \theta\left(\ell^+ - Q_T\right), \qquad (3.30)$$

$$T_{C_1} I = \int \frac{d\ell^+}{\ell^+} \, \theta \left(\ell^+ \right) \theta \left(Q - \ell^+ \right), \tag{3.31}$$

$$T_{C_1} T_S I = \int \frac{d\ell^+}{\ell^+} \, \theta \left(\ell^+ \right), \tag{3.32}$$

it follows from eq. (3.29) that

$$I = \int_{Q_T}^{\infty} \frac{d\ell^+}{\ell^+} + \int_0^Q \frac{d\ell^+}{\ell^+} - \int_0^{\infty} \frac{d\ell^+}{\ell^+}. \tag{3.33}$$

It is clear that in this split-up into different rapidity regions each contribution diverges – the integrals are *ill-defined*. Hence, we require a regularization prescription.

The fact that our regions analysis can introduce rapidity divergences is okay as long as they are properly regulated to give well-defined integrals. Rapidity divergences are *not* regulated by dimensional regularization and thus require their own regularization prescription. Multiple rapidity regulators have been described in the literature and in the factorization study of chapter 4, we choose a regulator of the following type [125, 126]:

$$\left| \frac{\ell^+}{\nu} \right|^{-\eta}. \tag{3.34}$$

This regularization prescription can be used for all the collinear, soft, and Glauber scalings discussed in subsection 3.3.1 and could be included in the definition of the region approximator T_R. The rapidity scale ν is a quantity of mass dimension 1 analogous to the renormalization scale μ in dimensional regularization, and the rapidity regulator η is analogous to the fractional dimension ε in dimensional regularization. Rapidity divergences manifest themselves as poles in η. Note that a full graph Γ does not have rapidity divergences – after all, they are merely an artifact of the regions analysis. This means that any rapidity regulator dependence must drop out on a graph-by-graph basis upon summing over all regions according to eq. (3.24). Hence, at the end one can safely take the limit $\eta \to 0$ for each graph individually.

Let us now come back again to the example. With the regularization prescription in (3.34), the three contributions to I in eq. (3.33) are given by

$$\int_{Q_T}^{\infty} \frac{d\ell^+}{\ell^+} \; \to \; \nu^\eta \int_{Q_T}^{\infty} d\ell^+ \, \frac{|\ell^+|^{-\eta}}{\ell^+} = \frac{1}{\eta} + \ln \left(\frac{\nu}{Q_T} \right) + \mathcal{O}(\eta), \tag{3.35}$$

$$\int_0^Q \frac{d\ell^+}{\ell^+} \; \to \; \nu^\eta \int_0^Q d\ell^+ \, \frac{|\ell^+|^{-\eta}}{\ell^+} = -\frac{1}{\eta} - \ln \left(\frac{\nu}{Q} \right) + \mathcal{O}(\eta), \tag{3.36}$$

$$\int_0^{\infty} \frac{d\ell^+}{\ell^+} \; \to \; \nu^\eta \int_0^{\infty} d\ell^+ \, \frac{|\ell^+|^{-\eta}}{\ell^+} = 0. \tag{3.37}$$

Since the subtraction term is given by a so-called scaleless integral, it vanishes, see e.g. [127]. We have expanded the first two integrals in the regulator η. Indeed,

the rapidity divergences are now properly regulated, giving rise to well-defined integrals. Note that upon summing together both contributions, the poles in η precisely cancel. After performing the limit $\eta \to 0$, the sum gives the finite result $\ln(Q/Q_T)$, consistent with the original integral in eq. (3.28).

3.4 COMPLICATIONS FROM GLAUBER MODES

We have discussed how to partition the leading-power DY graphs into various subgraphs that each characterize a particular momentum scaling, see figure 3.3. To establish factorization, one needs to sum over all possible gluon attachments between those subgraphs. Before doing so, so-called Grammer-Yennie approximations [128] can be applied to soft gluons and unphysically polarized collinear gluons. These approximations are essential to give rise to the necessary Wilson line operators through non-Abelian Ward identities. Unfortunately, the Grammer-Yennie approximation is not valid for partons with Glauber scaling. In this section we review the complications that arise from Glauber modes and how they are dealt with in the CSS proof. In subsection 3.4.3 we explain that to establish factorization one is not only required to sum over all possible graphs and leading regions, but also over all possible final-state cuts.

3.4.1 *Grammer-Yennie approximations*

For gluons that connect the different subgraphs one can make certain approximations. The following discussion summarizes the reviews presented in [102, 103]. Let us first consider a right-moving collinear gluon with momentum ℓ connecting the A and H subgraphs, which is of the form $A^\mu(\ell) \, H_\mu(\ell)$, where μ is the polarization index of the gluon. For this collinear line it is sufficient to only keep the large light-cone component in H, i.e. $\ell \to \hat{\ell} \equiv (\ell^+, 0, \mathbf{0})$. Since $|A^+| \gg |A^-|, |\mathbf{A}|$ (A^μ scales as a right-moving collinear momentum [30, 129]), and all components of H are generically of the same order, we approximate $A^\mu(\ell) \, H_\mu(\ell)$ as follows:

$$A^\mu(\ell) \, H_\mu(\hat{\ell}) \approx A^+(\ell) \, \frac{\ell^+ v^-}{\ell^+ v^- + i\epsilon} \, H^-(\hat{\ell}) \approx A_\mu(\ell) \, \frac{v^\mu}{\ell \cdot v + i\epsilon} \, \hat{\ell}^\nu H_\nu(\hat{\ell}), \qquad (3.38)$$

where $v \equiv (v^+, v^-, \mathbf{0})$ is an auxiliary vector with either $|v^-| \gg |v^+|$ or $|v^-| \sim |v^+|$. Note that in the second step we added an $i\epsilon$ prescription to the denominator to regulate the pole at $\ell \cdot v = 0$ (this prescription is commented on in section 3.5). Eq. (3.38) is the Grammer-Yennie approximation appropriate for right-moving collinear gluons with unphysical polarization,[6] and the right-hand side is suitable for the use of a

---

6 In their factorization proof, CSS adopt the Feynman gauge (in part I of this thesis we do the same), and, as it turns out, in this gauge only collinear gluons with unphysical polarization are leading. In a light-cone gauge, however, also physically polarized gluons can contribute at leading power [130], giving rise to transverse gauge links at minus light-cone infinity [70, 71].

Ward identity. Naturally, analogous approximations can be made for left-moving collinear gluons connecting the B and H subgraphs. Upon summing over all collinear gluons, the collinear subgraphs are stripped from the hard subgraph, and the former end up with past-pointing Wilson line operators as appropriate for DY scattering (see e.g. [102] and section 4.4 for more details).

One can also apply approximations to gluons that connect the soft and collinear subgraphs. Let us consider a soft gluon with momentum ℓ that connects the S and A subgraphs, which is of the form $S^\mu(\ell) A_\mu(\ell)$. We can neglect the plus component of ℓ compared to the large plus components in A, i.e. $\ell \to \tilde{\ell} \equiv (0, \ell^-, \boldsymbol{\ell})$. In this case the Grammer-Yennie approximation is given by

$$S^\mu(\ell) A_\mu(\tilde{\ell}) \approx S^-(\ell) \frac{\ell^- w^+}{\ell^- w^+ + i\epsilon} A^+(\tilde{\ell}) \approx S_\mu(\ell) \frac{w^\mu}{\ell \cdot w + i\epsilon} \tilde{\ell}^\nu A_\nu(\tilde{\ell}), \qquad (3.39)$$

where $w \equiv (w^+, w^-, \boldsymbol{0})$ is an auxiliary vector with $|w^+| \gg |w^-|$. Again, we have provided an $i\epsilon$ prescription to the denominator in the second step (see section 3.5). The right-hand side of eq. (3.39) is appropriate for the use of a Ward identity. Analogous approximations can be applied to soft gluons that connect the S and B subgraphs. Upon summing over all soft gluon attachments, the soft subgraph is stripped from the collinear subgraphs and is given by the vacuum expectation value of a product of Wilson lines (see e.g. [30, 102] for more details).

If this were the full story then one would have now established factorization of the DY cross section into soft, collinear, and hard subgraphs. However, there is a complication that arises from Glauber modes that are initially contained in the soft subgraph. The Grammer-Yennie approximation for soft gluons, eq. (3.39), does *not* hold for Glauber modes. If $\ell^- \sim \lambda^2 Q$ and $|\boldsymbol{\ell}| \sim \lambda Q$, then the approximation $\ell^- A^+ \approx \tilde{\ell}^\nu A_\nu$ fails as the transverse contribution to $\tilde{\ell}^\nu A_\nu$ cannot be neglected. A similar complication arises when considering Glauber attachments to B.

These problems can be overcome if the integrations over Glauber momenta could be deformed into a region of the complex plane (soft or collinear) where the Grammer-Yennie approximation is valid. Contour deformations are only possible if (some of) the integrations over the Glauber momentum components are *not* pinched. If no deformations into a different momentum region are possible, we say that the integration is 'trapped' in the Glauber region. In the next subsection we consider the pole structure associated to Glauber exchanges and provide examples of both a trapped and an untrapped Glauber gluon.

3.4.2 *Trapped in the Glauber region?*

Let us consider again the single-gluon exchange diagram given on the right in figure 3.4. In subsection 3.3.1 we showed with a power counting analysis that the C_1, U, G_1, and G regions contribute at leading power. Here we show explicitly that the loop integration contours are not trapped in the two Glauber regions G_1 and G,

implying that they could be deformed into different momentum regions. To this end, we investigate the poles of this diagram in the light-cone components of ℓ.

At leading power in the G region, the active-spectator gluon exchange graph has the following structure (we only consider the relevant pieces here, i.e. the factors that depend on ℓ^+ and ℓ^-):

$$
\int \frac{d\ell^+ d\ell^-}{(2\pi)^2} \frac{i}{(k_1 - \ell)^2 + i\epsilon} \frac{i}{(p_1 - k_1 + \ell)^2 + i\epsilon} \frac{i}{(k_2 + \ell)^2 + i\epsilon}
$$

$$
= \left[\int \frac{d\ell^-}{2\pi} \frac{i}{2k_1^+(k_1^- - \ell^-) - (k_1 - \ell)^2 + i\epsilon} \right.
$$

$$
\left. \times \frac{i}{2(p_1^+ - k_1^+)(p_1^- - k_1^- + \ell^-) - (k_1 - \ell)^2 + i\epsilon} \right]
$$

$$
\times \left[\int \frac{d\ell^+}{2\pi} \frac{i}{2k_2^-(k_2^+ + \ell^+) - (k_2 + \ell)^2 + i\epsilon} \right]. \tag{3.40}
$$

For the G_1 region the result is the same, except that on the last line the k_2^+ component can in that case be neglected compared to ℓ^+. From the structure of the propagators it follows that for both Glauber regions there are two poles in ℓ^- (labeled by a, b) and there is one pole in ℓ^+:

$$
\ell_a^\pm = \mathcal{O}(\lambda^2) - i\epsilon, \qquad \ell_b^- = \mathcal{O}(\lambda^2) + i\epsilon. \tag{3.41}
$$

Whereas the ℓ^- contour is pinched at the origin between two poles, the ℓ^+ contour is not pinched, see also figure 3.5. The pole ℓ_a^+ is also referred to as an 'initial-state' pole as it is consistent with the ultimate formation of an initial-state Wilson line (a pole on the opposite side of the real axis is also referred to as a 'final-state' pole). Since the ℓ^+ component only has a pole in the lower half plane, we can freely deform the ℓ^+ contour into the upper half plane according to Cauchy's theorem. In fact, if we deform it into a semicircle around the origin with radius of order Q, we enter the C_1 momentum region for which $\ell \sim (1, \lambda^2, \lambda)Q$. Hence, the two Glauber modes can be absorbed into a collinear region contribution where the appropriate Grammer-Yennie approximation can be applied. Note that since the ℓ^- contour is pinched, a deformation into the central soft region is not possible in this case.

For the next example, let us consider the spectator-spectator gluon exchange diagram in figure 3.6 (we employ the same model). Using the power counting techniques from subsection 3.3.1, it is straightforward to show that this graph receives a leading-power contribution from the central Glauber region G; there are three additional propagators with respect to the tree-level graph and combined they give rise to a $1/\lambda^6$ enhancement, and, at the same time, the Glauber phase space $\int d^4\ell$ gives a λ^6 suppression. Hence, the overall scaling of the graph is λ^0, implying that the region G contributes at leading power.

Figure 3.5: The pole structure for the light-cone components of the gluon momentum ℓ appearing in the active-spectator gluon exchange graph on the right in figure 3.4. The integration contour for the minus component is pinched between two poles (indicated by red crosses), whereas the one for the plus component is not. The latter contour can be deformed in the upper half plane into a semicircle around the origin with radius of order Q (indicated by the dashed line).
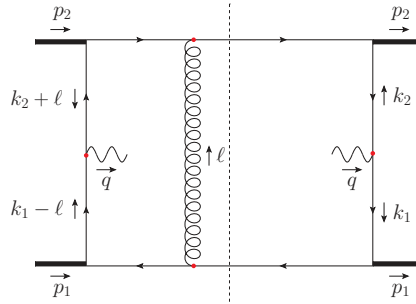
Figure 3.6: A spectator-spectator gluon exchange diagram that receives a leading contribution when the gluon momentum ℓ has (central) Glauber scaling. The integration over this loop momentum is trapped in the Glauber region.

At leading power in the G region, the spectator-spectator exchange graph has the following structure (we again consider the relevant pieces only):

$$\int \frac{d\ell^+ d\ell^-}{(2\pi)^2} \frac{i}{(k_1 - \ell)^2 + i\epsilon} \frac{i}{(p_1 - k_1 + \ell)^2 + i\epsilon} \frac{i}{(k_2 + \ell)^2 + i\epsilon} \frac{i}{(p_2 - k_2 - \ell)^2 + i\epsilon}$$

$$= \left[\int \frac{d\ell^-}{2\pi} \frac{i}{2k_1^+ (k_1^- - \ell^-) - (\boldsymbol{k}_1 - \boldsymbol{\ell})^2 + i\epsilon} \right.$$

$$\times \left. \frac{i}{2(p_1^+ - k_1^+)(p_1^- - k_1^- + \ell^-) - (\boldsymbol{k}_1 - \boldsymbol{\ell})^2 + i\epsilon} \right]$$

$$\times \left[\int \frac{d\ell^+}{2\pi} \frac{i}{2k_2^- (k_2^+ + \ell^+) - (\boldsymbol{k}_2 + \boldsymbol{\ell})^2 + i\epsilon} \right.$$

$$\times \left. \frac{i}{2(p_2^- - k_2^-)(p_2^+ - k_2^+ - \ell^+) - (\boldsymbol{k}_2 + \boldsymbol{\ell})^2 + i\epsilon} \right]. \tag{3.42}$$

From the structure of the propagators we infer that there are two poles for each light-cone component of ℓ, namely

$$\ell_a^\pm = \mathcal{O}(\lambda^2) - i\epsilon, \qquad \ell_b^\pm = \mathcal{O}(\lambda^2) + i\epsilon. \tag{3.43}$$

This means that both the ℓ^+ and ℓ^- contours are pinched at the origin between two poles situated on either side of the real axis, see also figure 3.7. Hence, neither contour can be deformed freely into the complex plane – the loop integration is *trapped* in the Glauber region.
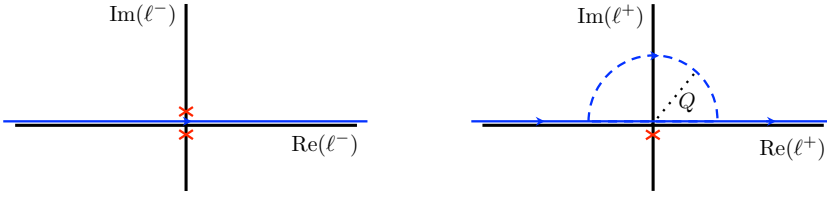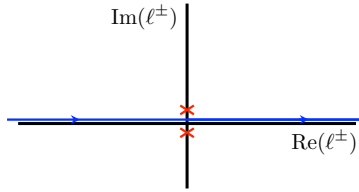

Figure 3.7: The pole structure for the light-cone components of the gluon momentum ℓ appearing in the spectator-spectator gluon exchange graph in figure 3.6. For each component, the integration contour is pinched between two poles at the origin.

With these two examples we have demonstrated that in DY both trapped and untrapped Glauber exchanges occur. The former situation generally results from spectator-spectator interactions [30] and seems problematic. Since the Grammer-Yennie approximations are key to establish factorization through Ward identities, we need to be able to deform *all* Glauber modes into soft or collinear regions where those approximations are valid. In the next subsection we will explain how the CSS approach takes care of this issue.

3.4.3 *Unitarity cancellation*

The Glauber modes prevent one from factorizing the soft, collinear, and hard sub-graphs upon summing over all soft and collinear attachments. The problem lies in the fact that in DY graphs both initial- and final-state poles appear, which is *not* the case in for example e^+e^- annihilation into two hadrons and in SIDIS [30, 113]. This extra difficulty in DY was overcome by CSS by summing over all kinematically allowed final-state cuts – a sum over different final states is indeed appropriate for *inclusive* DY scattering. Using light-cone perturbation theory, CSS proved to all orders that in fact all final-state poles cancel in the sum over cuts [30, 97, 98]. The sum over cuts is performed on a graph-by-graph basis and the pole cancellation is connected to unitarity of the theory through the Cutkosky cutting rules. Hence, the physical reason underlying this cancellation is conservation of probability. The unitarity cancellation implies that all pinches are lifted and that the Glauber modes become untrapped, allowing for contour deformations into soft or collinear regions

where the Grammer-Yennie approximations are valid. As the details of the proof are not relevant for our factorization analysis in chapter 4, we will simply clarify the sum-over-cuts argument by providing an explicit example.

Let us consider again the single-gluon exchange graph in figure 3.6. In the previous subsection we showed that for this spectator-spectator exchange the loop integration is trapped in the Glauber region. However, according to CSS, we need to sum over all unobserved final states for any given graph and leading region. For this particular graph, in the Glauber region, there are two possible final-state cuts, namely one directly to the left and one directly to the right of the exchanged gluon. Note that we should disregard a cut through the Glauber gluon itself, as a Glauber mode is an *off-shell* momentum mode.
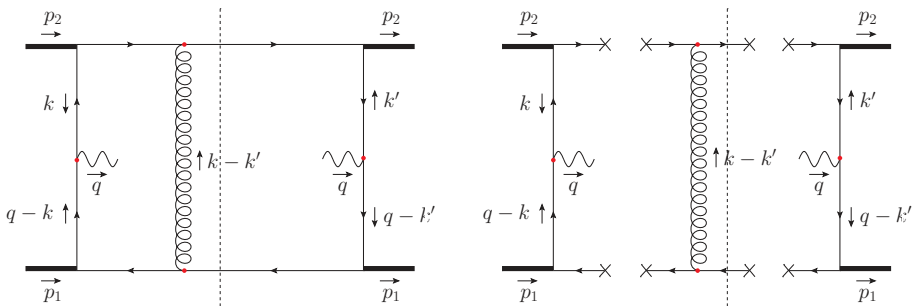
Figure 3.8: The single-Glauber exchange graph; only one of the two cut positions is shown. At leading power, the graph on the left can be replaced by the one on the right (the lines with crosses denote on-shell lines).

The following discussion is based on reviews in [30, 104] and for convenience we consider the same momentum routings. The authors showed that for each cut the spectator propagators that are not already on shell due to the cut can be replaced by on-shell delta functions. Pictorially this means that the graph on the left in figure 3.8 can be replaced by the one on the right, and a similar replacement holds for the other cut position. To motivate this replacement, we study the leading-power structure of the diagram on the left in terms of the light-cone components of k. Note that in the diagram the momentum k only appears to the left of the cut. The lines k and $q - k$ are left- and right-moving collinear momenta respectively, whereas q is hard. This means that k^+ and $q^- - k^-$ are negligible compared to the large components q^+ and q^- respectively. The precise numerator factors are not relevant in this discussion and will therefore be ignored. Employing Cauchy's residue theorem, we can write

$$\int \frac{dk^+ dk^-}{(2\pi)^2} \frac{i}{(q-k)^2 + i\epsilon} \frac{i}{(p_1 - q + k)^2 + i\epsilon} \frac{i}{k^2 + i\epsilon} \frac{i}{(p_2 - k)^2 + i\epsilon}$$

$$= \left[\int \frac{dk^-}{2\pi} \frac{i}{2q^+(q^- - k^-) - (\boldsymbol{q} - \boldsymbol{k})^2 + i\epsilon} \right.$$

$$\times \frac{i}{2(p_1^+ - q^+)(p_1^- - q^- + k^-) - (\boldsymbol{q} - \boldsymbol{k})^2 + i\epsilon} \Bigg]$$

$$\times \left[\int \frac{dk^+}{2\pi} \frac{i}{2q^- k^+ - \boldsymbol{k}^2 + i\epsilon} \frac{i}{2(p_2^- - q^-)(p_2^+ - k^+) - \boldsymbol{k}^2 + i\epsilon} \right]$$

$$= \left[\frac{1}{2(p_1^+ - q^+)} \frac{i}{2q^+(q^- - k_a^-) - (\boldsymbol{q} - \boldsymbol{k})^2 + i\epsilon} \right] \left[\frac{1}{2(p_2^- - q^-)} \frac{i}{2q^- k_a^+ - \boldsymbol{k}^2 + i\epsilon} \right]$$

$$= \left[\int \frac{dk^-}{2\pi} \frac{i}{2q^+(q^- - k^-) - (\boldsymbol{q} - \boldsymbol{k})^2 + i\epsilon} \right.$$

$$\left. \times 2\pi\delta \left(2(p_1^+ - q^+)(p_1^- - q^- + k^-) - (\boldsymbol{q} - \boldsymbol{k})^2 \right) \right]$$
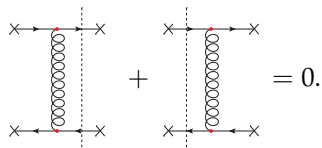
$$\times \left[\int \frac{dk^+}{2\pi} \frac{i}{2q^- k^+ - \boldsymbol{k}^2 + i\epsilon} 2\pi\delta \left(2(p_2^- - q^-)(p_2^+ - k^+) - \boldsymbol{k}^2 \right) \right]$$

$$= \int \frac{dk^+ dk^-}{(2\pi)^2} \frac{i}{(q - k)^2 + i\epsilon} 2\pi\delta \left((p_1 - q + k)^2 \right) \frac{i}{k^2 + i\epsilon} 2\pi\delta \left((p_2 - k)^2 \right), \quad (3.44)$$

where the poles $k_a^{\pm}$ appearing in the third step are given by

$$k_a^- = q^- - p_1^- + \frac{(\boldsymbol{q} - \boldsymbol{k})^2}{2(p_1^+ - q^+)} - i\epsilon, \qquad k_a^+ = p_2^+ - \frac{\boldsymbol{k}^2}{2(p_2^- - q^-)} + i\epsilon. \qquad (3.45)$$

Based on eq. (3.44) we conclude that the effect of the k^- and k^+ integrations is to set the lines $p_1 - q + k$ and $p_2 - k$ on shell, respectively. Hence, we can replace the left graph in figure 3.8 by the one on the right. In the same way one can show that for the other cut position (i.e. the one directly to the left of the Glauber gluon), the effect of the k'^- and k'^+ integrations is to set the lines $p_1 - q + k'$ and $p_2 - k'$ on shell, respectively.

For both cut positions the left- and rightmost factors are the same – together those factors constitute the tree-level DY graph (albeit with slightly different momenta to the left and right of the cut). Hence, upon summing over the two cuts, we have a tree-level contribution multiplied by a sum over the two external cuts for the t-channel gluon exchange subgraph. The latter sum vanishes due to the Cutkosky rules [131]:

$$= 0. \qquad (3.46)$$

In the following we elaborate on this result and argue that it follows from unitarity of the theory, which in turn relies on conservation of probability.

The probability P that a (normalized) initial state $|i\rangle$ transforms into some (normalized) final state $|f\rangle$ is given by

$$P(i \to f) = |\langle f| S |i\rangle|^2,$$ (3.47)

where S is the scattering operator. Probability conservation dictates that for a given initial state $|i\rangle$, the sum over all possible final states $|f\rangle$ gives unity. Making use of eq. (3.47), this condition translates into

$$1 = \sum_f P(i \to f) = \langle i| S^\dagger S |i\rangle.$$ (3.48)

Since the expectation value on the right-hand side must hold for any $|i\rangle$, it follows that $S^\dagger S = 1$. Similarly, any given final state must arise (i.e. with unit probability) from some initial state; this leads to the condition $SS^\dagger = 1$. Hence, S is a unitary operator. Conventionally, we decompose S into the identity operator I and the Hermitian transition operator T as $S = I + iT$. From the unitarity of S, it follows that

$$T - T^\dagger = iT^\dagger T.$$ (3.49)

The transition operator T, sandwiched between the states $\langle f|$ and $|i\rangle$, is related to the scattering amplitude $\mathcal{M}$:

$$\langle f| T |i\rangle = (2\pi)^4 \delta^{(4)}(p_i - p_f) i\mathcal{M}(i \to f).$$ (3.50)

From eqs. (3.49) and (3.50) we obtain the so-called generalized optical theorem:

$$i[\mathcal{M}(i \to f) - \mathcal{M}^*(f \to i)] = -\sum_X \int d\Phi_X \, \mathcal{M}(i \to X)\mathcal{M}^*(f \to X),$$ (3.51)

where X is any possible intermediate state and $d\Phi_X$ denotes the corresponding phase space element. The Cutkosky rules are a generalization of eq. (3.51) – they hold for *individual* Feynman diagrams with the sum over X on the right-hand side replaced by a sum over internal cuts [131, 132] (see also e.g. [104, 133]).

The left-hand side of eq. (3.46) is then precisely described by the left-hand side of eq. (3.51) for a given graph. After all, the Glauber exchange subgraph with the external cut to the right is just the graph itself, $i\mathcal{M}(i \to f)$, while the one with the external cut to the left is given by the complex conjugate of the time-reversed process, $-i\mathcal{M}^*(f \to i)$. As there are no possible internal cuts for our Glauber exchange subgraph, the sum over external cuts evaluates to zero according to the Cutkosky rules. Hence, for the spectator-spectator gluon exchange diagram of this example, the contribution from the Glauber region completely cancels out.[7] The general principle illustrated here also applies to more complicated diagrams featuring multiple Glauber exchanges, see e.g. [104].

7 Strictly speaking, as the Glauber subgraphs are actually imaginary [134] (and the tree-level skeleton is real), this particular spectator-spectator exchange graph must always cancel to end up with a real cross section, regardless of the Cutkosky rules.

In the previous section we discussed the complications that arise from Glauber modes regarding the factorization procedure, and how they can be overcome. CSS proved to all orders in perturbation theory that upon summing over all final-state cuts, the final-state poles cancel out due to unitarity of the theory. In some cases, though, this cancellation goes even further – in the previous section, namely, we showed that for e.g. the single-gluon exchange diagram in figure 3.8 the Glauber contribution completely vanishes. For the graphs for which this is not the case, the unitarity cancellation of final-state poles allows for contour deformations out of the Glauber region, away from initial-state poles, into soft or collinear regions where the Grammer-Yennie approximations can be applied. After application of those approximations, the contours can be deformed back to the real axes again. The $i\epsilon$ prescriptions in eqs. (3.38) and (3.39) are chosen such that this is always possible and are consistent with the formation of initial-state Wilson lines, as appropriate for DY.

Effectively what happens is that part of the Glauber contributions is cancelled while the other part can be absorbed in the soft or collinear subgraphs. Upon summing over all collinear and soft gluon attachments, Ward identities ensure that the various subgraphs (S, A, B, and H) factorize – this is ultimately linked to gauge invariance of the theory. The result of the factorization procedure is shown schematically in figure 3.9. The collinear subgraphs A and B each contain a past-pointing Wilson line, and the soft subgraph S is given by the vacuum expectation value of a product of Wilson lines.
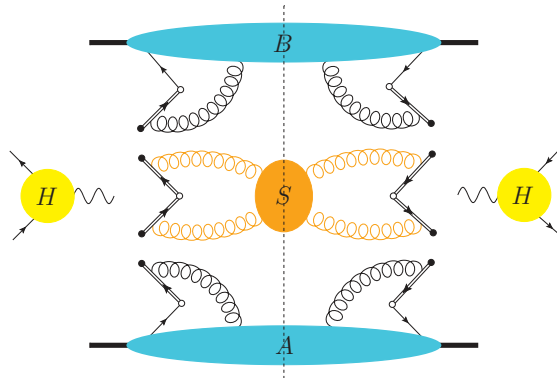
Figure 3.9: Factorized soft, collinear, and hard subgraphs.

In contrast to H, each of S, A, and B contains nonperturbative contributions that are not predicted by the theory; those contributions need to be fitted to data. However, since S appears inside a product with two collinear factors, it could never

be measured independently.[8] Hence, it was proposed in [30] to combine each of the subgraphs *A* and *B* with a square root of *S*. Today, those combinations are considered the proper definitions of TMD parton densities. Another advantage of this definition is that, as it turns out, rapidity divergences precisely cancel in the TMDs, while *S*, *A*, and *B* separately are divergent. The cancellation of rapidity divergences was recently proven to hold to all orders in perturbation theory [136]. Thus, the result of the CSS procedure is a factorized cross section with two TMDs and a hard function, see figure 3.10.
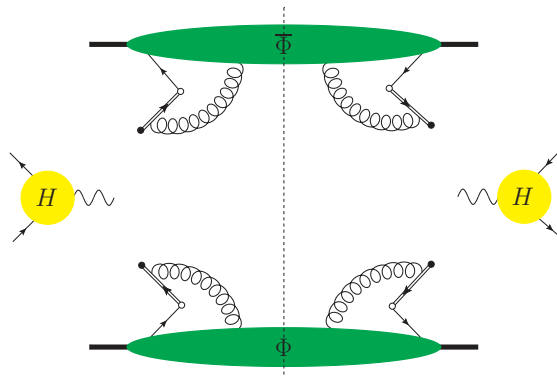
Figure 3.10: Final factorized form of the DY process. The soft factor has been absorbed in the quark and antiquark TMD correlators.

3.6 POSSIBLE ISSUES

In this chapter we have discussed various key elements of the CSS factorization proof for DY scattering. We have argued that upon summing over all graphs, leading momentum regions, and final-state cuts, the DY cross section factorizes into two TMDs and a hard function, see eq. (3.1) and figure 3.10. As discussed in subsection 2.3.3, the necessary gauge links in the TMDs are not unique, but depend on the process under consideration. The gauge links are understood to track the color flow in the hard process, which is e.g. entirely incoming in DY and entirely outgoing in SIDIS. Consequently, in DY the two TMDs come with past-pointing gauge links, whereas in SIDIS the links are future-pointing. Whenever the color flow is both incoming and outgoing, such as in back-to-back hadron pair production in *pp* collisions, it has been argued that the gauge links can no longer be disentangled, preventing factorization in terms of separately color gauge invariant

---

8 Although the soft factor is universal [135], it could not be independently extracted from other processes either (for which TMD factorization has been proven), as it is in fact always accompanied by two collinear factors.

factors containing TMDs [137–139]. This entanglement of color structures hampers the theoretical prediction of the associated observables.

In [101] it was suggested that color-entangled contributions can also arise in DY scattering. Those contributions, at best, would come in a factorized form with a color factor different from $1/N_c$ (as predicted by the DY factorization theorem). The affected contributions involve only T-odd TMDs, such as the BM or Sivers functions. As explained in subsection 2.3.3, those functions are special in the sense that their existence completely relies on the presence of gauge links, which causes them to change sign between DY and SIDIS. In [101], however, it was claimed that the process dependence of double T-odd contributions goes beyond this sign flip: as a result of color entanglement, those contributions are suppressed and change sign in DY due to an additional color factor of $-1/(N_c^2 - 1)$. This color-entanglement effect would signal a loophole in the TMD factorization proof for double T-odd contributions that involve polarization. In [101] factorization was neither proven nor disproven. Since it is unclear how color-entanglement effects would manifest themselves beyond the lowest order considered, it means that factorization could potentially be violated. At lowest order in α_s, an entangled color structure contributing to the dBM term for example arises in the graph in figure 3.11 with a single-gluon exchange between each correlator and the active parton coming from the other side.
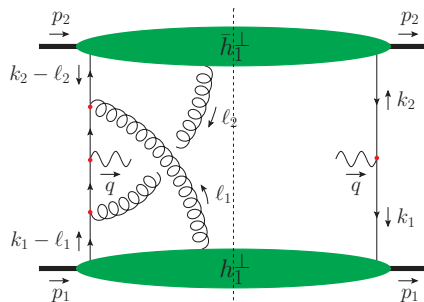
Figure 3.11: An example of a lowest-order graph with a color-entangled structure that gives a nonzero contribution to the dBM term in the DY cross section.

The question that is addressed in detail in the next chapter is whether the color-entanglement effect described in [101] survives after summing over all relevant graphs, leading momentum regions, and final-state cuts, as required according to CSS to obtain factorization. To this end, we perform an explicit calculation of the dBM contribution to DY scattering using a spectator model. The calculation is performed up to the first order at which color entanglement is supposed to appear according to [101], i.e. up to $\mathcal{O}(\alpha_s^2)$. Note that the $\mathcal{O}(\alpha_s^2)$ level corresponds to two gluon exchanges and thus includes for example the diagram in figure 3.11.

4

DISENTANGLING COLORS FOR AZIMUTHAL ASYMMETRIES

In this chapter we will investigate whether double T-odd contributions to the DY cross section factorize in the way predicted by CSS, or instead suffer from color-entanglement effects as described in [101]. As explained in section 3.6, those effects would suppress the double T-odd contributions by an additional color factor and could even lead to a breakdown of factorization. Since the affected contributions are measured in experiments, the factorization property of the cross section is highly essential. According to [101], the different color structures should be visible already at the lowest nontrivial order, i.e. at the level of two gluon exchanges. At this level, we perform an explicit model calculation of the dBM contribution to the DY cross section to investigate its color structure (in section 4.6 we also briefly discuss the results for the Sivers function). The computation is based on our work in [2] and proceeds largely along the lines of the CSS factorization proof as outlined in the previous chapter. The main differences with the CSS approach are that we will treat the Glauber region distinctly from the soft region in the Collins subtraction procedure (i.e. with its own approximator) and that we will not perform any contour deformations.

4.1 STATUS OF THE BOER-MULDERS EFFECT

The BM function $h_1^\perp$ encodes correlations between the transverse polarization and the transverse momentum of a quark inside an unpolarized hadron. As can be seen from eq. (3.1), this TMD contributes to a $\cos(2\phi)$ azimuthal asymmetry in the DY cross section. Experimentally, the dBM term can be extracted as follows:

$$\mathcal{F}\left[w(\boldsymbol{k}_1, \boldsymbol{k}_2)\, h_1^\perp \bar{h}_1^\perp \right] \propto \int_0^{2\pi} d\phi\, \cos(2\phi) \left(\frac{d\sigma}{d\Omega\, dx_1 dx_2\, d^2\boldsymbol{q}} \right). \tag{4.1}$$

Thus, by measuring the $\cos(2\phi)$ angular dependence, DY provides an opportunity to probe the transverse polarization of quarks inside unpolarized hadrons.

There is yet another process that can be used to extract the BM function, namely SIDIS. In the unpolarized SIDIS process, the BM function appears in a convolution with the chiral-odd Collins FF which describes the fragmentation of a transversely polarized quark into an unpolarized hadron [140]. This contribution gives rise to a $\cos(2\phi)$ angular dependence as well, where now ϕ is the azimuthal angle of the produced hadron with respect to the lepton plane. In both DY and SIDIS, the $\cos(2\phi)$ asymmetries are referred to as BM effects as they can be ascribed to a nonzero BM function.

The fact that intrinsic partonic transverse momenta can give rise to these asymmetries, even at leading twist, was noted in the 1990s by Boer in [116] for DY and by Boer and Mulders in [63] for SIDIS. Those observations sparked a lot of interest in the field of spin physics. Besides lattice QCD calculations of the BM function [141, 142], there have been many studies of the BM function in the context of QCD-inspired models, see e.g. [117, 143–154]. Since the BM function cannot be calculated using perturbative QCD methods, model calculations are used to obtain more insight into its nonperturbative properties, such as its size and sign. Moreover, they can be instructive as to what functional form should be used for the experimental extractions.

Perhaps the most studied model is the so-called spectator model in which the hadron is composed of an active quark that participates in the hard scattering, as well as a spectator that groups together the remaining partonic degrees of freedom. In the case of a proton, the spectator is also referred to as a diquark as it serves to represent a combination of two valence quarks. Naturally, the diquark can be of spin 0 (a scalar) or spin 1 (an axial vector) – both cases are studied in the literature, with the former being the simplest. The factorization calculation in this chapter is performed using the scalar spectator model (see the next section for more details).

In unpolarized hadron-hadron collisions, the $\cos(2\phi)$ asymmetry has been measured by the NA10 experiment at CERN [155, 156] and the E615 [157] and E866/NuSea experiments at Fermilab [158, 159], and is currently under active investigation by the COMPASS experiment at CERN [89] and the E906/SeaQuest experiment at Fermilab [160, 161]. Using the available data on pd and pp collisions in [158, 159], both the quark [162, 163] and antiquark [164] BM functions for the proton have been extracted. For the fits, the x- and k-dependent parts were assumed to factorize, and for the latter a Gaussian parametrization was considered. The results in [163] agree with the positive signs for both the up and down quark BM functions as predicted by almost all model calculations. Furthermore, they satisfy the appropriate positivity bounds given in [68]. However, as was recently argued in [165], for proper extractions of BM functions also next-to-leading order perturbative QCD corrections should be incorporated.

For unpolarized SIDIS, two different extractions of the proton BM function have been performed [166, 167]. The latest extraction [167] is based on charged hadron production through $e^{\pm}p$ and $e^{\pm}d$ collisions measured by the HERMES experiment at DESY [168, 169], as well as on charged hadron production through $\mu^{\pm}d$ collisions measured by the COMPASS experiment at CERN [170, 171]. Similar to the extraction from DY, the x- and k-dependent parts of the BM function were assumed to factorize, with a Gaussian profile for the latter. As it turns out, data at higher values of Q^2 is needed to properly separate the BM effect from higher-twist contributions [166].

Unfortunately, more data is needed to claim the observation of the expected sign flip of the BM function between DY and SIDIS. Many precision measurements of the $\cos(2\phi)$ asymmetries are planned at existing or future experimental facilities (see

e.g. the reviews [59, 172]), enabling more precise extractions of the BM functions in both processes.

Since the color-entanglement result for DY as described in [101] is at odds with the CSS factorization theorem, it is important to check this result in an explicit calculation. After all, the various ongoing and planned experimental extractions of the BM function crucially rely on factorization. To this end, we perform in this chapter an explicit model calculation of the dBM term in the DY cross section, following closely the steps of the CSS factorization proof as reviewed in the previous chapter. In the next section we discuss the details of the spectator model that we adopt for the calculation.

4.2 MODEL SETUP

For the factorization study in this chapter we employ a spectator model in which the colorless spin-$\frac{1}{2}$ proton couples to a spin-$\frac{1}{2}$ quark and a scalar spectator (see e.g. [75, 148, 173–175]). The quark is in the color triplet representation with electrical charge $e_q = 1$, and the scalar is in the color antitriplet representation and is electrically neutral. We take the proton-quark-scalar coupling to be a constant for simplicity (as one would obtain for a fundamental Yukawa-type fermion-fermion-scalar coupling); for convenience this vertex factor will be set to unity in the calculations. The proton and scalar are taken massive with masses M and m_s respectively, and the quark is taken massless.[1] The antiproton is treated using the same spectator model as the proton, albeit with quantum numbers appropriately conjugated.

In this chapter we adopt the momenta conventions as specified in sections 3.2 and 3.3. This means that in the cross section calculation we consider a proton colliding with an antiproton (in the CM frame of the collision), with right- and left-moving collinear momenta p_1 and p_2 respectively, see (3.17). In this collision, we consider the DY production of an off-shell photon with a hard momentum q, which occurs via $q\bar{q}$ fusion. The scalars coupling to either hadron are spectators. To enable the hard scattering, the active quarks must carry right- and left-moving collinear momenta k_1 and k_2 respectively, see (3.18).

In the model calculation we consider QCD corrections to tree-level DY production. The coupling of gluons to quarks, antiquarks, and scalars proceeds via the standard (fermionic or scalar) QCD Feynman rules, see appendix C. By using the standard couplings, we ensure that the model obeys important physical principles such as unitarity.[2] The unitarity property of the model will play an important role in the calculations of this chapter. In this part of the thesis we adopt the Feynman gauge.

Since we are interested in the dBM contribution to the cross section, both the quark and antiquark need to be transversely polarized, see eq. (3.6). We will for

---

[1] To avoid issues related to proton decay, we take $m_s > M$ in our calculations.
[2] In general, vertices respect unitarity if the corresponding interaction Lagrangian is Hermitian (and this is indeed the case for our model).

each diagram encounter a Dirac trace of the form $\text{Tr}(\Phi H_1 \overline{\Phi} H_2)$, where H_1 and H_2 represent hard scattering matrices, and Φ and $\overline{\Phi}$ are matrices for the proton and antiproton pieces, respectively. We can decompose Φ on a basis of Dirac structures as follows [102,176]:

$$
\begin{aligned}
\Phi_{\alpha\beta} = {} & \text{Tr}\left(\tfrac{1}{2}\Phi\right)\tfrac{1}{2}\delta_{\alpha\beta} + \text{Tr}\left(\tfrac{1}{2}\gamma^5\Phi\right)\tfrac{1}{2}(\gamma^5)_{\alpha\beta} + \text{Tr}\left(\tfrac{1}{2}\gamma_\mu\Phi\right)\tfrac{1}{2}(\gamma^\mu)_{\alpha\beta} \\
& + \text{Tr}\left(\tfrac{1}{2}\gamma^5\gamma_\mu\Phi\right)\tfrac{1}{2}(\gamma^\mu\gamma^5)_{\alpha\beta} + \text{Tr}\left(\tfrac{1}{2}i\sigma_{\nu\mu}\gamma^5\Phi\right)\tfrac{1}{4}i(\sigma^{\mu\nu}\gamma^5)_{\alpha\beta}.
\end{aligned}
\tag{4.2}
$$

At leading power, the trace with the hard scattering matrix H_i is given by

$$
\begin{aligned}
\text{Tr}\left(\Phi H_i\right) = {} & \text{Tr}\left(\tfrac{1}{2}\gamma^+\Phi\right)\text{Tr}\left(\tfrac{1}{2}\gamma^- H_i\right) + \text{Tr}\left(\tfrac{1}{2}\gamma^+\gamma^5\Phi\right)\text{Tr}\left(\tfrac{1}{2}\gamma^5\gamma^- H_i\right) \\
& + \text{Tr}\left(\tfrac{1}{2}i\sigma^{j+}\gamma^5\Phi\right)\text{Tr}\left(\tfrac{1}{2}i\sigma^{j-}\gamma^5 H_i\right).
\end{aligned}
\tag{4.3}
$$

Doing the same for $\overline{\Phi}$, we can decompose the Dirac trace $\text{Tr}(\Phi H_1 \overline{\Phi} H_2)$ as follows:

$$
\begin{aligned}
\text{Tr}\left(\Phi H_1 \overline{\Phi} H_2\right) = {} & \text{Tr}\left(\Gamma_U \Phi\right)\text{Tr}\left(\overline{\Gamma}_U \overline{\Phi}\right)\text{Tr}\left(\overline{\Gamma}_U H_1 \Gamma_U H_2\right) \\
& + \text{Tr}\left(\Gamma_T^j \Phi\right)\text{Tr}\left(\overline{\Gamma}_T^k \overline{\Phi}\right)\text{Tr}\left(\overline{\Gamma}_{Tj} H_1 \Gamma_{Tk} H_2\right) + \dots,
\end{aligned}
\tag{4.4}
$$

where the Dirac projectors Γ_U and Γ_T^j select unpolarized and transversely polarized quarks respectively, see eq. (3.7) (a bar on the projector indicates that plus components are replaced by minus components). The ellipsis in eq. (4.4) represents the double longitudinally polarized contribution as well as mixed terms. For the calculation of the dBM contribution, we only need to consider the second term in this sum.

4.3 GRAPHS AND MOMENTUM REGIONS

In this section we discuss which graphs and momentum regions need to be considered for the calculation of the dBM contribution to the DY cross section. We are interested in its precise color structure to check the color-entanglement result of [101]. We consider the employed spectator model rich enough in structure to settle the issue – in particular, the color factors involved are the same as those appearing in a full QCD calculation. The study is performed at leading power in Λ/Q and at the $\mathcal{O}(\alpha_s^2)$ level, being the lowest order at which color-entanglement effects are supposed to appear according to [101].

4.3.1 *Relevant graphs*

As will be argued in section 4.4 based on the factorization theorem, at least one gluon attachment to the spectator is needed for a nonzero BM function. At $\mathcal{O}(\alpha_s)$, we find that the only graph contributing to a nonzero dBM effect is the one that features a

(central) Glauber gluon exchange between the two scalar spectators. However, as we illustrated in subsection 3.4.3, this graph vanishes upon summing over the two possible final-state cuts. Hence, there are no graphs expected to contribute at the $\mathcal{O}(\alpha_s)$ level.

The lowest order at which graphs contribute to a nonzero dBM term is the $\mathcal{O}(\alpha_s^2)$ level, which corresponds to the exchange of two gluons. In our analysis, we distinguish between graphs that have a 'color-entangled' structure and those that do not. By definition, the latter type of graphs, the 'unentangled' ones, come with the color factor

$$\mathrm{Tr}(t^a t^a t^b t^b) = \frac{1}{N_c}\,C_\Phi^2\,, \tag{4.5}$$

where the two factors of $C_\Phi \equiv \mathrm{Tr}(t^a t^a)$ are absorbed in the BM functions (see section 4.4), and the $1/N_c$ color factor is precisely the overall color factor that is expected from the CSS factorization theorem; it also appears in the leading-order result in eq. (3.1). The deviating color factor that was found in [101] for double T-odd contributions originates from diagrams that have a color structure different from that in eq. (4.5).

We find by explicit calculation that the (relevant) color-entangled graphs to consider are diagrams (a)–(e) in figure 4.1, plus graphs related to these by Hermitian conjugation or a vertical proton-antiproton flip (denoted by $p \leftrightarrow \bar{p}$).[3] Diagrams (d) and (e) also have 'seagull' versions where the two gluon attachments to the lower scalar spectator line are merged into one. There are also diagrams that only involve spectator-spectator exchanges. However, the leading-power contribution from this class of diagrams cancels after the sum over possible final-state cuts, analogous to the cancellation of the single-gluon spectator-spectator exchange graph. Besides the color-entangled graphs, there are graphs that already have the color structure anticipated by the factorization theorem. An example of one of those unentangled graphs is diagram (f) in figure 4.2. Note that for none of the graphs in figures 4.1 and 4.2 we have specified any final-state cuts. The reason for this is that the possible cut positions depend on the considered scalings of the gluon momenta; the latter is discussed in the next subsection.

As it turns out, the contributions from diagrams (d) and (e) and their seagull versions cancel after the sum over cuts, along with all unentangled diagrams (except diagram (f) and its Hermitian conjugate); we refer to appendix A of [2] for a review of these cancellations. Hence, for the calculation of the dBM term we are left with diagrams (a)–(c), and (f). Since diagram (f) already comes with the expected $1/N_c$ color factor, we do not discuss this graph explicitly in the remainder of this chapter.

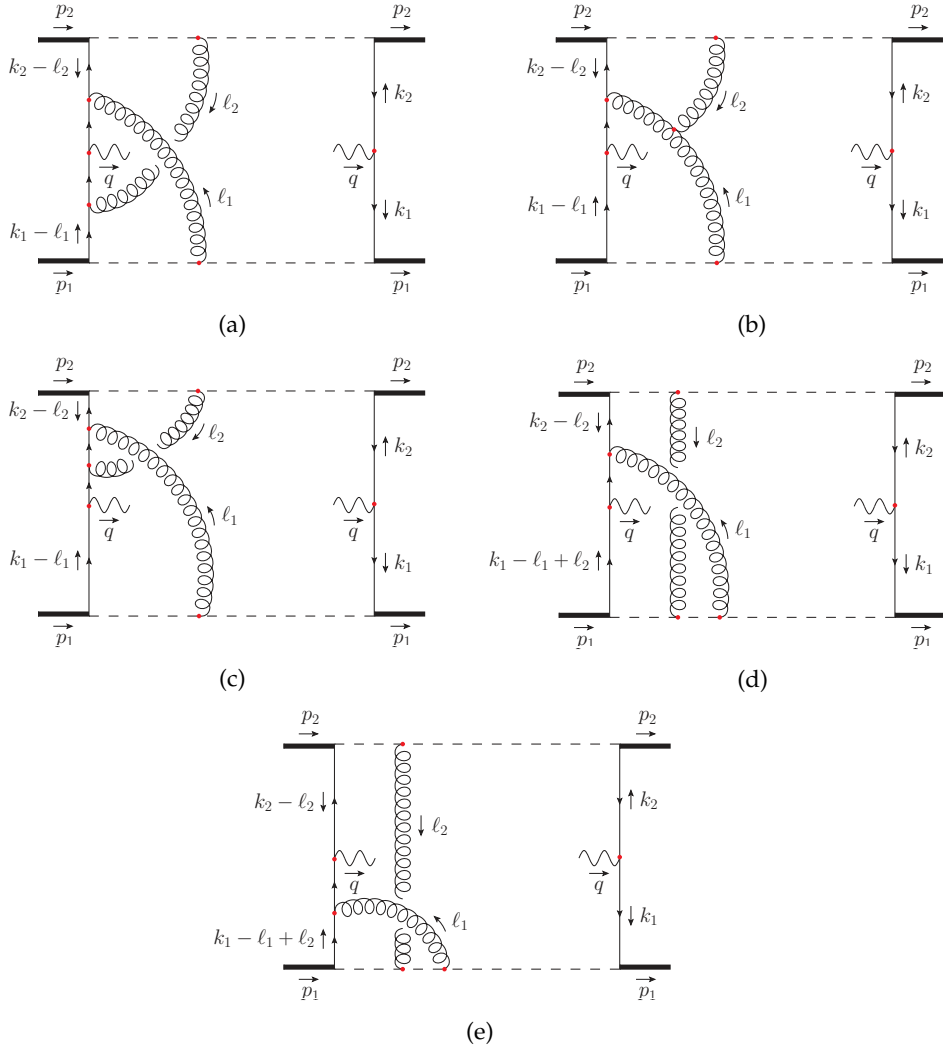3 For the $p \leftrightarrow \bar{p}$ versions we also flip the particle labels, i.e. $1 \leftrightarrow 2$.

Figure 4.1: Color-entangled diagrams contributing to the dBM part of the DY cross section in the model at $\mathcal{O}(\alpha_s^2)$. This set is supplemented by graphs that can be obtained by $p \leftrightarrow \bar{p}$ or Hermitian conjugation, and for diagrams (d) and (e) there are also seagull versions.
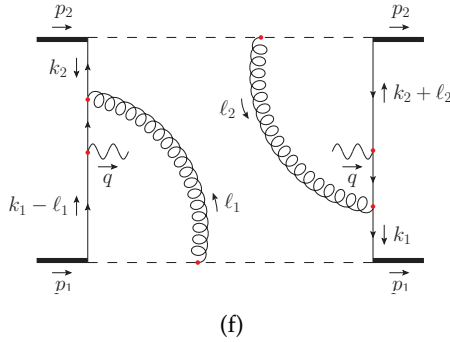
(f)

Figure 4.2: An example of an unentangled diagram contributing to the dBM part of the DY cross section in the model at $\mathcal{O}(\alpha_s^2)$.

4.3.2 Leading regions

As discussed in section 3.3, we can identify leading momentum regions for each graph to simplify the calculation. The individual region contributions can be summed together using the Collins subtraction procedure to reconstruct the full graph at leading power. From the power counting analysis applied to the active-spectator gluon exchange graph in figure 3.4, it is easy to understand which gluon scalings give a leading-power contribution to diagram (a), which features two active-spectator exchanges. Recall that for the single-gluon exchange graph the leading regions are U, G, G_1, and C_1, in the *spin-independent* case. However, now we are interested in contributions to the dBM term; selecting only transversely polarized quarks and antiquarks according to eq. (4.4), it turns out that the U region is power suppressed due to the numerator structure. Hence, in diagram (a) the leading regions for the gluon with momentum ℓ_1 are G, G_1, and C_1, while for the gluon with momentum ℓ_2 they are G, G_2, and C_2. Even though we know from subsection 3.4.2 that for diagram (a) the integration contours are not trapped in the Glauber regions, we will consider Glauber scalings explicitly in our analysis. As will become clear later, rather than summing each diagram individually over all leading regions and subsequently summing the full graph contributions, it turns out to be convenient to first sum all diagrams for a given region and then combine the various region contributions. Since trappings in the Glauber region occur for other relevant graphs, it is important to consider the Glauber scalings explicitly for all graphs.

For diagrams (a)–(c) and (f) we identify four common momentum regions for the gluon loop momenta that give leading-power contributions. We use the notation AB to describe the various regions, where A denotes the momentum scaling of the gluon with momentum ℓ_1 and B that of the gluon with momentum ℓ_2. The four leading regions are G_1G_2, C_1G, GC_2, and C_1C_2. For particular graphs one can identify other regions that give a leading-power contribution. However, the

contributions from those regions can be absorbed into the contributions from the regions considered (this is analogous to what happens to the U region in eq. (3.25)). We find that diagram (c) only gives a leading contribution in the GC_2 region, so we will not consider this graph explicitly in the other regions (and likewise for the $p \leftrightarrow \bar{p}$ version of diagram (c), which only receives a leading contribution from the $C_1 G$ region).

As prescribed by the Collins subtraction procedure (see subsection 3.3.2), we consider the contributions from each region with subtractions from the smaller regions. The region $G_1 G_2$ is the smallest one, in the sense that the pinch surface it corresponds to is the single point $\ell_1 = \ell_2 = 0$ in the eight-dimensional space spanned by ℓ_1 and ℓ_2. The regions $C_1 G$ and GC_2 are larger than this – their pinch surfaces are lines that intersect at $\ell_1 = \ell_2 = 0$. Finally, $C_1 C_2$ is the largest region with a pinch surface that is a plane. For each momentum region AB we apply an appropriate approximator T_{AB} to the graph Γ that reduces to the unit operator (up to power corrections) in the design region AB. From eq. (3.23) it follows that the contributions from each leading region are computed as

$$C_{G_1 G_2} \Gamma = T_{G_1 G_2} \Gamma, \tag{4.6}$$

$$C_{C_1 G} \Gamma = T_{C_1 G} \left(1 - T_{G_1 G_2} \right) \Gamma, \tag{4.7}$$

$$C_{GC_2} \Gamma = T_{GC_2} \left(1 - T_{G_1 G_2} \right) \Gamma, \tag{4.8}$$

$$C_{C_1 C_2} \Gamma = T_{C_1 C_2} \left(1 - T_{C_1 G} - T_{GC_2} \right) \left(1 - T_{G_1 G_2} \right) \Gamma. \tag{4.9}$$

The regions $C_1 G$, $G_1 G_2$, and GC_2 are all of the same virtuality (the $C_1 C_2$ region is of higher virtuality), meaning that they have the same number of powers of λ in their phase space $\int d^4 \ell_1 \, d^4 \ell_2$. In a sense, they are just separated by rapidity. What we mean by this is easiest understood in the context of diagram (b), where we have a gluon with momentum $\ell_1 + \ell_2$ produced by the three-gluon vertex. This gluon has the same virtuality in the $C_1 G$, $G_1 G_2$, and GC_2 regions, but moves in rapidity space from being C_1 in the $C_1 G$ region, S in the $G_1 G_2$ region, and finally C_2 in the GC_2 region. We know from subsection 3.3.3 that whenever regions are only separated by rapidity, we need to include a rapidity regulator to avoid ill-defined results. For all regions we use regulators of the type (3.34), and we choose slightly different regulators for the different graphs:

$$\text{diagram (a):} \quad \frac{1}{2} \left(\left| \frac{\ell_1^+}{\nu} \right|^{-\eta_1} \left| \frac{\ell_2^-}{\nu} \right|^{-\eta_2} + \left| \frac{\ell_1^+}{\nu} \right|^{-\eta_1} \left| \frac{\ell_2^-}{\nu} \right|^{-\eta_2} \right), \tag{4.10}$$

$$\text{diagram (b):} \quad \left| \frac{\ell_1^+}{\nu} \right|^{-\eta_1} \left| \frac{\ell_2^-}{\nu} \right|^{-\eta_2} \quad \text{with } \eta_1 \gg \eta_2, \tag{4.11}$$

$$\text{diagram (b) with } p \leftrightarrow \bar{p}: \quad \left| \frac{\ell_1^+}{\nu} \right|^{-\eta_{\bar{1}}} \left| \frac{\ell_2^-}{\nu} \right|^{-\eta_2} \quad \text{with } \eta_{\bar{1}} \ll \eta_{\bar{2}}. \tag{4.12}$$

Choosing regulators on a graph-by-graph basis is allowed as a full graph Γ does not suffer from rapidity divergences and thus does not require a rapidity regulator. At the end of the calculation we can take the limits $\eta_i \to 0$. Note however that for the three-gluon vertex graphs we have to take these limits in a specific order to obtain a well-defined result. Once we sum all graphs, diagram (b) needs to be combined with the first term in (4.10), while its $p \leftrightarrow \bar{p}$ version needs to be combined with the second term. Alternative rapidity regulators are discussed in appendix B of [2].

In section 4.5 we will show that upon summing over all graphs, leading regions, and possible final-state cuts we obtain a disentangled result with the standard $1/N_c$ color factor, in agreement with the TMD factorization theorem. As it turns out, the *full* contribution to the dBM term comes from the $G_1 G_2$ region – the contributions from the $C_1 G$, $G C_2$, and $C_1 C_2$ regions all vanish at leading power. This, however, is a regulator-dependent statement – it is possible to shift contributions between regions by choosing different rapidity regulators. The fact that the dBM contribution at the considered order can be entirely ascribed to the $G_1 G_2$ region implies that these Glauber effects can be absorbed in the TMDs – after all, there is no explicit Glauber function in the factorization formula. Indeed, we will find that the Glauber gluons can be absorbed in the collinear subgraphs by deformations of the loop momentum contours, fully consistent with the CSS approach; this is discussed in subsection 4.5.3.

4.4 THE BOER-MULDERS FUNCTION

In our $\mathcal{O}(\alpha_s^2)$ calculation of the dBM contribution in the next section, we will not assume but rather *derive* factorization. To be able to later identify the pieces that represent the quark and antiquark BM functions, we calculate $h_1^\perp$ and $\bar{h}_1^\perp$ to leading power in the model. However, before we resort to the model setting, let us first look at the quark BM function in QCD. The operator definition of $h_1^\perp$ is given in eq. (3.6), where the past-pointing gauge link $U_{[0,z]}^{[-]}$ reflects a sum over infinitely many gluon emissions. The $\mathcal{O}(g)$ contribution in the expansion of the gauge link (we will refer to this as the first order), which corresponds to the emission of a single gluon, is given by the diagram in figure 4.3 (plus its Hermitian conjugate).
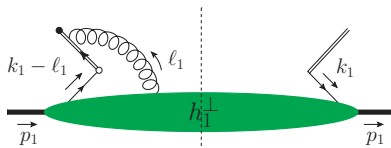

Figure 4.3: The first-order contribution to the quark BM function $h_1^\perp$ (also the Hermitian conjugate graph is needed).

Whether one ends up with a contribution to either a future- or a past-pointing gauge link is determined by the sign of the $i\epsilon$ term in the denominator of the

eikonal propagator. This sign is different for DY and SIDIS, as it depends on the direction of color flow. In the following we will argue that a plus sign results in a past-pointing gauge link. Using the Feynman rules for eikonal lines and vertices given in appendix C, the BM function in DY is given by[4]

$$
\begin{aligned}
\frac{\widetilde{k}_{1T}^{j}}{M}\, h_1^{\perp}(x_1, \mathbf{k}_1^2) =& \int \frac{dz^- d^2\mathbf{z}}{(2\pi)^3}\, e^{ik_1 \cdot z}\, \langle p_1 | \overline{\psi}(0)\, \Gamma_T^j\, \psi(z)\, | p_1 \rangle \Big|_{z^+=0} \\
&+ \left\{ (-ig t^a n^\mu) \int d\ell_1^+ d^2\boldsymbol{\ell}_1 \left(-\frac{i}{\ell_1 \cdot n + i\epsilon} \right) \int \frac{dz^- d^2\mathbf{z}}{(2\pi)^3} \int \frac{dz'^- d^2\mathbf{z}'}{(2\pi)^3} \right. \\
&\quad \times\, e^{ik_1 \cdot z}\, e^{i\ell_1 \cdot (z'-z)}\, \langle p_1 | \overline{\psi}(0)\, A_\mu^a(z')\, \Gamma_T^j\, \psi(z)\, | p_1 \rangle \Big|_{z^+=z'^+=0} \\
&+ \text{h.c.} \Bigg\} + \mathcal{O}(g^2) \\
=& \int \frac{dz^- d^2\mathbf{z}}{(2\pi)^3}\, e^{ik_1 \cdot z}\, \langle p_1 | \overline{\psi}(0)\, \Gamma_T^j\, \psi(z)\, | p_1 \rangle \Big|_{z^+=0} \\
&+ \left\{ \int \frac{dz^- d^2\mathbf{z}}{(2\pi)^3}\, e^{ik_1 \cdot z}\, \langle p_1 | \overline{\psi}(0)\, \Gamma_T^j \left[ig \int \frac{dz'^- d^2\mathbf{z}'}{(2\pi)^3} \right. \right. \\
&\quad \times \left. \int d\ell_1^+ d^2\boldsymbol{\ell}_1\, \frac{i}{\ell_1 \cdot n + i\epsilon}\, e^{i\ell_1 \cdot (z'-z)}\, n \cdot A(z') \right] \psi(z)\, | p_1 \rangle \Big|_{z^+=z'^+=0} \\
&+ \text{h.c.} \Bigg\} + \mathcal{O}(g^2).
\end{aligned}
\tag{4.13}
$$

The term on the first line represents the $\mathcal{O}(g^0)$ contribution, corresponding to no gluon emissions. Since this term does not depend on an eikonal propagator (in other words, it is *independent* of the direction of color flow), it is the same for DY and SIDIS and thus it must vanish. However, since we will identify this term with the $\mathcal{O}(g^0)$ contribution to the gauge link, it is better to keep it here. The terms in between curly brackets make up the $\mathcal{O}(g)$ contribution. They contain the Fourier transform of the trilocal quark-gluon-quark matrix element, evaluated at the light-front $z^+ = z'^+ = 0$. Making use of the integral representation

$$
\frac{i}{x \pm i\epsilon} = \pm \int_0^\infty d\zeta\, e^{i\zeta(\pm x + i\epsilon)},
\tag{4.14}
$$

the quantity in between square brackets in eq. (4.13) can be rewritten as

$$
\begin{aligned}
& ig \int \frac{dz'^- d^2\mathbf{z}'}{(2\pi)^3} \int d\ell_1^+ d^2\boldsymbol{\ell}_1\, \frac{i}{\ell_1 \cdot n + i\epsilon}\, e^{i\ell_1 \cdot (z'-z)}\, n \cdot A(z') \Big|_{z^+=z'^+=0} \\
&= ig \int_0^\infty d\zeta \int \frac{dz'^- d^2\mathbf{z}'}{(2\pi)^3} \int d\ell_1^+ d^2\boldsymbol{\ell}_1\, e^{i(\zeta n + z' - z) \cdot \ell_1}\, n \cdot A(z') \Big|_{z^+=z'^+=0}
\end{aligned}
$$

---

[4] For convenience we suppress in this section any reference to quark flavors.

$$= ig \int_0^\infty d\zeta \, n \cdot A(z - \zeta n) \Big|_{z^+ = 0}$$

$$= ig \int_{-\infty}^{z^-} d\eta^- A^+(\eta^+ = 0, \eta^-, \boldsymbol{\eta} = \boldsymbol{z}), \tag{4.15}$$

where in the last step we defined $\eta \equiv z - \zeta n$. Once we also include the graph with one gluon emission on either side of the final-state cut, eq. (4.13) becomes

$$\frac{\widetilde{k}_{1T}^j}{M} h_1^\perp(x_1, \boldsymbol{k}_1^2) = \int \frac{dz^- d^2 \boldsymbol{z}}{(2\pi)^3} \, e^{ik_1 \cdot z} \, \langle p_1 | \, \overline{\psi}(0) \, \Gamma_T^j$$

$$\times \left[1 + ig \int_0^{-\infty} d\eta^- A^+(\eta^+ = 0, \eta^-, \boldsymbol{\eta} = \boldsymbol{0}) \right]$$

$$\times \left[1 + ig \int_{-\infty}^{z^-} d\eta^- A^+(\eta^+ = 0, \eta^-, \boldsymbol{\eta} = \boldsymbol{z}) \right] \psi(z) \, | p_1 \rangle \Big|_{z^+ = 0}$$

$$+ \mathcal{O}(g^2). \tag{4.16}$$

The two factors in between square brackets correspond to the expansions up to $\mathcal{O}(g)$ of the Wilson lines $U_{[0,-\infty;\boldsymbol{0}]}^n$ and $U_{[-\infty,z^-;\boldsymbol{z}]}^n$ defined in eq. (2.18). Since we have adopted the Feynman gauge, there is no contribution from the transverse link $U_{[-\infty;\boldsymbol{0},\boldsymbol{z}]}^T$ defined in eq. (2.44). Note that up to the considered order, the path ordering of the Wilson lines does not play a role (this starts at $\mathcal{O}(g^2)$). Summing over all gluon emissions (see e.g. [33]), one indeed precisely recovers eq. (3.6), featuring the past-pointing gauge link $U_{[0,z]}^{[-]}$.

If we had taken the $i\epsilon$ term in the eikonal propagator with a minus sign instead, corresponding to the SIDIS case, we would have ended up with the future-pointing gauge link $U_{[0,z]}^{[+]}$. Hence, the final gauge link structure is indeed connected to the $i\epsilon$ term in the eikonal propagator denominator $\ell_1^+ \pm i\epsilon$. Since the BM function is process dependent (it changes sign between DY and SIDIS), it requires sensitivity to the imaginary part of this propagator, which is only the case when $\ell_1^+ \to 0$. As will be shown below, this is consistent with the fact that the first-order contribution to the BM function in the model turns out to originate from the region where the gluon has Glauber scaling.

Now let us switch to the spectator model. Naively, one would need to compute the BM functions up to the order at which we work, namely $\mathcal{O}(\alpha_s^2)$. However, since the BM functions are T-odd they each require a gluon attachment to the eikonal line – thus, there is no tree-level contribution and it suffices to compute each function only up to $\mathcal{O}(\alpha_s)$. At the $\mathcal{O}(\alpha_s)$ level, there are two contributions to $h_1^\perp$, see figure 4.4 (also the Hermitian conjugate graphs are needed). We note that there is no contribution coming from the diagram on the right where the gluon attaches to the active quark line – its numerator vanishes:

$$\text{Tr} \left(\Gamma_T^j \, \slashed{k}_1 \, (\slashed{p}_1 + M) \, \slashed{k}_1 \, (\slashed{k}_1 - \slashed{\ell}_1) \, \slashed{n} \right) = 0. \tag{4.17}$$

By contrast, the graph on the left does not vanish and gives the first-order contribution to the past-pointing Wilson line.
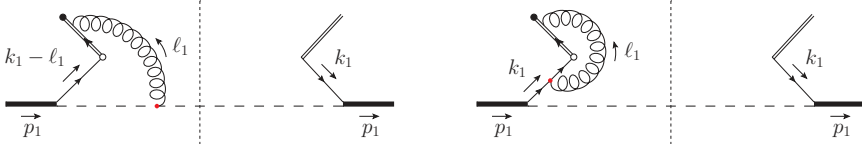

Figure 4.4: The two $\mathcal{O}(\alpha_s)$ contributions to the quark BM function $h_1^\perp$ in the scalar spectator model (also the Hermitian conjugate graphs are needed).

To calculate the BM function $h_1^\perp$ in the model, i.e. the diagram on the left in figure 4.4 plus its conjugate, we first identify two (nontrivial) momentum regions for the gluon momentum ℓ_1 that could give a leading-power contribution, namely G_1 and C_1 (the power counting is similar to that for the single-gluon exchange graph in figure 3.4). The region G can be absorbed in G_1, and U gives a power-suppressed contribution for the same reason as discussed earlier in subsection 4.3.2. From eqs. (3.23) and (3.24) it follows that the sum over the G_1 and C_1 region contributions is given by

$$C_{G_1}\Gamma + C_{C_1}\Gamma = T_{G_1}\Gamma + T_{C_1}\left(1 - T_{G_1}\right)\Gamma. \tag{4.18}$$

As mentioned already in subsection 4.3.2 and as we will show explicitly in section 4.5, for our choice of rapidity regulators the full contribution to the dBM term turns out to come from the G_1G_2 region. Hence, we expect only the first term in eq. (4.18) to be nonzero.

Before applying any momentum approximations, the quark BM function in the spectator model is given by

$$\frac{\widetilde{k}_{1T}^j}{M} h_1^\perp(x_1, \mathbf{k}_1^2) = -iC_\Phi \int \frac{d\ell_1^+}{2\pi} \, (2p_1 - 2k_1 + \ell_1) \cdot n \, \chi^j(x_1, \mathbf{k}_1) \frac{v^{\eta_1} |\ell_1 \cdot n|^{-\eta_1}}{\ell_1 \cdot n + i\epsilon} + \text{h.c.}, \tag{4.19}$$

where we have included the rapidity regulator η_1 (which can be sent to zero at the end of this calculation), as well as the rapidity scale v. All relevant Feynman rules are listed in appendix C. Note that $h_1^\perp$ is manifestly real due to the presence of the Hermitian conjugate term (denoted by 'h.c.'). The color factor C_Φ is defined in eq. (4.5), and we have defined

$$\chi^j(x_1, \mathbf{k}_1) \equiv \pi g^2 \int \frac{dk_1^-}{(2\pi)^4} \, \theta\big((p_1 - k_1)^0\big) \, \delta\big((p_1 - k_1)^2 - m_s^2\big) \int \frac{d\ell_1^-}{2\pi} \int \frac{d^2\ell_1}{(2\pi)^2}$$

$$\times \frac{D_1^j}{[(k_1 - \ell_1)^2 + i\epsilon] \, [(p_1 - k_1 + \ell_1)^2 - m_s^2 + i\epsilon] \, [\ell_1^2 + i\epsilon] \, [k_1^2 - i\epsilon]} \,. \tag{4.20}$$

The Dirac trace D_1^j is given by

$$D_1^j \equiv \text{Tr}\left(\Gamma_T^j \left(\slashed{k}_1 - \slashed{\ell}_1\right) (\slashed{p}_1 + M) \slashed{k}_1\right) = 2iM\left(x_1 p_1^+ \widetilde{\ell}_{1T}^j - \ell_1^+ \widetilde{k}_{1T}^j\right). \tag{4.21}$$

Let us first calculate the contribution from the G_1 region. We expand $h_1^\perp$ up to leading power in λ and subsequently perform the integrals over k_1^- and ℓ_1^-. The delta function $\delta\big((p_1 - k_1)^2 - m_s^2\big)$ is used for the integration over k_1^- and for the integration over ℓ_1^- we invoke Cauchy's residue theorem. To leading power, the BM function is given by

$$\frac{\widetilde{k}_{1T}^j}{M} h_1^\perp(x_1, k_1^2) = -2i\, C_\Phi\, (1 - x_1) p_1^+ \chi^j(x_1, k_1) \int \frac{d\ell_1^+}{2\pi} \frac{\nu^{\eta_1} |\ell_1^+|^{-\eta_1}}{\ell_1^+ + i\epsilon} + \text{h.c.}, \qquad (4.22)$$

where, using the shorthand notation $\Lambda_1^2 \equiv x_1 m_s^2 - x_1(1 - x_1)M^2$,

$$\chi^j(x_1, k_1) = \frac{ig^2}{64\pi^3} \int \frac{d^2\ell_1}{(2\pi)^2} \frac{\theta(x_1)\, \theta(1 - x_1)\, D_1^j}{(p_1^+)^2 \left[(k_1 - \ell_1)^2 + \Lambda_1^2\right] (k_1^2 + \Lambda_1^2)\, \ell_1^2}, \qquad (4.23)$$

and

$$D_1^j = 2iM\, x_1 p_1^+ \widetilde{\ell}_{1T}^j. \qquad (4.24)$$

Inasmuch as the quantity χ is real, only the imaginary part of the ℓ_1^+ integral contributes to $h_1^\perp$ – its real part is canceled by the Hermitian conjugate term. This imaginary part comes from the region where ℓ_1^+ is sensitive to the $i\epsilon$ term in the denominator, which is the case for $\ell_1^+ \to 0$, i.e. when ℓ_1 has Glauber scaling. Note that this is consistent with our earlier observation – the sensitivity to the $i\epsilon$ term is connected to the required process dependence of this T-odd function.

Now let us consider the C_1 momentum region. Here, it is sufficient to consider what happens at a fixed nonzero value of ℓ_1^+, as the small ℓ_1^+ region is suppressed due to the subtraction (this is explained in subsection 3.3.2). This means that we can drop the $i\epsilon$ term in the eikonal propagator, making it purely real. We consider the subtraction and naive graph terms ($T_{C_1} T_{G_1} \Gamma$ and $T_{C_1} \Gamma$, respectively) separately at this nonzero ℓ_1^+. The contribution to $h_1^\perp$ from the subtraction term is also given by eq. (4.22), but vanishes at finite ℓ_1^+ due to the cancellation between the amplitude and its conjugate. At leading power, the contribution from the naive graph term is given by

$$\frac{\widetilde{k}_{1T}^j}{M} h_1^\perp(x_1, k_1^2) = -i\, C_\Phi \int \frac{d\ell_1^+}{2\pi} \left[1 + \frac{2(1 - x_1)p_1^+}{\ell_1^+}\right] \left|\frac{\ell_1^+}{\nu}\right|^{-\eta_1} \chi^j(x_1, k_1, \ell_1^+) + \text{h.c.} \quad (4.25)$$

The precise form for χ is not relevant here – what matters is that it is again real-valued for nonzero ℓ_1^+ such that the amplitude and conjugate contributions cancel. Hence, the C_1 region does not contribute to the BM function and the full contribution comes from the G_1 region only.

The G_1 region result in eq. (4.22) can be simplified further. The integral over ℓ_1 can be performed using standard techniques, including Passarino-Veltman reduction [177] and the use of Feynman parameters. Performing also the integral over

ℓ_1^+ and subsequently taking the limit $\eta_1 \to 0$, we find that the BM function in our spectator model is given by

$$h_1^\perp(x_1, \boldsymbol{k}_1^2) = \frac{g^2}{16\pi^2}\, C_\Phi\, \frac{x_1(1-x_1)M^2}{\boldsymbol{k}^2(\boldsymbol{k}^2 + \Lambda_1^2)}\, \ln\left(\frac{\boldsymbol{k}^2 + \Lambda_1^2}{\Lambda_1^2}\right),\qquad (4.26)$$

which is consistent with the results in [149,178]. The fully integrated result for $h_1^\perp$ in eq. (4.26) is not needed for our calculations in the next section; the form in eq. (4.22) turns out to be sufficient for our purposes.

In the same way we can obtain the BM function for the antiquark, where the full contribution comes from the G_2 region. Since the kinematical setup is invariant under the simultaneous interchange of plus and minus indices and the particle labels 1 and 2, $\bar{h}_1^\perp$ is simply obtained from $h_1^\perp$ by the two substitutions $+ \to -$ and $1 \to 2$.

4.5 CALCULATION OF THE CROSS SECTION

At the $\mathcal{O}(\alpha_s^2)$ level there are various graphs that potentially contribute to the dBM term in the DY cross section, see figures 4.1 and 4.2. In subsection 4.3.1 we argued that some of these graphs vanish upon summing over cuts, and in subsection 4.3.2 we already mentioned that in all regions the sum over graphs and cuts turns out to give zero except in the $G_1 G_2$ region. In this section we present the leading-power calculation of all nonvanishing color-entangled graphs, considering all possible cut positions. To this end, we use the decomposition in eq. (4.4) to select transversely polarized quarks and antiquarks. Flavor labels, as well as the sum over different flavors, will be implicit (so the quantity e^2 should be thought of as $\sum_q e_q^2$). We discuss the $G_1 G_2$ region computation in detail and then comment briefly on the other region contributions. First, let us rearrange the various integrations that appear in the differential cross section contributions and discuss some overall factors.

Each diagram contains integrations over the internal four-momenta ℓ_1, ℓ_2, k_1, k_2, and q. The integrations over ℓ_1 and ℓ_2, as well as over the components k_1^- and k_2^+, are contained in the quark and antiquark BM functions, as discussed in the previous section. The remaining integrals, including the two-particle phase space element, can be cast into the following form at leading power:

$$d\sigma_{\text{dBM}} \propto \int dk_1^+ d^2\boldsymbol{k}_1 \int dk_2^- d^2\boldsymbol{k}_2\, (2\pi)^4\, \delta^{(4)}(k_1 + k_2 - q) \int \frac{d^4q}{(2\pi)^4}$$

$$\times (2\pi)^4\, \delta^{(4)}(\ell + \ell' - q)\, \frac{d^3\vec{l}}{(2\pi)^3\, 2l^0}\, \frac{d^3\vec{l}'}{(2\pi)^3\, 2l'^0}$$

$$= \frac{1}{8} \int dk_1^+ d^2\boldsymbol{k}_1 \int dk_2^- d^2\boldsymbol{k}_2\, (2\pi)^4\, \delta^{(4)}(k_1 + k_2 - q)\, \frac{d^4q}{(2\pi)^4}\, \frac{d\Omega}{(2\pi)^2}$$

$$= \frac{p_1^+ p_2^-}{32\pi^2} \int d^2\boldsymbol{k}_1 \int d^2\boldsymbol{k}_2\, \delta^{(2)}(\boldsymbol{k}_1 + \boldsymbol{k}_2 - \boldsymbol{q})\, d\Omega\, dx_1 dx_2\, d^2\boldsymbol{q}\,,\qquad (4.27)$$

where $l^0 = \sqrt{\vec{l}^2}$, $l'^0 = \sqrt{\vec{l}'^2}$, and $d\Omega \equiv \sin\theta\, d\theta\, d\phi$, in terms of the CS angles θ and ϕ (see figure 3.2).

For our cross section calculation we include the usual flux factor, which to leading power is given by

$$\frac{1}{4\sqrt{(p_1 \cdot p_2)^2 - M^4}} = \frac{1}{4p_1^+ p_2^-} \, . \tag{4.28}$$

The averaging over initial-state polarizations leads to a factor of $\frac{1}{4}$ (each hadron can have two spin polarizations); this factor has been absorbed in the BM functions.

4.5.1 *Contribution from the $G_1 G_2$ region*

(a,i)   (b,i)

(b,v)
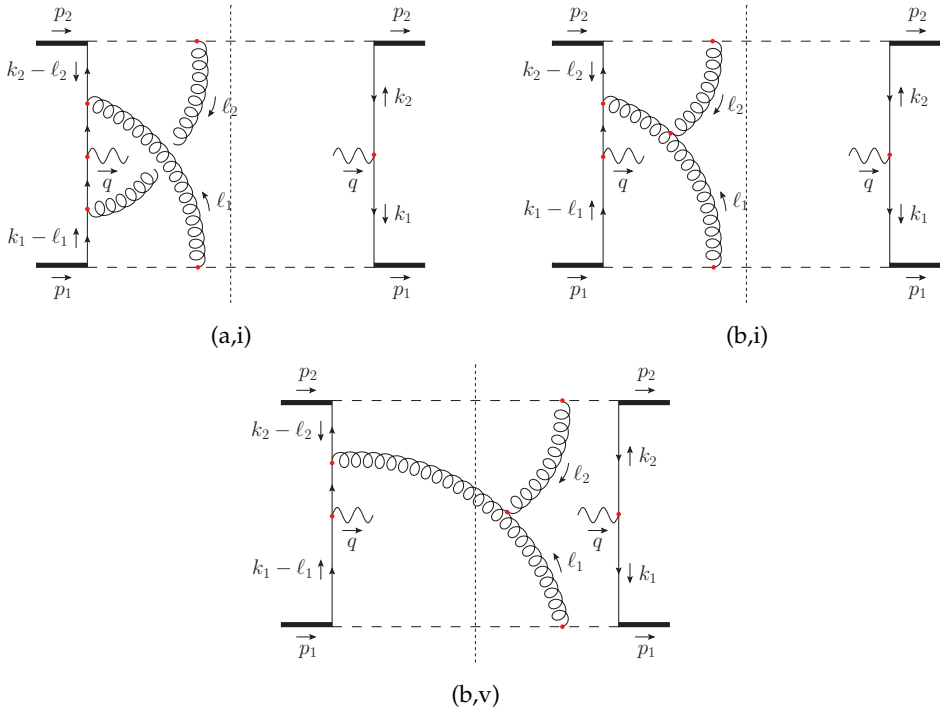
Figure 4.5: The color-entangled two-gluon exchange graphs that contribute to the dBM part of the DY cross section in the $G_1 G_2$ region. This set is supplemented by graphs that can be obtained by $p \leftrightarrow \bar{p}$ or Hermitian conjugation.

For the $G_1 G_2$ region, we need to consider all possible final-state cuts for graphs (a) and (b) given in figure 4.1. Since cuts through Glauber modes are not permitted, there is only one possible cut position for graph (a), whereas there are two for graph (b). Denoting the cut positions with roman numerals, the three diagrams that need

to be computed are (a,i), (b,i), and (b,v), see figure 4.5 (in subsection 4.5.2 it becomes clear why we refer to the cut through the soft gluon line in graph (b) as the fifth cut). In addition to these diagrams, also the ones that can be obtained through $p \leftrightarrow \bar{p}$ or Hermitian conjugation are needed.

The differential cross section can be computed by summing over the relevant cut graphs. For the $G_1 G_2$ region, the sum can be written as follows:

$$
\frac{d\sigma_{\text{dBM}}}{d\Omega\, dx_1 dx_2\, d^2\boldsymbol{q}} = 2 \left[\frac{1}{2} \left(\frac{d\sigma_{\text{dBM}}}{d\Omega\, dx_1 dx_2\, d^2\boldsymbol{q}} \right)_{(a,i)} + \left(\frac{d\sigma_{\text{dBM}}}{d\Omega\, dx_1 dx_2\, d^2\boldsymbol{q}} \right)_{(b,i)} \right.
$$
$$
\left. + \left(\frac{d\sigma_{\text{dBM}}}{d\Omega\, dx_1 dx_2\, d^2\boldsymbol{q}} \right)_{(b,v)} + \frac{1}{2} \left(\frac{d\sigma_{\text{dBM}}}{d\Omega\, dx_1 dx_2\, d^2\boldsymbol{q}} \right)_{(f,i)} + \text{h.c.} \right] . \quad (4.29)
$$

The factor of 2 in front arises from taking into account the graphs that can be obtained from (b,i) and (b,v) (and their Hermitian conjugates) by $p \leftrightarrow \bar{p}$. Diagram (f,i) is given in figure 4.6. Although this graph already comes with the expected $1/N_c$ color factor and will not play any role in disentangling the color structures of diagrams (a,i), (b,i), and (b,v), it is needed to obtain the full contribution from the $G_1 G_2$ region. This graph is not considered here explicitly.

(f,i)
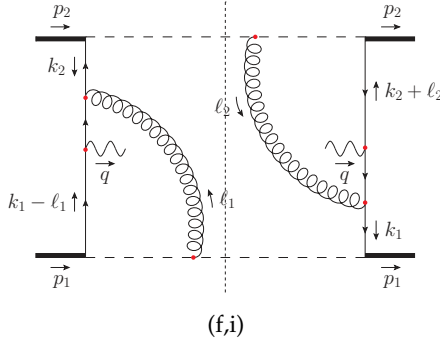
Figure 4.6: The only unentangled diagram that contributes to the dBM part of the DY cross section in the $G_1 G_2$ region (also its Hermitian conjugate is needed).

We now proceed with the leading-power calculation of the dBM contributions from diagrams (a,i), (b,i), and (b,v) in the $G_1 G_2$ region. In the following, all diagrams are expressed in terms of the functions χ (defined in eq. (4.20) and simplified in eq. (4.23)) and $\overline{\chi}$ (the antiquark analogue of χ).

Diagram (a,i). The dBM contribution from diagram (a,i) is given by

$$
\frac{1}{2} \left(\frac{d\sigma_{\text{dBM}}}{d\Omega\, dx_1 dx_2\, d^2\boldsymbol{q}} \right)_{(a,i)} = -\frac{1}{16} \frac{\alpha^2}{q^4}\, e^2\, C_{(a)} \int d^2\boldsymbol{k}_1 \int \frac{d\ell_1^+}{2\pi}\, \chi^j(x_1, \boldsymbol{k}_1) \int d^2\boldsymbol{k}_2 \int \frac{d\ell_2^-}{2\pi}
$$

$$\times \, \overline{\chi}^k(x_2, \mathbf{k}_2) \, \frac{v^{\eta_1 + \eta_2} |\ell_1^+|^{-\eta_1} |\ell_2^-|^{-\eta_2} \, R_{(a)jk;\mu\nu} \, L^{\mu\nu}}{[(k_2 - \ell_2 + \ell_1)^2 + i\epsilon] \, [(k_1 - \ell_1 + \ell_2)^2 + i\epsilon]}$$
$$\times \, \delta^{(2)}(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{q}), \tag{4.30}$$

where only the first term of the regulator defined in (4.10) appears. As explained in subsection 4.3.2, the other 'half' of diagram (a,i) comes with the second term in (4.10) and is ultimately combined with the $p \leftrightarrow \bar{p}$ versions of diagrams (b,i) and (b,v). The color factor $C_{(a)}$ is given by[5]

$$C_{(a)} \equiv \mathrm{Tr}\left(t^a t^b t^a t^b\right) = -\frac{1}{N_c^2 - 1} \frac{1}{N_c} C_\Phi^2 \,. \tag{4.31}$$

Note that this is precisely the color factor that has been predicted to appear in the final factorized result in [101]. The Dirac trace $R_{(a)jk;\mu\nu}$ in eq. (4.30) is defined as

$$R_{(a)jk;\mu\nu} \equiv \mathrm{Tr}\left(\overline{\Gamma}_{Tj} \, \gamma_\mu \, \Gamma_{Tk} \, (2\slashed{p}_1 - 2\slashed{k}_1 + \slashed{\ell}_1) \, (\slashed{k}_2 - \slashed{\ell}_2 + \slashed{\ell}_1)\right.$$
$$\left. \times \, \gamma_\nu \, (\slashed{k}_1 - \slashed{\ell}_1 + \slashed{\ell}_2) \, (2\slashed{p}_2 - 2\slashed{k}_2 + \slashed{\ell}_2)\right), \tag{4.32}$$

and the spin-averaged leptonic tensor $L^{\mu\nu}$ is given by

$$L^{\mu\nu} \equiv \mathrm{Tr}\left(\slashed{l} \, \gamma^\mu \, \slashed{l}' \, \gamma^\nu\right) = 4 \left(l^\mu l'^\nu + l^\nu l'^\mu - l \cdot l' g^{\mu\nu}\right). \tag{4.33}$$

Up to leading power, its contraction with $R_{(a)jk;\mu\nu}$ reads

$$R_{(a)jk;\mu\nu} L^{\mu\nu} = -128 \, H_{jk} \, x_1(1 - x_1) x_2 (1 - x_2)(p_1^+ p_2^-)^2, \tag{4.34}$$

where

$$H_{jk} \equiv l_{Tj} l'_{Tk} + l_{Tk} l'_{Tj} + \boldsymbol{l} \cdot \boldsymbol{l}' \, g_{Tjk} \,. \tag{4.35}$$

Hence, to leading power, eq. (4.30) becomes

$$\frac{1}{2} \left(\frac{d\sigma_{\mathrm{dBM}}}{d\Omega \, dx_1 dx_2 \, d^2 \boldsymbol{q}}\right)_{(a,i)} = 2 \frac{\alpha^2}{q^4} e^2 \, C_{(a)} \, (1 - x_1)(1 - x_2) p_1^+ p_2^- \int d^2 \boldsymbol{k}_1 \, \chi^j(x_1, \boldsymbol{k}_1)$$
$$\times \int d^2 \boldsymbol{k}_2 \, \overline{\chi}^k(x_2, \boldsymbol{k}_2) \, I_{(a,i)} \, H_{jk} \, \delta^{(2)}(\boldsymbol{k}_1 + \boldsymbol{k}_2 - \boldsymbol{q}), \tag{4.36}$$

where $I_{(a,i)}$ is an integral over ℓ_1^+ and ℓ_2^-, given by

$$I_{(a,i)} \equiv \int \frac{d\ell_1^+}{2\pi} \frac{v^{\eta_1} |\ell_1^+|^{-\eta_1}}{\ell_1^+ + i\epsilon} \int \frac{d\ell_2^-}{2\pi} \frac{v^{\eta_2} |\ell_2^-|^{-\eta_2}}{\ell_2^- + i\epsilon} \,. \tag{4.37}$$

Note that the integrand has initial-state poles in both ℓ_1^+ and ℓ_2^-.

---

5 Whenever a quantity is independent of the cut position, the roman numeral is simply omitted in the subscript.

Diagram (b,i). The dBM contribution from diagram (b,i) is given by

$$
\left(\frac{d\sigma_{\text{dBM}}}{d\Omega\, dx_1 dx_2\, d^2\boldsymbol{q}} \right)_{\text{(b,i)}} = -\frac{1}{8}\frac{\alpha^2}{q^4}\, e^2\, C_{\text{(b)}} \int d^2\boldsymbol{k}_1 \int \frac{d\ell_1^+}{2\pi}\, \chi^j(x_1,\boldsymbol{k}_1) \int d^2\boldsymbol{k}_2 \int \frac{d\ell_2^-}{2\pi}
$$

$$
\times\, \bar{\chi}^k(x_2,\boldsymbol{k}_2)\, \frac{\nu^{\eta_1+\eta_2}|\ell_1^+|^{-\eta_1}|\ell_2^-|^{-\eta_2}\, R_{\text{(b)}jk;\mu\nu}\, L^{\mu\nu}}{[(k_2+\ell_1)^2+i\epsilon]\,[(\ell_1+\ell_2)^2+i\epsilon]}
$$

$$
\times\, \delta^{(2)}(\boldsymbol{k}_1+\boldsymbol{k}_2-\boldsymbol{q}). \tag{4.38}
$$

The color factor $C_{\text{(b)}}$ reads

$$
C_{\text{(b)}} \equiv \text{Tr}\left(i f^{abc} t^a t^b t^c \right) = -\frac{N_c^2}{N_c^2-1}\frac{1}{N_c}\, C_\Phi^2, \tag{4.39}
$$

and the Dirac trace $R_{\text{(b)}jk;\mu\nu}$ is defined as

$$
R_{\text{(b)}jk;\mu\nu} \equiv \text{Tr}\left(\bar{\Gamma}_{Tj}\, \gamma_\mu\, \Gamma_{Tk} \left[(2p_1-2k_1+\ell_1)\cdot(2p_2-2k_2+\ell_2)\,(\ell_1-\ell_2) \right.\right.
$$

$$
+ (2p_1-2k_1+\ell_1)\cdot(\ell_1+2\ell_2)\,(2\slashed{p}_2-2\slashed{k}_2+\slashed{\ell}_2)
$$

$$
\left.\left. - (2p_2-2k_2+\ell_2)\cdot(2\ell_1+\ell_2)\,(2\slashed{p}_1-2\slashed{k}_1+\slashed{\ell}_1) \right]\,(\slashed{k}_2+\slashed{\ell}_1)\,\gamma_\nu \right). \tag{4.40}
$$

Up to leading power, its contraction with the leptonic tensor is given by

$$
R_{\text{(b)}jk;\mu\nu}\, L^{\mu\nu} = 64\, H_{jk}\,(1-x_1)x_2(1-x_2)p_1^+(p_2^-)^2\ell_1^+. \tag{4.41}
$$

Hence, to leading power, eq. (4.38) becomes

$$
\left(\frac{d\sigma_{\text{dBM}}}{d\Omega\, dx_1 dx_2\, d^2\boldsymbol{q}} \right)_{\text{(b,i)}} = -2\frac{\alpha^2}{q^4}\, e^2\, C_{\text{(b)}}\,(1-x_1)(1-x_2)p_1^+ p_2^- \int d^2\boldsymbol{k}_1\, \chi^j(x_1,\boldsymbol{k}_1)
$$

$$
\times \int d^2\boldsymbol{k}_2\, \bar{\chi}^k(x_2,\boldsymbol{k}_2)\, I_{\text{(b,i)}}\, H_{jk}\, \delta^{(2)}(\boldsymbol{k}_1+\boldsymbol{k}_2-\boldsymbol{q}), \tag{4.42}
$$

where $I_{\text{(b,i)}}$ is an integral over ℓ_1^+ and ℓ_2^-, given by

$$
I_{\text{(b,i)}} \equiv \int \frac{d\ell_1^+}{2\pi}\frac{\nu^{\eta_1}|\ell_1^+|^{-\eta_1}}{\ell_1^++i\epsilon} \int \frac{d\ell_2^-}{2\pi}\frac{2\ell_1^+\nu^{\eta_2}|\ell_2^-|^{-\eta_2}}{2\ell_1^+\ell_2^- - (\ell_1+\ell_2)^2+i\epsilon}. \tag{4.43}
$$

Note that the integrand has an initial-state pole in ℓ_1^+, and, depending on the sign of ℓ_1^+, the pole in ℓ_2^- is either an initial- or a final-state one.

Diagram (b,v). The dBM contribution from diagram (b,v) is given by

$$
\left(\frac{d\sigma_{\text{dBM}}}{d\Omega\, dx_1 dx_2\, d^2\boldsymbol{q}} \right)_{\text{(b,v)}} = -\frac{\pi^2}{8}\frac{\alpha^2}{q^4}\, e^2 g^4\, C_{\text{(b)}} \int \frac{dk_1^-\, d^2\boldsymbol{k}_1}{(2\pi)^4} \int \frac{d^4\ell_1}{(2\pi)^4}
$$

$$\times \, \theta\big((p_1 - k_1 + \ell_1)^0\big) \, \delta\big((p_1 - k_1 + \ell_1)^2 - m_s^2\big)$$

$$\times \, \frac{D_1^j}{[(k_1 - \ell_1)^2 + i\epsilon] \, [(p_1 - k_1)^2 - m_s^2 - i\epsilon] \, [\ell_1^2 - i\epsilon] \, [k_1^2 - i\epsilon]}$$

$$\times \int \frac{dk_2^+ d^2\boldsymbol{k}_2}{(2\pi)^4} \int \frac{d^4\ell_2}{(2\pi)^4} \, \theta\big((p_2 - k_2 + \ell_2)^0\big)$$

$$\times \, \delta\big((p_2 - k_2 + \ell_2)^2 - m_s^2\big)$$

$$\times \, \frac{D_2^k}{[(k_2 - \ell_2)^2 + i\epsilon] \, [(p_2 - k_2)^2 - m_s^2 - i\epsilon] \, [\ell_2^2 - i\epsilon] \, [k_2^2 - i\epsilon]}$$

$$\times \, \frac{v^{\eta_1} |\ell_1^+|^{-\eta_1} \, R_{(\text{b})jk;\mu\nu} L^{\mu\nu}}{(k_2 + \ell_1)^2 + i\epsilon} \cdot 2\pi i \, \theta\big(-(\ell_1 + \ell_2)^0\big)$$

$$\times \, \delta\big((\ell_1 + \ell_2)^2\big) \, v^{\eta_2} |\ell_2^-|^{-\eta_2} \, \delta^{(2)}(\boldsymbol{k}_1 + \boldsymbol{k}_2 - \boldsymbol{q}). \qquad (4.44)$$

We now expand this expression to leading power in λ and perform the integrals over k_1^-, ℓ_1^-, k_2^+, and ℓ_2^+. The two delta functions $\delta\big((p_1 - k_1 + \ell_1)^2 - m_s^2\big)$ and $\delta\big((p_2 - k_2 + \ell_2)^2 - m_s^2\big)$ are used for the integrations over ℓ_1^- and ℓ_2^+, and for the integrations over k_1^- and k_2^+ we invoke Cauchy's residue theorem. To leading power, eq. (4.44) becomes

$$\left(\frac{d\sigma_{\text{dBM}}}{d\Omega \, dx_1 dx_2 \, d^2\boldsymbol{q}} \right)_{(\text{b,v})} = -2 \frac{\alpha^2}{q^4} e^2 \, C_{(\text{b})} \, (1 - x_1)(1 - x_2) p_1^+ p_2^- \int d^2\boldsymbol{k}_1 \, \chi^j(x_1, \boldsymbol{k}_1)$$

$$\times \int d^2\boldsymbol{k}_2 \, \overline{\chi}^k(x_2, \boldsymbol{k}_2) \, I_{(\text{b,v})} \, H_{jk} \, \delta^{(2)}(\boldsymbol{k}_1 + \boldsymbol{k}_2 - \boldsymbol{q}), \qquad (4.45)$$

where $I_{(\text{b,v})}$ is an integral over ℓ_1^+ and ℓ_2^-, given by

$$I_{(\text{b,v})} \equiv 4\pi i \int \frac{d\ell_1^+}{2\pi} \, \frac{v^{\eta_1} |\ell_1^+|^{-\eta_1}}{\ell_1^+ + i\epsilon} \int \frac{d\ell_2^-}{2\pi} \, \theta(-\ell_1^+) \, \ell_1^+$$

$$\times \, \delta\big(2\ell_1^+ \ell_2^- - (\boldsymbol{\ell}_1 + \boldsymbol{\ell}_2)^2\big) \, v^{\eta_2} |\ell_2^-|^{-\eta_2}. \qquad (4.46)$$

Note that the integrand has an initial-state pole in ℓ_1^+, and both an initial- and a final-state pole in ℓ_2^-. The latter can be seen from the identity

$$2\pi i \, \delta(x) = \frac{1}{x - i\epsilon} - \frac{1}{x + i\epsilon}, \qquad (4.47)$$

which is consistent with eq. (4.14).

Sum of the diagrams. We can use eq. (4.29) to combine the results from diagrams (a,i), (b,i), and (b,v) to obtain the full $G_1 G_2$ region contribution to the DY cross section. Employing eqs. (4.36), (4.42), (4.45), and that $C_{(\text{b})} = N_c^2 \, C_{(\text{a})}$, gives to leading power

$$\frac{d\sigma_{\text{dBM}}}{d\Omega \, dx_1 dx_2 \, d^2\boldsymbol{q}} = 4 \frac{\alpha^2}{q^4} e^2 \, C_{(\text{a})} \, (1 - x_1)(1 - x_2) p_1^+ p_2^- \int d^2\boldsymbol{k}_1 \, \chi^j(x_1, \boldsymbol{k}_1)$$

$$\times \int d^2 \boldsymbol{k}_2 \, \bar{\chi}^k(x_2, \boldsymbol{k}_2) \left[I_{(a,i)} - N_c^2 \left(I_{(b,i)} + I_{(b,v)} \right) \right] H_{jk}$$
$$\times \delta^{(2)}(\boldsymbol{k}_1 + \boldsymbol{k}_2 - \boldsymbol{q}) + \ldots + \text{h.c.}, \tag{4.48}$$

where the ellipsis refers to the contribution from diagram (f,i) which we have not considered explicitly.

Let us now have a closer look at the integrals $I_{(b,i)}$ and $I_{(b,v)}$. Performing the integrations over ℓ_2^- and expanding in the regulator η_2 (following our regulator prescription in (4.11)), we obtain

$$I_{(b,i)} = \frac{i}{2} \int \frac{d\ell_1^+}{2\pi} \left[\theta(-\ell_1^+) - \theta(\ell_1^+) \right] \frac{\nu^{\eta_1} |\ell_1^+|^{-\eta_1}}{\ell_1^+ + i\epsilon} + \mathcal{O}(\eta_2), \tag{4.49}$$

$$I_{(b,v)} = -i \int \frac{d\ell_1^+}{2\pi} \, \theta(-\ell_1^+) \frac{\nu^{\eta_1} |\ell_1^+|^{-\eta_1}}{\ell_1^+ + i\epsilon} + \mathcal{O}(\eta_2). \tag{4.50}$$

Both $I_{(b,i)}$ and $I_{(b,v)}$ have poles in η_1, which disappear once they are summed together:

$$I_{(b,i)} + I_{(b,v)} = -\frac{i}{2} \int \frac{d\ell_1^+}{2\pi} \frac{\nu^{\eta_1} |\ell_1^+|^{-\eta_1}}{\ell_1^+ + i\epsilon} + \mathcal{O}(\eta_2) = I_{(a,i)}. \tag{4.51}$$

As expected, summing over all allowed cuts of graph (b), i.e. adding up diagrams (b,i) and (b,v), cancels out the final-state poles present in $I_{(b,i)}$ and $I_{(b,v)}$. We are now left with $I_{(a,i)}$ that has initial-state poles only, consistent with the formation of initial-state Wilson lines. This cancellation of final-state poles is a consequence of the unitarity property of our model. From the integral identification (4.51), it follows that

$$C_{(a)} \left[I_{(a,i)} - N_c^2 \left(I_{(b,i)} + I_{(b,v)} \right) \right] = \frac{1}{N_c} C_\Phi^2 \, I_{(a,i)}. \tag{4.52}$$

Hence, we see that in the $G_1 G_2$ region the different color structures completely disentangle, yielding the expected $1/N_c$ color factor.

4.5.2 Contributions from other regions

According to the Collins subtraction procedure, the full leading-power approximation of the dBM term is given by the sum of the various region contributions given in eqs. (4.6)–(4.9). In the previous subsection we calculated the first contribution $C_{G_1 G_2} \Gamma$, and, as it turns out, this will be the only nonvanishing contribution. In the following we will argue that (for our choice of rapidity regulators) the dBM term does not receive any contributions from the $C_1 G$, $G C_2$, and $C_1 C_2$ regions.

When the gluons have collinear scalings, more cuts are possible compared to the $G_1 G_2$ case. For the $C_1 C_2$ region, there are four possible cut positions for both diagrams (a) and (b). Besides (a,i) and (b,i), this region requires consideration of the six cut graphs given in figure 4.7, and, as before, we also need to take into account
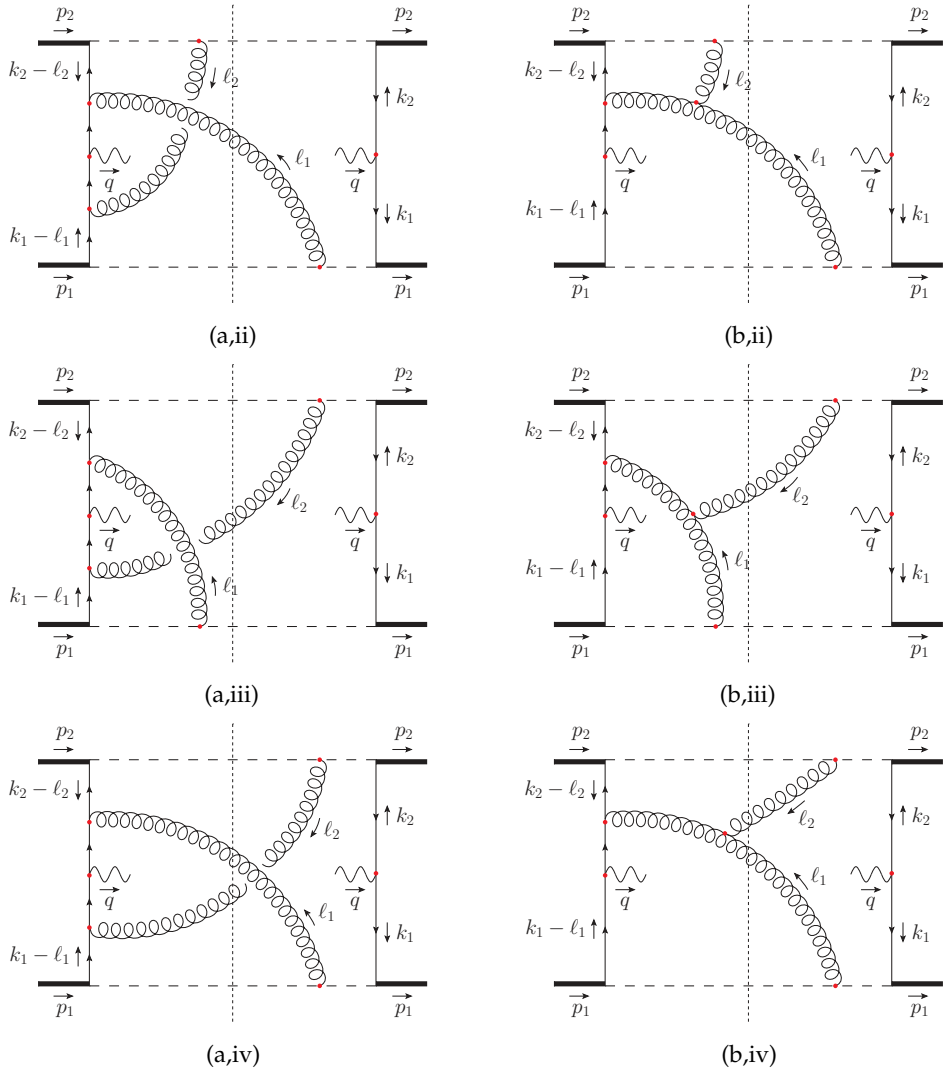
Figure 4.7: The color-entangled two-gluon exchange graphs that contribute to the dBM part of the DY cross section in the C_1C_2 region. This set is supplemented by diagrams (a,i) and (b,i) given in figure 4.5, and by graphs that can be obtained by $p \leftrightarrow \bar{p}$ or Hermitian conjugation.

their $p \leftrightarrow \bar{p}$ versions and Hermitian conjugates. Note that graph (b,v) is not allowed this time (the hard gluon line with momentum $\ell_1 + \ell_2$ is off shell).

The cross section contribution from the $C_1 C_2$ region is given by eq. (4.9). In this region, we can consider the naive graph terms and subtractions separately for fixed nonzero values of ℓ_1^+ and ℓ_2^-, as the subtractions remove the regions where ℓ_1^+ or ℓ_2^- are zero. This implies that we can ignore the $i\epsilon$ terms in the eikonal propagators (note that this is analogous to our treatment of the C_1 region contribution in section 4.4).

Let us consider the naive graph term $T_{C_1 C_2}\Gamma$ explicitly here. The starting expressions for diagrams (a,i) and (b,i) are given by eqs. (4.30) and (4.38). At leading power, the Dirac traces $R_{(a)jk;\mu\nu}$ and $R_{(b)jk;\mu\nu}$, contracted with the leptonic tensor $L^{\mu\nu}$, are now given by

$$R_{(a)jk;\mu\nu}L^{\mu\nu} = -\frac{2(\ell_1^+ - x_1 p_1^+)(\ell_2^- - x_2 p_2^-)}{\ell_1^+ x_2 p_2^-} R_{(b)jk;\mu\nu}L^{\mu\nu}, \tag{4.53}$$

$$R_{(b)jk;\mu\nu}L^{\mu\nu} = 16\, H_{jk}\, \ell_1^+ x_2 p_2^- \left[\ell_1^+ + 2(1-x_1)p_1^+\right]\left[\ell_2^- + 2(1-x_2)p_2^-\right], \tag{4.54}$$

so that eqs. (4.30) and (4.38) become

$$\frac{1}{2}\left(\frac{d\sigma_{\text{dBM}}}{d\Omega\, dx_1 dx_2\, d^2q}\right)_{(a,i)} = \frac{1}{2}\frac{\alpha^2}{q^4} e^2\, C_{(a)} \int d^2k_1 \int \frac{d\ell_1^+}{2\pi}\left[1 + \frac{2(1-x_1)p_1^+}{\ell_1^+}\right]$$

$$\times\, \chi^j(x_1, k_1, \ell_1^+) \int d^2k_2 \int \frac{d\ell_2^-}{2\pi}\left[1 + \frac{2(1-x_2)p_2^-}{\ell_2^-}\right]$$

$$\times\, \bar{\chi}^k(x_2, k_2, \ell_2^-)\, H_{jk}\, \delta^{(2)}(k_1 + k_2 - q), \tag{4.55}$$

$$\left(\frac{d\sigma_{\text{dBM}}}{d\Omega\, dx_1 dx_2\, d^2q}\right)_{(b,i)} = -\frac{1}{2}\frac{\alpha^2}{q^4} e^2\, C_{(b)} \int d^2k_1 \int \frac{d\ell_1^+}{2\pi}\left[1 + \frac{2(1-x_1)p_1^+}{\ell_1^+}\right]$$

$$\times\, \chi^j(x_1, k_1, \ell_1^+) \int d^2k_2 \int \frac{d\ell_2^-}{2\pi}\left[1 + \frac{2(1-x_2)p_2^-}{\ell_2^-}\right]$$

$$\times\, \bar{\chi}^k(x_2, k_2, \ell_2^-)\, H_{jk}\, \delta^{(2)}(k_1 + k_2 - q). \tag{4.56}$$

Summing these two contributions and using that $C_{(b)} = N_c^2\, C_{(a)}$, gives

$$\frac{1}{2}\left(\frac{d\sigma_{\text{dBM}}}{d\Omega\, dx_1 dx_2\, d^2q}\right)_{(a,i)} + \left(\frac{d\sigma_{\text{dBM}}}{d\Omega\, dx_1 dx_2\, d^2q}\right)_{(b,i)}$$

$$= \frac{1}{2}\frac{\alpha^2}{q^4} e^2 \left(1 - N_c^2\right) C_{(a)} \int d^2k_1 \int \frac{d\ell_1^+}{2\pi}\left[1 + \frac{2(1-x_1)p_1^+}{\ell_1^+}\right]\chi^j(x_1, k_1, \ell_1^+)$$

$$\times \int d^2k_2 \int \frac{d\ell_2^-}{2\pi}\left[1 + \frac{2(1-x_2)p_2^-}{\ell_2^-}\right]\bar{\chi}^k(x_2, k_2, \ell_2^-)\, H_{jk}\, \delta^{(2)}(k_1 + k_2 - q). \tag{4.57}$$

Note that since

$$\left(1 - N_c^2\right) C_{(a)} = \frac{1}{N_c} C_\Phi^2 , \tag{4.58}$$

we again obtain the standard $1/N_c$ color factor. This disentanglement of color is consistent with the expectations based on the non-Abelian Ward identity [30].

In eq. (4.57) we recognize the first term on the right-hand side of eq. (4.25). Recall from section 4.4 that the C_1 contribution to the quark BM function vanishes due to the cancellation between the amplitude and its conjugate (and likewise for the antiquark BM function). Hence, it follows that after properly taking into account the contributions from diagram (f,i) and all Hermitian conjugates, also eq. (4.57) vanishes. Naturally, we arrive at the same conclusion when considering the $p \leftrightarrow \bar{p}$ version of diagram (b,i).

In the preceding discussion on diagrams (a,i) and (b,i), the precise position of the cut was actually unimportant as the Dirac traces and color factors are independent of the cut. This means that the color disentanglement occurs in exactly the same way for the combination of diagrams (a,ii) and (b,ii), and for (a,iii) and (b,iii), etc. – in other words, the disentanglement proceeds in a *cut-by-cut* fashion. Again, these contributions are not only disentangled but in fact completely cancel out.

The subtraction terms can be dealt with in exactly the same way. Also here we find that the different color structures disentangle and ultimately vanish. Hence, the complete region contribution $C_{C_1 C_2} \Gamma$ to the dBM term is zero.

For a discussion on the $C_1 G$ and $G C_2$ regions we refer to [2]. As it turns out, also for these two regions the color completely disentangles for both the naive graph and subtraction terms separately. Once the Hermitian conjugate diagrams are included, these region contributions precisely cancel out as well. Hence, the dBM term can be completely ascribed to the $G_1 G_2$ region.

4.5.3 *The final result*

We have seen that with the regulators in (4.10)–(4.12), the cancellation of the color entanglement occurs on a region-by-region basis, and the full $\mathcal{O}(\alpha_s^2)$ contribution to the dBM term ends up in the $G_1 G_2$ region. With different regulators one can shift contributions between regions; the final result, however, obtained from the sum over all graphs and leading regions, remains the same. We refer to appendix B of [2] for examples of regulator choices for which the $G_1 G_2$ region does not give the full contribution.

We can now compare the result of our model calculation to the leading-order factorization formula (3.1). The sum of the diagrams in eq. (4.48), together with eq. (4.52), gives an expression consistent with taking the dBM part of eq. (3.1) and inserting the model expression for the quark BM function in eq. (4.22), as well as the analogous expression for the antiquark BM function. Diagram (f,i) and its conjugate precisely account for the 'cross terms', i.e. the combinations of the first/second term

in eq. (4.22) for the quark with the second/first term of the analogous expression for the antiquark. Upon summing all contributions, our factorized result is given by:[6]

$$\frac{d\sigma_{\mathrm{dBM}}}{d\Omega\,dx_1 dx_2\,d^2\boldsymbol{q}} = \frac{\alpha^2}{N_c\,q^4}\sum_q e_q^2 \int d^2\boldsymbol{k}_1 \frac{\widetilde{k}_{1T}^{j}}{M} h_{1,q}^{\perp}(x_1,\boldsymbol{k}_1^2) \int d^2\boldsymbol{k}_2 \frac{\widetilde{k}_{2T}^{k}}{M} \bar{h}_{1,q}^{\perp}(x_2,\boldsymbol{k}_2^2)$$

$$\times H_{jk}\,\delta^{(2)}(\boldsymbol{k}_1+\boldsymbol{k}_2-\boldsymbol{q})$$

$$= \frac{\alpha^2}{N_c\,q^2}\sum_q e_q^2\,B(\theta)\cos(2\phi)\,\mathcal{F}\left[w(\boldsymbol{k}_1,\boldsymbol{k}_2)\,h_1^{\perp}\bar{h}_1^{\perp}\right]. \tag{4.59}$$

Although we have worked in the CM frame of the $p\bar{p}$ collision, we have expressed our final result in terms of the CS angles [180].

We conclude that the dBM term in the DY cross section precisely factorizes as already anticipated by the CSS works; no loophole in their original proof for this double T-odd contribution is found. In contrast to the findings in [101], we do *not* find an additional color factor on top of the standard $1/N_c$ one. We have demonstrated that the entangled color structures are completely disentangled after summing over all relevant diagrams and cuts. The graph containing the three-gluon vertex, diagram (b) in figure 4.1, plays a vital role in this cancellation. This diagram was not taken into account in [101].

In our calculation it is possible to identify two mechanisms that underlie the cancellation of the color entanglement. The first of these is the unitarity cancellation of final-state poles after the sum over cuts of a particular diagram. This mechanism gives a finite result with initial-state poles in ℓ_1^+ and ℓ_2^- for diagram (b) after the sum over cuts. The second is the non-Abelian Ward identity. For each region, this identity ensures that when the diagrams are combined we end up with the standard $1/N_c$ color factor. We know from chapter 3 that these two principles are also essential in the all-order CSS factorization proof.

As mentioned already in subsection 4.3.2, the fact that our calculation agrees with the factorization formula that contains only TMDs (and a hard function) implies that the Glauber contributions may be absorbed into the TMDs. After the sum over cuts of diagrams (a), (b), and (f), the integrations over the gluon momentum components ℓ_1^+ and ℓ_2^- are no longer pinched in the Glauber region. At the same time, the integrations over ℓ_1^- and ℓ_2^+ are always pinched, even after the sum over cuts. In the case of diagram (b), the numerator structure appears to be responsible for this. Hence, analogous to our example of the active-spectator graph in subsection 3.4.2, we may deform ℓ_1 into the C_1 region while ℓ_2 may be deformed into the C_2 region. The fact that the Glauber contributions can be absorbed into other region contributions is consistent with the factorization proof. However, the CSS works on the deformation of soft momenta out of the Glauber region are not prescriptive about which momenta

---

6 Since $\bar{h}_{1,q}^{\perp}$ pertains to a left-moving hadron rather than a right-moving one, it comes with a minus sign relative to $h_{1,q}^{\perp}$ from swapping plus and minus indices in the antisymmetric Levi-Civita symbol (see e.g. [179]).

can be deformed into the collinear, and which into the central soft regions. With our calculation we have shown that certain Glauber momenta cannot be deformed into the central soft region, but instead must be deformed into a collinear region.

4.6 GENERALIZATIONS

In this final section of part I we comment on the scope of our analysis. Although, for definiteness, we have only considered collisions between a proton and an antiproton, our calculation of the dBM effect would proceed in exactly the same way for other hadron-hadron collisions (in an appropriately modified spectator model). In particular, the obtained disentanglement result would remain unchanged.

So far we have only discussed the dBM contribution to DY, even though according to [101] color entanglement is supposed to affect *all* double T-odd contributions. For collisions between spin-$\frac{1}{2}$ hadrons there is one other double T-odd contribution, namely the one that involves the quark and antiquark Sivers functions. The Sivers function $f_{1T}^{\perp}$ quantifies the correlation between the transverse momentum of the parton and the transverse spin of the parent hadron, see eq. (2.36). Hence, to probe the double Sivers effect, the hadron beams need to be transversely polarized [116,181]. It is straightforward to extend the analysis described in this chapter to include this particular contribution as well. Similarly to the dBM part of the cross section, our spectator model also gives a nonzero result for the double Sivers part at $\mathcal{O}(\alpha_s^2)$, coming from the same diagrams as discussed in subsection 4.3.1. Moreover, the BM and Sivers functions have been shown to be identical in this model [117,143,149,174, 178]. Via the same steps and mechanisms as for the dBM case, we find here no color entanglement, consistent with the CSS factorization works. Furthermore, by again choosing the regulators given in (4.10)–(4.12), also the double Sivers term completely originates from the $G_1 G_2$ momentum region.

As explained in subsection 4.5.3, our final results at $\mathcal{O}(\alpha_s^2)$ rely on the unitarity cancellation of final-state poles and the non-Abelian Ward identity. Although we cannot say anything about the possible appearance of color-entangled structures beyond the lowest order considered, we have no reason to doubt that these two mechanisms also drive the cancellation of the color entanglement at higher orders, in line with the arguments of CSS.

Finally, we do not expect the cancellation of the color entanglement to be restricted to the model studied, but rather to hold generally for DY in QCD. After all, the color factors involved in our calculation all appear exactly as in the analogous full QCD calculation. Furthermore, the key mechanisms that lead to our result are rather general field-theoretical principles and also play an essential role in the CSS proof.

4

Part II

DISTRIBUTIONS OF GLUONS INSIDE HADRONS

5

GLUON TMDS FOR HADRONS OF SPIN ≤ 1

This part of the thesis is about multidimensional gluon distributions. More knowl-
edge on gluon distributions could yield new insights into the internal dynamics of
hadrons. Fortunately, most of the formalism for quarks as introduced in chapter 2
can be straightforwardly extended to the gluon case. Whereas in chapter 6 we will
study gluon distributions at small x, here we will study gluon TMDs for general x.

For spin-$\frac{1}{2}$ hadrons, the first parametrization of the gluon-gluon TMD correlator
in terms of gluon TMDs was presented by Mulders and Rodrigues in 2000 in [182].
Here we extend their work to spin-1 hadrons about which little is known. Similarly
to the parametrization of the quark-quark TMD correlator in subsection 2.3.2, we will
parametrize the gluon-gluon TMD correlator for spin-1 hadrons in terms of leading-
twist gluon TMDs of definite rank; this is based on our work in [1]. Although the
TMDs that appear for tensor polarized hadrons are new, their collinear counterparts
were already introduced in 1989 in [183, 184]. Subsequently, in section 5.2, we will
derive positivity bounds for gluon TMDs for spin-0, spin-$\frac{1}{2}$, and spin-1 hadrons,
reflecting our work in [3]. These bounds are model-independent inequalities that
help relating and estimating the magnitude of the gluon TMDs about which very
little, or almost nothing, is currently known. In section 5.3 we discuss the pro-
cess dependence of gluon TMDs and in section 5.4 we briefly comment on some
experimental possibilities to study these functions.

5.1 PARAMETRIZATION OF THE GLUON-GLUON CORRELATION FUNCTION

In this section we provide a systematic and efficient parametrization of the gluon-
gluon correlation function for spin-1 hadrons. We start by giving a parametrization of
the unintegrated correlator and then we integrate over k^- to obtain a parametrization
of the light-front correlator in terms of TMDs.

5.1.1 *The unintegrated correlator*

The unintegrated gluon-gluon correlator for spin-1 hadrons is defined as

$$\Gamma^{[U,U']\,\mu\nu;\rho\sigma}(k, P, S, T, n)$$
$$\equiv \int \frac{d^4z}{(2\pi)^4}\, e^{ik\cdot z}\, \langle P; S, T| \operatorname{Tr}\left(F^{\mu\nu}(0)\, U_{[0,z]}\, F^{\rho\sigma}(z)\, U'_{[z,0]}\right) |P; S, T\rangle\,, \qquad (5.1)$$

where the trace reflects a summation over color degrees of freedom. Information on hadron spin has been incorporated in a way analogous to eq. (2.27). The gauge links $U_{[0,z]}$ and $U'_{[z,0]}$ guarantee color gauge invariance. Similarly to the quark case, these (calculable) gauge links give rise to a process dependence; this will be commented on in section 5.3. Even though without specifying a process the path integrations could run along arbitrary paths, we have already included a dependence on the four-vector n, which enters upon consideration of the staple-like gauge links defined in eq. (2.43).

Since the correlator in eq. (5.1) has open Lorentz indices, it should be parametrized in terms of Lorentz structures. Those structures need to be antisymmetric in μ, ν as well as in ρ, σ (this follows from the properties of the field strength tensor given in eq. (2.3)) and can be built from the Levi-Civita symbol ϵ, the metric tensor g, and the available vectors and tensors k, P, S, T (they are given in eqs. (2.9), (2.10), (2.24), and (2.25)), and n.[1] The different Lorentz structures will be labeled by the coefficient functions $A_i = A_i(k^2, k{\cdot}P, k{\cdot}n, P{\cdot}n)$ with $[A_i] = -4$, and the hadron mass M will be used to give the terms the correct mass dimension, consistent with $[\Gamma] = -2$. As the spin vector and spin tensor appear linearly in the spin density matrix in eq. (2.28), they can occur at most linearly in the parametrization. Similarly to the quark case, the correlator is constrained by requirements that arise from Hermiticity and parity. For any gauge link structure, these requirements are respectively given by

$$\Gamma^{\mu\nu;\rho\sigma}(k, P, S, T, n) = \Gamma^{\rho\sigma;\mu\nu*}(k, P, S, T, n), \tag{5.2}$$

$$\Gamma^{\mu\nu;\rho\sigma}(k, P, S, T, n) = \Gamma_{\mu\nu;\rho\sigma}(\bar{k}, \bar{P}, -\bar{S}, \bar{T}, \bar{n}). \tag{5.3}$$

The Hermiticity constraint requires the coefficient functions A_i to be real and the parity constraint forbids the presence of parity-odd terms (parity-odd structures include ϵ and S, which for example implies that S can only appear together with ϵ).

As in the case of quarks, time reversal transformations relate correlators with time-reversed gauge link structures and thus do not constrain the actual form of the parametrization. Considering staple-like gauge links only, the time reversal constraints are given by

$$\Gamma^{[\pm,\pm]\,\mu\nu;\rho\sigma*}(k, P, S, T, n) = \Gamma^{[\mp,\mp]}_{\mu\nu;\rho\sigma}(\bar{k}, \bar{P}, \bar{S}, \bar{T}, \bar{n}), \tag{5.4}$$

$$\Gamma^{[\pm,\mp]\,\mu\nu;\rho\sigma*}(k, P, S, T, n) = \Gamma^{[\mp,\pm]}_{\mu\nu;\rho\sigma}(\bar{k}, \bar{P}, \bar{S}, \bar{T}, \bar{n}), \tag{5.5}$$

where Hermitian conjugation is implied for the second gauge link entry. Similarly to eqs. (2.48) and (2.49), the T-even and T-odd parts of Γ are given by [185]:

$$\Gamma^{(\text{T-even})}(k, P, S, T, n) = \frac{1}{2}\left[\Gamma^{[+,+]}(k, P, S, T, n) + \Gamma^{[-,-]}(k, P, S, T, n)\right], \tag{5.6}$$

---

1 The four-vector n was not included in the original parametrization in [182]. Similarly to the quark case, the inclusion of n does not give rise to any new leading-twist TMDs, which we have checked explicitly for unpolarized hadrons. As we are ultimately only interested in leading-twist functions, we will simply omit terms containing n in the parametrization of the unintegrated correlator.

$$\Gamma'^{(\text{T-even})}(k, P, S, T, n) = \frac{1}{2} \left[\Gamma^{[+,-]}(k, P, S, T, n) + \Gamma^{[-,+]}(k, P, S, T, n) \right], \qquad (5.7)$$

$$\Gamma^{(\text{T-odd})}_{(f)}(k, P, S, T, n) = \frac{1}{2} \left[\Gamma^{[+,+]}(k, P, S, T, n) - \Gamma^{[-,-]}(k, P, S, T, n) \right], \qquad (5.8)$$

$$\Gamma^{(\text{T-odd})}_{(d)}(k, P, S, T, n) = \frac{1}{2} \left[\Gamma^{[+,-]}(k, P, S, T, n) - \Gamma^{[-,+]}(k, P, S, T, n) \right], \qquad (5.9)$$

where we suppressed the Lorentz indices to lighten the notation. The two T-odd parts are also sometimes referred to as f-type and d-type contributions, as they, respectively, are related to the SU(3) structure constants f and d defined in eq. (2.6).

For convenience, let us split the parametrization up according to the type of hadron polarization (unpolarized, vector polarized, or tensor polarized):

$$\Gamma^{\mu\nu;\rho\sigma}(k, P, S, T, n) = \Gamma^{\mu\nu;\rho\sigma}(k, P, n) + \Delta\Gamma^{\mu\nu;\rho\sigma}(k, P, S, n) + \Delta\Gamma^{\mu\nu;\rho\sigma}(k, P, T, n). \quad (5.10)$$

A possible parametrization of the first term (for which $S = T = 0$) is given by

$$\begin{aligned}
\Gamma^{\mu\nu;\rho\sigma}(k, P, n) &= M^2 A_1 \, \epsilon^{\mu\nu\alpha\beta} \epsilon^{\rho\sigma}{}_{\alpha\beta} + A_2 \, P^{[\mu} g^{\nu][\rho} P^{\sigma]} + A_3 \, k^{[\mu} g^{\nu][\rho} k^{\sigma]} \\
&\quad + (A_4 + iA_5) \, P^{[\mu} g^{\nu][\rho} k^{\sigma]} + (A_4 - iA_5) \, k^{[\mu} g^{\nu][\rho} P^{\sigma]} \\
&\quad + \frac{A_6}{M^2} \, P^{[\mu} k^{\nu]} P^{[\rho} k^{\sigma]},
\end{aligned} \qquad (5.11)$$

where square brackets denote antisymmetrization of the indices. The term with coefficient A_5 is T-odd.

The second term in eq. (5.10) represents the vector polarized case and can be parametrized as follows:

$$\begin{aligned}
\Delta\Gamma^{\mu\nu;\rho\sigma}(k, P, S, n) &\equiv \frac{1}{2} \left[\Gamma^{\mu\nu;\rho\sigma}(k, P, S, T, n) - \Gamma^{\mu\nu;\rho\sigma}(k, P, -S, T, n) \right] \\
&= -2MA_7 \, \epsilon^{\mu\nu\rho\sigma} k \cdot S + iMA_8 \left(\epsilon^{\mu\nu P[\rho} S^{\sigma]} - \epsilon^{\rho\sigma P[\mu} S^{\nu]} \right) \\
&\quad + iMA_9 \left(\epsilon^{\mu\nu S[\rho} P^{\sigma]} - \epsilon^{\rho\sigma S[\mu} P^{\nu]} \right) \\
&\quad + iMA_{10} \left(\epsilon^{\mu\nu k[\rho} S^{\sigma]} - \epsilon^{\rho\sigma k[\mu} S^{\nu]} \right) \\
&\quad + iMA_{11} \left(\epsilon^{\mu\nu S[\rho} k^{\sigma]} - \epsilon^{\rho\sigma S[\mu} k^{\nu]} \right) \\
&\quad + \frac{iA_{12}}{M} \left(\epsilon^{\mu\nu P[\rho} P^{\sigma]} - \epsilon^{\rho\sigma P[\mu} P^{\nu]} \right) k \cdot S \\
&\quad + \frac{iA_{13}}{M} \left(\epsilon^{\mu\nu k[\rho} k^{\sigma]} - \epsilon^{\rho\sigma k[\mu} k^{\nu]} \right) k \cdot S \\
&\quad + \frac{iA_{14}}{M} \left(\epsilon^{\mu\nu P[\rho} k^{\sigma]} - \epsilon^{\rho\sigma P[\mu} k^{\nu]} \right) k \cdot S \\
&\quad + \frac{iA_{15}}{M} \left(\epsilon^{\mu\nu k[\rho} P^{\sigma]} - \epsilon^{\rho\sigma k[\mu} P^{\nu]} \right) k \cdot S \\
&\quad + \frac{A_{16} + iA_{17}}{M} \, \epsilon^{\mu\nu PS} k^{[\rho} P^{\sigma]} + \frac{A_{16} - iA_{17}}{M} \, \epsilon^{\rho\sigma PS} k^{[\mu} P^{\nu]}
\end{aligned}$$

$$+ \frac{A_{18} + iA_{19}}{M} \, \epsilon^{\mu\nu k S} k^{[\rho} P^{\sigma]} + \frac{A_{18} - iA_{19}}{M} \, \epsilon^{\rho\sigma k S} k^{[\mu} P^{\nu]}$$

$$+ \frac{A_{20} + iA_{21}}{M} \, \epsilon^{\mu\nu k P} P^{[\rho} S^{\sigma]} + \frac{A_{20} - iA_{21}}{M} \, \epsilon^{\rho\sigma k P} P^{[\mu} S^{\nu]}$$

$$+ \frac{A_{22} + iA_{23}}{M} \, \epsilon^{\mu\nu k P} k^{[\rho} S^{\sigma]} + \frac{A_{22} - iA_{23}}{M} \, \epsilon^{\rho\sigma k P} k^{[\mu} S^{\nu]}$$

$$+ \frac{A_{24} + iA_{25}}{M^3} \, \epsilon^{\mu\nu k P} k^{[\rho} P^{\sigma]} \, k{\cdot}S$$

$$+ \frac{A_{24} - iA_{25}}{M^3} \, \epsilon^{\rho\sigma k P} k^{[\mu} P^{\nu]} \, k{\cdot}S. \tag{5.12}$$

The terms with coefficients A_7, A_{16}, A_{18}, A_{20}, A_{22}, and A_{24} are T-odd, and we note that the ones with coefficients A_8 up to A_{15} are slightly different from those in the original parametrization in [182].

Finally, for the third term in eq. (5.10), representing tensor polarized hadrons, a possible parametrization reads

$$\Delta\Gamma^{\mu\nu;\rho\sigma}(k,P,T,n) \equiv \frac{1}{2} \left[\Gamma^{\mu\nu;\rho\sigma}(k,P,S,T,n) - \Gamma^{\mu\nu;\rho\sigma}(k,P,S,-T,n) \right]$$

$$= A_{26} \, k^{[\mu} T^{\nu][\rho} k^{\sigma]} + A_{27} \, P^{[\mu} T^{\nu][\rho} P^{\sigma]} + (A_{28} + iA_{29}) \, k^{[\mu} T^{\nu][\rho} P^{\sigma]}$$

$$+ (A_{28} - iA_{29}) \, P^{[\mu} T^{\nu][\rho} k^{\sigma]} + \frac{A_{30} + iA_{31}}{M^2} \, k_\alpha T^{\alpha[\mu} k^{\nu]} k^{[\rho} P^{\sigma]}$$

$$+ \frac{A_{30} - iA_{31}}{M^2} \, k_\alpha T^{\alpha[\rho} k^{\sigma]} k^{[\mu} P^{\nu]} + \frac{A_{32} + iA_{33}}{M^2} \, k_\alpha T^{\alpha[\mu} P^{\nu]} k^{[\rho} P^{\sigma]}$$

$$+ \frac{A_{32} - iA_{33}}{M^2} \, k_\alpha T^{\alpha[\rho} P^{\sigma]} k^{[\mu} P^{\nu]} + M^2 A_{34} \left(g^{\mu[\rho} T^{\sigma]\nu} - g^{\nu[\rho} T^{\sigma]\mu} \right)$$

$$+ (A_{35} + iA_{36}) \, k_\alpha T^{\alpha[\mu} g^{\nu][\rho} k^{\sigma]} + (A_{35} - iA_{36}) \, k_\alpha T^{\alpha[\rho} g^{\sigma][\mu} k^{\nu]}$$

$$+ (A_{37} + iA_{38}) \, k_\alpha T^{\alpha[\mu} g^{\nu][\rho} P^{\sigma]} + (A_{37} - iA_{38}) \, k_\alpha T^{\alpha[\rho} g^{\sigma][\mu} P^{\nu]}$$

$$+ A_{39} \, k_\alpha k_\beta T^{\alpha\beta} \epsilon^{\mu\nu\kappa\lambda} \epsilon^{\rho\sigma}{}_{\kappa\lambda} + \frac{A_{40}}{M^2} \, k_\alpha k_\beta T^{\alpha\beta} P^{[\mu} g^{\nu][\rho} P^{\sigma]}$$

$$+ \frac{A_{41}}{M^2} \, k_\alpha k_\beta T^{\alpha\beta} k^{[\mu} g^{\nu][\rho} k^{\sigma]} + \frac{(A_{42} + iA_{43})}{M^2} \, k_\alpha k_\beta T^{\alpha\beta} P^{[\mu} g^{\nu][\rho} k^{\sigma]}$$

$$+ \frac{(A_{42} - iA_{43})}{M^2} \, k_\alpha k_\beta T^{\alpha\beta} k^{[\mu} g^{\nu][\rho} P^{\sigma]}$$

$$+ \frac{A_{44}}{M^4} \, k_\alpha k_\beta T^{\alpha\beta} P^{[\mu} k^{\nu]} P^{[\rho} k^{\sigma]}, \tag{5.13}$$

where the terms with coefficients A_{29}, A_{31}, A_{33}, A_{36}, A_{38}, and A_{43} are T-odd.

In the next subsection these results will be used to derive a parametrization of the light-front correlator in terms of gluon TMDs.

5

The gluon-gluon TMD correlator can be obtained from the unintegrated correlator in eq. (5.1) upon integration over k^-; it is given by

$$
\begin{aligned}
&\Gamma^{[U,U']\,\mu\nu;\rho\sigma}(x,\boldsymbol{k}) \\
&\equiv \frac{2}{P^+} \int \frac{dz^-\,d^2\boldsymbol{z}}{(2\pi)^3}\, e^{ik\cdot z}\, \langle P;S,T|\,\mathrm{Tr}\left(F^{\mu\nu}(0)\,U_{[0,z]}\,F^{\rho\sigma}(z)\,U'_{[z,0]}\right)|P;S,T\rangle\Bigg|_{z^+=0}\,, \quad (5.14)
\end{aligned}
$$

where we also included a conventional factor of $2/P^+$. Counting powers of the inverse hard scale relevant in the process leads to the definition of the leading-twist correlator

$$
\Gamma^{ij}(x,\boldsymbol{k}) \equiv \Gamma^{+i;+j}(x,\boldsymbol{k}), \quad (5.15)
$$

where i,j are transverse indices. This correlator was first parametrized in terms of gluon TMDs in [182], covering both unpolarized and vector polarized hadrons. A few years later, a different nomenclature for those TMDs was proposed in [175], in close analogy to the nomenclature for quarks. We adopt the conventions of the latter work and extend it to spin-1 hadrons. As for the quark case, we employ symmetric traceless tensors in k_T (see appendix B) to ensure the TMDs are of definite rank. Integrating the sum of eqs. (5.11), (5.12), and (5.13) over k^-, we obtain the following parametrization of the light-front correlator in eq. (5.15) in terms of leading-twist, definite-rank gluon TMDs:

$$
\begin{aligned}
\Gamma^{ij}(x,\boldsymbol{k}) = \Gamma_U^{ij}(x,\boldsymbol{k}) + \Gamma_L^{ij}(x,\boldsymbol{k}) + \Gamma_T^{ij}(x,\boldsymbol{k}) \\
+ \Gamma_{LL}^{ij}(x,\boldsymbol{k}) + \Gamma_{LT}^{ij}(x,\boldsymbol{k}) + \Gamma_{TT}^{ij}(x,\boldsymbol{k}),
\end{aligned} \quad (5.16)
$$

where

$$
\Gamma_U^{ij}(x,\boldsymbol{k}) = x\left[\delta_T^{ij}\,f_1(x,\boldsymbol{k}^2) + \frac{k_T^{ij}}{M^2}\,h_1^\perp(x,\boldsymbol{k}^2)\right], \quad (5.17)
$$

$$
\Gamma_L^{ij}(x,\boldsymbol{k}) = x\left[i\epsilon_T^{ij} S_L\,g_1(x,\boldsymbol{k}^2) + \frac{\epsilon_T^{\{i}{}_\alpha k_T^{j\}\alpha} S_L}{2M^2}\,h_{1L}^\perp(x,\boldsymbol{k}^2)\right], \quad (5.18)
$$

$$
\begin{aligned}
\Gamma_T^{ij}(x,\boldsymbol{k}) = x\Bigg[\frac{\delta_T^{ij}\,\epsilon_T^{S_T k_T}}{M}\,f_{1T}^\perp(x,\boldsymbol{k}^2) + \frac{i\epsilon_T^{ij}\,\boldsymbol{k}\cdot\boldsymbol{S}_T}{M}\,g_{1T}(x,\boldsymbol{k}^2) \\
- \frac{\epsilon_T^{k_T\{i}S_T^{j\}} + \epsilon_T^{S_T\{i}k_T^{j\}}}{4M}\,h_1(x,\boldsymbol{k}^2) - \frac{\epsilon_T^{\{i}{}_\alpha k_T^{j\}\alpha S_T}}{2M^3}\,h_{1T}^\perp(x,\boldsymbol{k}^2)\Bigg], \quad (5.19)
\end{aligned}
$$

$$
\Gamma_{LL}^{ij}(x,\boldsymbol{k}) = x\left[\delta_T^{ij} S_{LL}\,f_{1LL}(x,\boldsymbol{k}^2) + \frac{k_T^{ij} S_{LL}}{M^2}\,h_{1LL}^\perp(x,\boldsymbol{k}^2)\right], \quad (5.20)
$$

$$\Gamma_{LT}^{ij}(x, \mathbf{k}) = x \left[\frac{\delta_T^{ij} \, \mathbf{k} \cdot \mathbf{S}_{LT}}{M} f_{1LT}(x, \mathbf{k}^2) + \frac{i\epsilon_T^{ij} \epsilon_T^{S_{LT}k_T}}{M} g_{1LT}(x, \mathbf{k}^2) \right.$$

$$\left. + \frac{S_{LT}^{\{i} k_T^{j\}}}{M} h_{1LT}(x, \mathbf{k}^2) + \frac{k_T^{ij\alpha} S_{LT\alpha}}{M^3} h_{1LT}^{\perp}(x, \mathbf{k}^2) \right], \tag{5.21}$$

$$\Gamma_{TT}^{ij}(x, \mathbf{k}) = x \left[\frac{\delta_T^{ij} k_T^{\alpha\beta} S_{TT\alpha\beta}}{M^2} f_{1TT}(x, \mathbf{k}^2) + \frac{i\epsilon_T^{ij} \epsilon_{T\gamma}^{\beta} k_T^{\gamma\alpha} S_{TT\alpha\beta}}{M^2} g_{1TT}(x, \mathbf{k}^2) \right.$$

$$+ S_{TT}^{ij} h_{1TT}(x, \mathbf{k}^2) + \frac{S_{TT\alpha}^{\{i} k_T^{j\}\alpha}}{M^2} h_{1TT}^{\perp}(x, \mathbf{k}^2)$$

$$\left. + \frac{k_T^{ij\alpha\beta} S_{TT\alpha\beta}}{M^4} h_{1TT}^{\perp\perp}(x, \mathbf{k}^2) \right]. \tag{5.22}$$

The expressions of the TMDs in terms of the coefficient functions A_i can be found in appendix D, and the equivalent parametrization in coordinate space is given in appendix E.

	δ_T^{ij}	$i\epsilon_T^{ij}$	k_T^i, k_T^{ij}, etc.
U	f_1		$h_1^{\perp}$
L		g_1	$h_{1L}^{\perp}$
T	$f_{1T}^{\perp}$	g_{1T}	h_1, $h_{1T}^{\perp}$
LL	f_{1LL}		$h_{1LL}^{\perp}$
LT	f_{1LT}	g_{1LT}	h_{1LT}, $h_{1LT}^{\perp}$
TT	f_{1TT}	g_{1TT}	h_{1TT}, $h_{1TT}^{\perp}$, $h_{1TT}^{\perp\perp}$

Table 5.1: An overview of the leading-twist gluon TMDs for unpolarized (U), vector polarized (L or T), and tensor polarized (LL, LT, or TT) hadrons. The functions indicated in boldface also occur as collinear PDFs, and the ones in red are T-odd. The Lorentz structures δ_T^{ij}, $i\epsilon_T^{ij}$, and k_T^i, k_T^{ij}, etc. correspond to unpolarized, circularly polarized, and linearly polarized gluons, respectively.

In table 5.1 we have organized the gluon TMDs appearing in eq. (5.16) by hadron and gluon polarizations and we have indicated which functions are T-odd and which ones have a collinear counterpart (the rank-0 TMDs). The gluon polarizations are projected out by the Lorentz structures that come with the functions (analogous to

the Dirac structures in the quark case); we use the letters f, g, and h to indicate unpolarized, circularly polarized, and linearly polarized gluons, respectively.

With respect to the nomenclature in [175], we have introduced the definite-rank function[2]

$$h_1 \equiv h_{1T} + h_{1T}^{\perp(1)}. \tag{5.23}$$

Furthermore, we have defined $g_1 \equiv g_{1L}$, the gluon helicity TMD. Unlike the quark TMDs, most of the gluon TMDs do not go by a special name. The most well-known T-odd gluon function is the gluon Sivers TMD $f_{1T}^{\perp}$. We note that the collinear PDF $h_{1TT}(x)$ appears in the structure function $\Delta(x)$ defined in [183] (this structure function is called $\Delta_2 G(x)$ in [184]).

5.2 POSITIVITY BOUNDS

In this section we derive bounds for the gluon TMDs that appear in eq. (5.16). Positivity bounds for gluon TMDs were derived in [182] for spin-$\frac{1}{2}$ hadrons and, by applying the same strategy, we extend their analysis to spin-1 hadrons. For completeness, we will also cover the bounds that apply to spin-$\frac{1}{2}$ hadrons, completing the study of [182] where T-odd functions were not included. We will also mention the spin-0 case. Bounds on collinear gluon PDFs for spin-1 hadrons can be found in [3, 186].

Similarly to the quark-quark correlator in eq. (2.27), the gluon-gluon correlator in eq. (5.15) is defined as

$$\Gamma^{ij}(x, \mathbf{k}, S, T) \equiv \mathrm{Tr}\left(\rho(S, T) N^{ij}(x, \mathbf{k})\right) = \rho_{s's}(S, T) N_{ss'}^{ij}(x, \mathbf{k}), \tag{5.24}$$

where s, s' label the hadronic polarization states. Our starting point is the idea that the correlator Γ can be seen as a 2×2 matrix in the two gluon polarizations, and that the quantity N can be regarded as a 6×6 matrix in gluon $\otimes$ hadron spin space. Choosing the same basis as in [182], namely given by the two circular gluon polarizations $|\pm\rangle = \mp \frac{1}{\sqrt{2}} (|1\rangle \pm i |2\rangle)$, we find that N is given by

$$N = x \begin{pmatrix} A & B \\ B^\dagger & C \end{pmatrix}, \tag{5.25}$$

where

$$A = \begin{pmatrix} f_1 + \frac{f_{1LL}}{2} - g_1 & \frac{|\mathbf{k}|e^{-i\phi}}{\sqrt{2}M}\left(\tilde{f} - \tilde{g} + h_{1LT}\right) & \frac{k^2 e^{-2i\phi}}{M^2}\left(f_{1TT} + ig_{1TT} - h_{1TT}^\perp\right) \\ \frac{|\mathbf{k}|e^{i\phi}}{\sqrt{2}M}\left(\tilde{f}^* - \tilde{g}^* + h_{1LT}\right) & f_1 - f_{1LL} & -\frac{|\mathbf{k}|e^{-i\phi}}{\sqrt{2}M}\left(\tilde{f}^* + \tilde{g}^* + h_{1LT}\right) \\ \frac{k^2 e^{2i\phi}}{M^2}\left(f_{1TT} - ig_{1TT} - h_{1TT}^\perp\right) & -\frac{|\mathbf{k}|e^{i\phi}}{\sqrt{2}M}\left(\tilde{f} + \tilde{g} + h_{1LT}\right) & f_1 + \frac{f_{1LL}}{2} + g_1 \end{pmatrix},$$

---

[2] Despite the notational similarities, this function is quite different from the quark transversity function h_1 defined in eq. (2.36).

$$
B = \begin{pmatrix}
-\frac{k^2 e^{-2i\phi}}{4M^2}\left(2h_1^\perp + h_{1LL}^\perp - 2ih_{1L}^\perp\right) & \frac{|k|^3 e^{-3i\phi}}{2\sqrt{2}M^3}\left(h_{1LT}^\perp + ih_{1T}^\perp\right) & -\frac{k^4 e^{-4i\phi}}{2M^4}\,h_{1TT}^{\perp\perp} \\[2mm]
-\frac{|k|e^{-i\phi}}{\sqrt{2}M}\left(2h_{1LT} - ih_1\right) & -\frac{k^2 e^{-2i\phi}}{2M^2}\left(h_1^\perp - h_{1LL}^\perp\right) & -\frac{|k|^3 e^{-3i\phi}}{2\sqrt{2}M^3}\left(h_{1LT}^\perp - ih_{1T}^\perp\right) \\[2mm]
-2h_{1TT} & \frac{|k|e^{-i\phi}}{\sqrt{2}M}\left(2h_{1LT} + ih_1\right) & -\frac{k^2 e^{-2i\phi}}{4M^2}\left(2h_1^\perp + h_{1LL}^\perp + 2ih_{1L}^\perp\right)
\end{pmatrix},
$$

$$
C = \begin{pmatrix}
f_1 + \frac{f_{1LL}}{2} + g_1 & \frac{|k|e^{-i\phi}}{\sqrt{2}M}\left(\tilde{f} + \tilde{g} + h_{1LT}\right) & \frac{k^2 e^{-2i\phi}}{M^2}\left(f_{1TT} - ig_{1TT} - h_{1TT}^\perp\right) \\[2mm]
\frac{|k|e^{i\phi}}{\sqrt{2}M}\left(\tilde{f}^* + \tilde{g}^* + h_{1LT}\right) & f_1 - f_{1LL} & -\frac{|k|e^{-i\phi}}{\sqrt{2}M}\left(\tilde{f}^* - \tilde{g}^* + h_{1LT}\right) \\[2mm]
\frac{k^2 e^{2i\phi}}{M^2}\left(f_{1TT} + ig_{1TT} - h_{1TT}^\perp\right) & -\frac{|k|e^{i\phi}}{\sqrt{2}M}\left(\tilde{f} - \tilde{g} + h_{1LT}\right) & f_1 + \frac{f_{1LL}}{2} - g_1
\end{pmatrix}.
$$

For convenience we have defined $\tilde{f} \equiv f_{1LT} + if_{1T}^\perp$ and $\tilde{g} \equiv g_{1T} + ig_{1LT}$. Furthermore, we have expressed k in terms of its polar coordinates $|k|$ and ϕ. From symmetry considerations it follows that blocks A and C are related by parity and that the off-diagonal blocks are Hermitian conjugates.

To make more apparent the properties of the matrix N, we write its elements in the following form:

$$
\begin{aligned}
N_{ss'}^{ij}(x, k) &\equiv \frac{2}{P^+}\int \frac{dz^-\, d^2z}{(2\pi)^3}\, e^{ik\cdot z}\, \langle P; s|\, \mathrm{Tr}\left(F^{+i}(0)\, F^{+j}(z)\right)|P; s'\rangle\Big|_{z^+=0} \\
&= \frac{2}{P^+}\sum_m \mathrm{Tr}\left(\langle P_m|\, F^{ni}(0)\,|P; s\rangle^*\, \langle P_m|\, F^{nj}(0)\,|P; s'\rangle\right) \\
&\quad \times\, \delta\left(P_m^+ - (1-x)P^+\right)\, \delta^{(2)}\left(P_m + k\right),
\end{aligned}
\tag{5.26}
$$

where we inserted a complete set of momentum eigenstates $\{|P_m\rangle\}$. From eq. (5.26) we infer that, for any basis, the diagonal elements are given by absolute squares. In particular, it follows that the eigenvalues of N in eq. (5.25) must be ≥ 0, or, equivalently, that N is *positive semidefinite*. This property can be used to set constraints on the TMDs. Given the limited amount of information we have on the gluon functions, we will refrain from diagonalizing the full 6×6 matrix and rather restrict ourselves to finding the eigenvalues of its 2×2 principal minors. Due to the symmetry properties of N, some minors yield the same bounds; we obtain the following nine inequalities for the spin-1 case:

$$
\frac{k^2}{2M^2}\,|h_1^\perp - h_{1LL}^\perp| \leq f_1 - f_{1LL}\,,
\tag{5.27}
$$

$$
\frac{k^4}{16M^4}\left[4(h_{1L}^\perp)^2 + (2h_1^\perp + h_{1LL}^\perp)^2\right] \leq f^+ f^-,
\tag{5.28}
$$

$$
\frac{k^2}{2M^2}\left(h_1^2 + 4h_{1LT}^2\right) \leq (f_1 - f_{1LL})\, f^+,
\tag{5.29}
$$

$$
\frac{k^6}{8M^6}\left[(h_{1T}^\perp)^2 + (h_{1LT}^\perp)^2\right] \leq (f_1 - f_{1LL})\, f^-,
\tag{5.30}
$$

$$
\frac{k^2}{2M^2}\left[(f_{1T}^\perp + g_{1LT})^2 + (f_{1LT} + g_{1T} + h_{1LT})^2\right] \leq (f_1 - f_{1LL})\, f^+,
\tag{5.31}
$$

$$\frac{k^2}{2M^2}\left[(f_{1T}^{\perp} - g_{1LT})^2 + (f_{1LT} - g_{1T} + h_{1LT})^2\right] \le (f_1 - f_{1LL})f^-, \qquad (5.32)$$

$$|h_{1TT}| \le \frac{f^+}{2}, \qquad (5.33)$$

$$\frac{k^4}{2M^4}|h_{1TT}^{\perp\perp}| \le f^-, \qquad (5.34)$$

$$\frac{k^4}{M^4}\left[g_{1TT}^2 + (f_{1TT} - h_{1TT}^{\perp})^2\right] \le f^+f^-, \qquad (5.35)$$

where for convenience we have defined $f^{\pm} \equiv f_1 + \frac{f_{1LL}}{2} \pm g_1$.

Now let us consider spin-$\frac{1}{2}$ hadrons. Bounds for this case have been derived already in [182], however using a different notation and leaving the T-odd TMDs aside. The parametrization of the correlator for a spin-$\frac{1}{2}$ hadron is given by the sum of eqs. (5.17)–(5.19), and the spin density matrix can be parametrized in terms of the spin vector only (using the identity matrix and the three Pauli matrices) and is a 2×2 matrix in hadron spin space. The quantity N is now a 4×4 matrix in gluon $\otimes$ hadron spin space and its explicit form (that *does* contain the T-odd functions) is given in [182]. From that matrix we can extract the following bounds from its 2×2 principal minors:

$$|g_1| \le f_1, \qquad (5.36)$$

$$\frac{k^4}{4M^4}\left[(h_{1L}^{\perp})^2 + (h_1^{\perp})^2\right] \le (f_1 + g_1)(f_1 - g_1), \qquad (5.37)$$

$$\frac{|k|}{M}|h_1| \le f_1 + g_1, \qquad (5.38)$$

$$\frac{|k|^3}{2M^3}|h_{1T}^{\perp}| \le f_1 - g_1, \qquad (5.39)$$

$$\frac{k^2}{M^2}\left[(f_{1T}^{\perp})^2 + g_{1T}^2\right] \le (f_1 + g_1)(f_1 - g_1). \qquad (5.40)$$

Upon omitting the functions related to tensor polarization in bounds (5.27)–(5.35), which is mathematically equivalent to considering the spin-$\frac{1}{2}$ case, one obtains a set of bounds that is less strict than (but consistent with) bounds (5.36)–(5.40). In general, bounds can be sharpened upon considering the eigenvalues of higher-dimensional principal minors.

Finally, let us comment on the spin-0 case. For spin-0 hadrons the parametrization of the correlator is given by the two functions in eq. (5.17). Now, as there is no spin density matrix (this means that $\Gamma = N$), we simply need to compute the eigenvalues of the 2×2 matrix Γ; this results in the single bound

$$\frac{k^2}{2M^2}|h_1^{\perp}| \le f_1. \qquad (5.41)$$

This inequality also follows from bounds (5.36)–(5.40) upon omitting the functions related to vector polarization.

Note that in eq. (5.26) we did not consider the process-dependent gauge link structure explicitly. In fact, the inequalities (5.27)–(5.41) do *not* hold generally true for any correlator – the matrix N is positive semidefinite only for field combinations, including gauge links, that 'factorize' into the form $O^\dagger(0)\,O(z)$. The simplest gauge link structures for which this holds are $[+,+]$, $[-,-]$, $[+,-]$, and $[-,+]$ (we will say more about these link structures in the next section). For the same links, one also has the constraint $f_1 \geq 0$ (note that for the spin-$\frac{1}{2}$ and spin-0 cases this follows already from the provided bounds).

One might worry about the effects of QCD evolution on the validity of the bounds. In the collinear case, the so-called Soffer bound involving three quark functions [187] has been shown to be preserved up to next-to-leading order accuracy [188–190]. However, to our knowledge, there are no studies yet on the stability of bounds under evolution concerning TMDs. The fact that the evolution kernel for TMDs is independent of spin [191, 192] might suggest that, in the appropriate regime of k where TMD factorization is valid, positivity bounds are respected also in this case. This topic remains open to further investigation.

Since the bounds follow from the positive semidefiniteness of N and rely primarily on the operator structure of the correlator, they can be considered as tests of the TMD formalism (provided the functions are compared at the same scales). In practical situations, however, the bounds will, rather than serving as tests of the framework, often be more useful as a check in model or lattice calculations, or as a way to obtain an order of magnitude estimate of certain TMDs. The latter is commonly done by saturating the bounds. Those estimates could serve, for instance, as input for an estimate of measurements of particular azimuthal asymmetries.

5.3 PROCESS DEPENDENCE

In general, gluon TMDs depend on the process under consideration through the gauge links. As explained in subsection 2.3.3, the integration paths of these links depend on the direction of color flow in the hard scattering process and are calculable. The four simplest and most common gauge links in the gluon case are $[+,+]$, $[-,-]$, $[+,-]$, and $[-,+]$, see figure 5.1. The $[-,-]$ gauge link appears in processes with color flow annihilated within the initial state, such as the (gluonic) DY process or Higgs production through gluon fusion ($gg \rightarrow h$) [192, 193]. The structure $[+,+]$, on the other hand, is related to color flow into the final state, which is the case for e.g. $q\bar{q}$ production in SIDIS [194]. When color flow involves both initial and final states, the gauge links $[+,-]$ and $[-,+]$ appear, which is for instance the case in processes with $qg \rightarrow qg$ and $\bar{q}g \rightarrow \bar{q}g$ partonic contributions, respectively [72].

The relations between gluon TMDs that feature different gauge link structures have been systematically studied in [79]. For example, for the T-even unpolarized function f_1 and the T-odd gluon Sivers function $f_{1T}^\perp$ we have the following relations [72,79,195]:

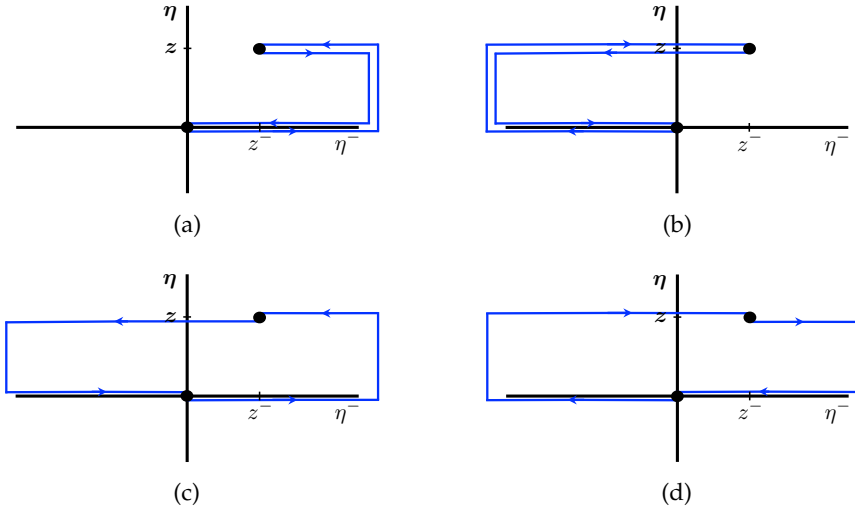$$f_1^{[+,+]} = f_1^{[-,-]}, \qquad f_1^{[+,-]} = f_1^{[-,+]}, \tag{5.42}$$

Figure 5.1: The integration paths of (a) the $[+,+]$ gauge link, (b) the $[-,-]$ gauge link, (c) the $[+,-]$ gauge link, and (d) the $[-,+]$ gauge link. The paths run along the light-front $\eta^+ = 0$ via plus and/or minus light-cone infinity.

$$f_{1T}^{\perp[+,+]} = -f_{1T}^{\perp[-,-]}, \qquad f_{1T}^{\perp[+,-]} = -f_{1T}^{\perp[-,+]}. \tag{5.43}$$

As it turns out, $f_1^{[+,+]}$ *cannot* be related to $f_1^{[+,-]}$, and likewise for $f_{1T}^\perp$. Thus, for some processes TMDs are completely independent from each other and encode distinct information. For this reason, different extractions could in fact be *complementary* [196].

In the context of small-x studies, the unpolarized distributions $f_1^{[+,+]}$ and $f_1^{[+,-]}$ are often respectively referred to as the (unintegrated) Weizsäcker-Williams (WW) and dipole distributions [197–199]. These distributions are considered as the two fundamental gluon distributions at small x. Although in some relatively complicated processes such as dijet production in pA collisions distributions with other gauge link structures can appear, it seems that in the large-N_c limit they can all be related to the WW and dipole distributions [199, 200]. We will refer to the $[+,+]$ and $[-,-]$ links as 'WW-type' gauge link structures, and to the $[+,-]$ and $[-,+]$ links as 'dipole-type' gauge link structures (irrespective of gluon or hadron polarization and for general x).

5.4 EXPERIMENTAL POSSIBILITIES

Over recent years, interest in gluon distributions has increased significantly. However, to this day, the extraction of gluon functions remains very challenging compared to quark functions, for example because they typically require higher-energy scattering

processes. We end this chapter with a brief discussion on experimental possibilities to study gluon TMDs.

The gluon TMDs for unpolarized and vector polarized hadrons could be investigated through processes at the Relativistic Heavy Ion Collider (RHIC), at the Large Hadron Collider (LHC), possibly at a future polarized fixed-target experiment at the LHC called AFTER@LHC [201], and at a future EIC [93]. In fact, recently, using LHCb data on $pp \rightarrow J/\Psi \, J/\Psi \, X$ and assuming a Gaussian profile, the transverse momentum dependence of the unpolarized gluon distribution $f_1^{[+,+]}$ was constrained for the first time [202]. As yet, there are no fits to data available for the gluon Sivers function $f_{1T}^{\perp}$. The most promising processes that give access to $f_{1T}^{\perp}$ are $p^{\uparrow}p \rightarrow \gamma \, \text{jet} \, X$ (dipole-type) at RHIC and at AFTER@LHC [203,204], $p^{\uparrow}p \rightarrow J/\Psi \, \gamma \, X$ (WW-type) [205] or $p^{\uparrow}p \rightarrow J/\Psi \, J/\Psi \, X$ (WW-type) at AFTER@LHC [206,207], and $e^{-}p^{\uparrow} \rightarrow e^{-}D\bar{D}X$ (WW-type) at an EIC [196], where in between brackets we indicated for each process the (dominant) gauge link structure that is probed. Today, no all-order TMD factorization proofs are available for these processes, which all involve at least three hadrons. For the same reason that the expected sign change between DY and SIDIS for the quark Sivers function is currently under investigation, it would be interesting from a theoretical point of view to compare extractions of the gluon Sivers function from different processes. The current knowledge on the gluon Sivers function as well as future prospects are reviewed in [196].

The study of gluon TMDs for tensor polarized hadrons would be possible at the experiments proposed to investigate polarized deuterons, e.g. at the EIC possibility at Jefferson Lab (JLEIC) [93,208,209], or at COMPASS at CERN [210]. Some experimental interest has been shown already in the collinear PDF $h_{1TT}(x)$ that represents the distribution of linearly polarized gluons in a transversely tensor polarized hadron. Its extraction has been proposed to occur at Jefferson Lab using nitrogen targets [211], and could also be achieved within the program of the EIC [91].

For more information on the phenomenology of gluon TMDs, see e.g. [93,95,194, 199,212–218].

GLUON DISTRIBUTIONS AT SMALL x

In this chapter we will discuss dipole-type gluon TMDs, GTMDs, and Wigner distributions in the small-x region where hadrons mainly consist of gluons. The field of research concerned with small-x physics is also referred to as high-energy QCD, for a high CM energy is required in collisions to probe small values of x. Most of this chapter is based on our work in [1,3,4].

First, we provide a parametrization of the gluon-gluon GTMD correlator for unpolarized hadrons in terms of four leading-twist gluon GTMDs. Then we show that for the dipole-type gauge link structure the GTMD correlator simplifies in the small-x limit to the Fourier transform of a hadronic matrix element containing a rectangular Wilson loop. This so-called off-forward Wilson loop correlator can be parametrized in terms of a single function (for unpolarized hadrons), which implies that at small x there is only one independent dipole-type GTMD and not four. This Wilson loop GTMD contains both so-called pomeron and odderon contributions. The same strategy can be applied to the TMD case. As it turns out, in this analysis the small-x limit is independent of hadron polarization, which means that the obtained results for the unpolarized case in fact apply to each individual hadronic polarization state. As a result, we obtain a greatly simplified picture for the TMDs that were introduced in subsection 5.1.2 for spin-1 hadrons. In the TMD case, the only odderon contribution that we identify originates from transversely polarized hadrons, consistent with the findings in [219]. We will also derive positivity bounds for TMDs that apply to the small-x region.

In the second part of this chapter we will study azimuthal correlations in hadron production from high-energy collisions that originate from dipole Wigner distributions of gluons in the nucleus. In section 6.3 we introduce the so-called McLerran-Venugopalan (MV) model (and an extension thereof) that will be used to calculate the expectation values of the pomeron and odderon operators in a nuclear environment. Those results will feature in the calculation of pomeron and odderon Wigner distributions that contribute to the cross section of dihadron production in pA collisions. We show that the odderon Wigner distribution can generate odd harmonics in the two-particle azimuthal correlations and we will provide a numerical estimate of the size of the first odd-harmonic contribution, quantified by the so-called directed flow coefficient.

Analogous to the quark-quark GTMD correlator in eq. (2.21), the gluon-gluon GTMD correlator for an unpolarized hadron is defined as [220]:

$$
\begin{aligned}
&G^{[U,U']\,\mu\nu;\rho\sigma}(x,\boldsymbol{k},\xi,\boldsymbol{\Delta}) \\
&\equiv \frac{2}{P^+} \int \frac{dz^-\,d^2\boldsymbol{z}}{(2\pi)^3}\, e^{ik\cdot z}\, \langle p'|\,\mathrm{Tr}\left(F^{\mu\nu}\!\left(-\tfrac{z}{2}\right) U_{\left[-\frac{z}{2},\frac{z}{2}\right]}\, F^{\rho\sigma}\!\left(\tfrac{z}{2}\right) U'_{\left[\frac{z}{2},-\frac{z}{2}\right]}\right)|p\rangle\Big|_{z^+=0}.
\end{aligned}
\tag{6.1}
$$

The off-forwardness $\Delta \equiv p' - p$ is given in eq. (2.22). Similarly to eq. (5.15), the leading-twist correlator is given by

$$
G^{ij}(x,\boldsymbol{k},\xi,\boldsymbol{\Delta}) \equiv G^{+i;+j}(x,\boldsymbol{k},\xi,\boldsymbol{\Delta}),
\tag{6.2}
$$

where i,j are transverse indices. This correlator can be parametrized in terms of GTMDs, in a way analogous to the parametrization of the TMD correlator in terms of TMDs in chapter 5. This has been done already in [220] based on the light-front formalism (in fact, up to all powers in the inverse hard scale and including vector polarization) and here we present an alternative, but equivalent, parametrization. To ensure that the distribution functions are of definite rank, we use symmetric traceless tensors in both k_T and Δ_T. As explained in [220], requirements from Hermiticity and time reversal (unlike parity) do not affect the Lorentz structure of the parametrization – they rather impose constraints on the GTMDs. A possible parametrization of the correlator in eq. (6.2) is given by[1]

$$
\begin{aligned}
G^{ij}(x,\boldsymbol{k},\xi,\boldsymbol{\Delta}) = x\,\Bigg[&\delta_T^{ij}\,\mathcal{F}_1(x,\boldsymbol{k}^2,\xi,\boldsymbol{\Delta}^2,\boldsymbol{k}\cdot\boldsymbol{\Delta}) + \frac{k_T^{ij}}{M^2}\,\mathcal{F}_2(x,\boldsymbol{k}^2,\xi,\boldsymbol{\Delta}^2,\boldsymbol{k}\cdot\boldsymbol{\Delta}) \\
&+ \frac{\Delta_T^{ij}}{M^2}\,\mathcal{F}_3(x,\boldsymbol{k}^2,\xi,\boldsymbol{\Delta}^2,\boldsymbol{k}\cdot\boldsymbol{\Delta}) + \frac{k_T^{[i}\Delta_T^{j]}}{M^2}\,\mathcal{F}_4(x,\boldsymbol{k}^2,\xi,\boldsymbol{\Delta}^2,\boldsymbol{k}\cdot\boldsymbol{\Delta})\Bigg],
\end{aligned}
\tag{6.3}
$$

where the functions $\mathcal{F}_{1-4}$ are (complex-valued) GTMDs. In contrast to the TMD case, no positivity bounds exist for GTMDs. After all, the GTMD analogue of N in eq. (5.26) is *not* positive semidefinite due to Δ being nonzero.

In the forward limit, the GTMDs in eq. (6.3) are related to the TMDs for unpolarized hadrons in eq. (5.17); to be specific:

$$
\lim_{\Delta\to0}\mathcal{F}_1 = f_1\,, \qquad \lim_{\Delta\to0}\mathcal{F}_2 = h_1^\perp\,, \qquad \lim_{\Delta\to0}\mathcal{F}_3 = \lim_{\Delta\to0}\mathcal{F}_4 = 0.
\tag{6.4}
$$

Upon integration over $\boldsymbol{k}$, one recovers from the GTMDs the leading-twist GPDs [175].

---

[1] In principle one could also have a function that comes with the symmetric and traceless Lorentz structure $k_T^{\{i}\Delta_T^{j\}} + (\boldsymbol{k}\cdot\boldsymbol{\Delta})\,g_T^{ij}$. However, this function would not be independent from the other ones; more specifically, it could be eliminated by suitable redefinitions of the functions $\mathcal{F}_2$ and $\mathcal{F}_3$ through the (two-dimensional) relation $\boldsymbol{k}\cdot\boldsymbol{\Delta}\left[k_T^{\{i}\Delta_T^{j\}} + (\boldsymbol{k}\cdot\boldsymbol{\Delta})\,g_T^{ij}\right] = \boldsymbol{\Delta}^2 k_T^{ij} + \boldsymbol{k}^2\Delta_T^{ij}$.

The parametrization in eq. (6.3) holds for general x and for any gauge link structure. In this section we will study the small-x limit of the correlator for the dipole-type gauge link structure. We will start with the case of GTMDs and discuss the forward limit in subsection 6.2.2. Hadron polarization will not be considered until subsection 6.2.3.

6.2.1 *The off-forward Wilson loop correlator*

The dipole-type GTMD correlator is given by

$$
\begin{aligned}
G^{[+,-]ij}(x,\mathbf{k},\xi,\mathbf{\Delta}) &= \frac{2}{P^+} \int \frac{dz^- d^2\mathbf{z}}{(2\pi)^3}\, e^{ik\cdot z} \\
&\quad \times \langle p' | \mathrm{Tr} \left(F^{+j}\!\left(\tfrac{z}{2}\right) U^{[-]\dagger}_{[-\frac{z}{2},\frac{z}{2}]} F^{+i}\!\left(-\tfrac{z}{2}\right) U^{[+]}_{[-\frac{z}{2},\frac{z}{2}]} \right) | p \rangle \bigg|_{z^+=0} \\
&= \frac{2}{P^+} \int \frac{dz^- d^2\mathbf{z}}{(2\pi)^3}\, e^{ik\cdot z - i\Delta\cdot b} \\
&\quad \times \langle p' | \mathrm{Tr} \left(F^{+j}(x) U^{[-]\dagger}_{[y,x]} F^{+i}(y) U^{[+]}_{[y,x]} \right) | p \rangle \bigg|_{x^+=y^+=b^+=z^+=0}, \quad (6.5)
\end{aligned}
$$

where $x \equiv b + \tfrac{z}{2}$ and $y \equiv b - \tfrac{z}{2}$. Using that the limit $\Delta \to 0$ of

$$
\langle p' | p \rangle = (2\pi)^3\, 2p^+ \delta(\Delta^+)\, \delta^{(2)}(\mathbf{\Delta}) \tag{6.6}
$$

is given by $\langle P | P \rangle = 2P^+ \int db^- d^2\mathbf{b}$, we can write

$$
\begin{aligned}
G^{[+,-]ij}(x,\mathbf{k},\xi,\mathbf{\Delta}) &= \frac{4}{\langle P|P\rangle} \int \frac{dx^- d^2\mathbf{x}\, dy^- d^2\mathbf{y}}{(2\pi)^3}\, e^{ik\cdot(x-y) - i\Delta\cdot\frac{x+y}{2}} \\
&\quad \times \langle p' | \mathrm{Tr} \left(F^{+j}(x) U^{[-]}_{[x,y]} F^{+i}(y) U^{[+]}_{[y,x]} \right) | p \rangle \bigg|_{x^+=y^+=0}. \quad (6.7)
\end{aligned}
$$

Now, let us consider the region where $x \sim \xi$ is small. To that end, we perform the limit

$$
G^{[+,-]ij}(\mathbf{k},\mathbf{\Delta}) \equiv \lim_{x,\xi\to 0} G^{[+,-]ij}(x,\mathbf{k},\xi,\mathbf{\Delta}), \tag{6.8}
$$

in which case the x^- and y^- integrations can be performed. Those integrations can be used to express the result in terms of the operators

$$
G^k_T(z) \equiv \frac{1}{2} \int_{-\infty}^{\infty} d\eta^- \, U^n_{[z^-,\eta^-;z]} \, F^{+k}(z^+,\eta^-,\mathbf{z}) \, U^n_{[\eta^-,z^-;z]}. \tag{6.9}
$$

Upon using the relation [79]

$$
\left[i\partial^k_z, U^{[\pm]}_{[a,z]} \right] = \mp g U^{[\pm]}_{[a,z]} G^k_T(z), \tag{6.10}
$$

the dipole-type GTMD correlator can be written as the correlator of a Wilson loop:

$$
\begin{aligned}
G^{[+,-]\,ij}(\boldsymbol{k},\boldsymbol{\Delta}) &= \frac{16}{\langle P|P\rangle} \int \frac{d^2\boldsymbol{x}\, d^2\boldsymbol{y}}{(2\pi)^3}\, e^{-i\boldsymbol{k}\cdot(\boldsymbol{x}-\boldsymbol{y})+i\boldsymbol{\Delta}\cdot\frac{\boldsymbol{x}+\boldsymbol{y}}{2}} \\
&\quad \times \langle p'|\,\mathrm{Tr}\left(G_T^{j}(x)\, U_{[x,y]}^{[-]}\, G_T^{i}(y)\, U_{[y,x]}^{[+]}\right)|p\rangle\,\bigg|_{x^+=y^+=0} \\
&= \frac{4}{g^2\langle P|P\rangle} \int \frac{d^2\boldsymbol{x}\, d^2\boldsymbol{y}}{(2\pi)^3}\, e^{-i\boldsymbol{k}\cdot(\boldsymbol{x}-\boldsymbol{y})+i\boldsymbol{\Delta}\cdot\frac{\boldsymbol{x}+\boldsymbol{y}}{2}} \\
&\quad \times \langle p'|\,\partial_x^j\partial_y^i\,\mathrm{Tr}\left(U_{[y,x]}^{[\Box]}\right)|p\rangle\,\bigg|_{x^+=y^+=0},
\end{aligned}
\tag{6.11}
$$

where now the off-forwardness is only transverse ($\boldsymbol{\Delta}$) and where

$$
U_{[y,x]}^{[\Box]} \equiv U_{[y,x]}^{[+]}\, U_{[x,y]}^{[-]}
\tag{6.12}
$$

is a rectangular Wilson loop with transverse distance $\boldsymbol{x}-\boldsymbol{y}$, stretched to infinity in the minus direction (see figure 6.1).
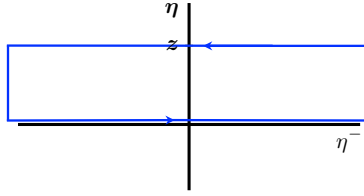
Figure 6.1: The integration path of the Wilson loop $U_{[0,z]}^{[\Box]}$ with transverse extent $|z|$. The closed path runs along the light-front $\eta^+ = 0$ via plus and minus light-cone infinity.

Next, we can apply partial integration twice and write eq. (6.11) as[2]

$$
G^{[+,-]\,ij}(\boldsymbol{k},\boldsymbol{\Delta}) = \frac{1}{2\pi g^2}\left[\frac{1}{2}\left(k^2 - \frac{\Delta^2}{4}\right)\delta_T^{ij} + k_T^{ij} - \frac{\Delta_T^{ij}}{4} - \frac{k_T^{[i}\Delta_T^{j]}}{2}\right] G^{[\Box]}(\boldsymbol{k},\boldsymbol{\Delta}),
\tag{6.13}
$$

where

$$
G^{[\Box]}(\boldsymbol{k},\boldsymbol{\Delta}) \equiv \frac{4N_c}{\langle P|P\rangle}\int \frac{d^2\boldsymbol{x}\, d^2\boldsymbol{y}}{(2\pi)^2}\, e^{-i\boldsymbol{k}\cdot(\boldsymbol{x}-\boldsymbol{y})+i\boldsymbol{\Delta}\cdot\frac{\boldsymbol{x}+\boldsymbol{y}}{2}}\,\langle p'|\,S^{[\Box]}(\boldsymbol{x},\boldsymbol{y})|p\rangle\,\bigg|_{x^+=y^+=0},
\tag{6.14}
$$

is the Fourier transform of the Wilson loop operator

$$
S^{[\Box]}(\boldsymbol{x},\boldsymbol{y}) \equiv \frac{1}{N_c}\,\mathrm{Tr}\left(U_{[y,x]}^{[\Box]}\right).
\tag{6.15}
$$

---

2 Eq. (6.13) is consistent with the result in [221] where only the term with δ_T^{ij} was considered.

We will refer to $G^{[\Box]}$ as the off-forward Wilson loop correlator. Naturally, also this correlator can be parametrized in terms of GTMDs, analogous to the parametrization in eq. (6.3); we choose the following parametrization:

$$G^{[\Box]}(k, \Delta) = \frac{2\pi g^2}{M^2} \, \mathcal{E}(k^2, \Delta^2, k \cdot \Delta),$$ (6.16)

where $\mathcal{E}$ is a (complex-valued) Wilson loop GTMD.

From eq. (6.16) we infer something important, namely that at vanishing x and ξ the picture is very simple – there is only *one* independent GTMD. From eqs. (6.13) and (6.16) it follows that in the limit of small x and ξ the GTMDs defined in eq. (6.3) are related as follows:

$$\lim_{x,\xi \to 0} x\mathcal{F}_1 = \lim_{x,\xi \to 0} x\mathcal{F}_2^{\{1\}} = -4 \lim_{x,\xi \to 0} x\mathcal{F}_3^{\{1\}} = -2 \lim_{x,\xi \to 0} x\mathcal{F}_4^{\{1\}} = \mathcal{E}^{\{1\}},$$ (6.17)

where we used the shorthand notation $\mathcal{F}_i^{\{n\}} \equiv [(k^2 - \Delta^2/4)/(2M^2)]^n \, \mathcal{F}_i$, which reduces to eq. (2.40) in the forward limit.

Let us now have a closer look at the Wilson loop operator $S^{[\Box]}$. This operator can be written in terms of its real and imaginary parts as follows:

$$S^{[\Box]}(x, y) = \mathcal{P}(x, y) + i\mathcal{O}(x, y),$$ (6.18)

with

$$\mathcal{P}(x, y) \equiv \frac{1}{2N_c} \operatorname{Tr}\left(U_{[y,x]}^{[\Box]} + U_{[y,x]}^{[\Box]\dagger} \right),$$ (6.19)

$$\mathcal{O}(x, y) \equiv \frac{1}{2iN_c} \operatorname{Tr}\left(U_{[y,x]}^{[\Box]} - U_{[y,x]}^{[\Box]\dagger} \right).$$ (6.20)

As it turns out, these parts behave differently under charge conjugation (C): the real part $\mathcal{P}$ is C-even and called the pomeron, while the imaginary part $\mathcal{O}$ is C-odd and called the odderon [222]. These charge conjugation properties can be easily understood starting from the transformation of the gauge field:

$$C A_\mu C^{-1} = -A_\mu.$$ (6.21)

It then follows that the gauge links transform as

$$C U_{[x,y]}^{[\pm]} C^{-1} = U_{[x,y]}^{[\pm]\dagger}, \qquad C U_{[y,x]}^{[\Box]} C^{-1} = U_{[y,x]}^{[\Box]\dagger},$$ (6.22)

so that, indeed,

$$C \mathcal{P}(x, y) C^{-1} = \mathcal{P}(x, y), \qquad C \mathcal{O}(x, y) C^{-1} = -\mathcal{O}(x, y).$$ (6.23)

In this chapter special attention will be given to the odderon, which was introduced already in 1973 in [223]. At lowest order in perturbation theory, an odderon

corresponds to three gluons in a symmetric color state [224]. First experimental indications for the existence of odderon effects have come from the recent results on pp collisions by the TOTEM experiment at CERN [225–227], combined with data on $p\bar{p}$ collisions by the DØ experiment at Fermilab [228].

Similar to eq. (5.9), the T-odd contribution to the dipole-type GTMD correlator in eq. (6.13) is given by:

$$
\begin{aligned}
G_{(d)}^{(\text{T-odd})\,ij}(\boldsymbol{k},\boldsymbol{\Delta}) &\equiv \frac{1}{2}\left[G^{[+,-]\,ij}(\boldsymbol{k},\boldsymbol{\Delta}) - G^{[-,+]\,ij}(\boldsymbol{k},\boldsymbol{\Delta})\right] \\
&= \frac{1}{4\pi g^2}\left[\frac{1}{2}\left(\boldsymbol{k}^2 - \frac{\boldsymbol{\Delta}^2}{4}\right)\delta_T^{ij} + k_T^{ij} - \frac{\Delta_T^{ij}}{4} - \frac{k_T^{[i}\Delta_T^{j]}}{2}\right] \\
&\quad \times \left[G^{[\Box]}(\boldsymbol{k},\boldsymbol{\Delta}) - G^{[\Box^\dagger]}(\boldsymbol{k},\boldsymbol{\Delta})\right] \\
&= \frac{2iN_c}{\pi g^2\,\langle P|P\rangle}\left[\frac{1}{2}\left(\boldsymbol{k}^2 - \frac{\boldsymbol{\Delta}^2}{4}\right)\delta_T^{ij} + k_T^{ij} - \frac{\Delta_T^{ij}}{4} - \frac{k_T^{[i}\Delta_T^{j]}}{2}\right] \\
&\quad \times \int \frac{d^2\boldsymbol{x}\,d^2\boldsymbol{y}}{(2\pi)^2}\, e^{-ik\cdot(x-y)+i\Delta\cdot\frac{x+y}{2}}\,\langle p'|\,\mathcal{O}(\boldsymbol{x},\boldsymbol{y})\,|p\rangle\bigg|_{x^+=y^+=0}, \quad (6.24)
\end{aligned}
$$

which is proportional to a matrix element of the odderon operator. Thus, the odderon contribution to the dipole-type GTMD correlator is not only C-odd but also T-odd. In the same way it follows that the T-even contribution is proportional to a matrix element of the pomeron operator. From the Hermiticity and time reversal constraints, respectively given by

$$
G^{[\Box]*}(\boldsymbol{k},\boldsymbol{\Delta}) = G^{[\Box]}(\boldsymbol{k},-\boldsymbol{\Delta}), \qquad G^{[\Box]*}(\boldsymbol{k},\boldsymbol{\Delta}) = G^{[\Box^\dagger]}(-\boldsymbol{k},-\boldsymbol{\Delta}), \qquad (6.25)
$$

it follows that the combination $G^{[\Box]}(\boldsymbol{k},\boldsymbol{\Delta}) - G^{[\Box^\dagger]}(\boldsymbol{k},\boldsymbol{\Delta})$ appearing on the second line of eq. (6.24) only contains *odd* powers of $\boldsymbol{k}\cdot\boldsymbol{\Delta}$. Likewise, the T-even contribution only contains *even* powers of $\boldsymbol{k}\cdot\boldsymbol{\Delta}$. Hence, if we were to expand $\mathcal{E}$ in powers of $\boldsymbol{k}\cdot\boldsymbol{\Delta}$ and extract the angle $\phi \equiv \phi_k - \phi_\Delta$ (we will do something similar for the gluon Wigner distributions in subsection 6.5.2), then the pomeron and odderon contributions would, respectively, be constituted by all even harmonics $\cos(2n\phi)$ and all odd harmonics $\cos((2n+1)\phi)$, where $n \geq 0$.

6.2.2 *The forward Wilson loop correlator*

In the remainder of this section we will consider the TMD case. In the forward limit, the dipole-type GTMD correlator reduces to the dipole-type TMD correlator:

$$
\lim_{\Delta \to 0} G^{[+,-]\,ij}(x,\boldsymbol{k},\xi,\boldsymbol{\Delta}) = \Gamma^{[+,-]\,ij}(x,\boldsymbol{k}). \qquad (6.26)
$$

Employing eqs. (6.8), (6.13), and the $\Delta \to 0$ limit of (6.6), it follows that in the small-x limit the TMD correlator is given by

$$
\begin{aligned}
\lim_{x \to 0} \Gamma^{[+,-]ij}(x, \boldsymbol{k}) &= \lim_{\boldsymbol{\Delta} \to 0} G^{[+,-]ij}(\boldsymbol{k}, \boldsymbol{\Delta}) \\
&= \frac{1}{2\pi g^2 L} \left(\frac{1}{2} \boldsymbol{k}^2 \delta_T^{ij} + k_T^{ij} \right) \Gamma^{[\square]}(\boldsymbol{k}),
\end{aligned}
\tag{6.27}
$$

where $L \equiv 2\pi \, \delta(\Delta^+ = 0)$ and

$$
\begin{aligned}
\Gamma^{[\square]}(\boldsymbol{k}) &\equiv \lim_{\boldsymbol{\Delta} \to 0} L \, G^{[\square]}(\boldsymbol{k}, \boldsymbol{\Delta}) \\
&= \frac{2N_c}{P^+} \int \frac{d^2 z}{(2\pi)^2} \, e^{-i\boldsymbol{k} \cdot \boldsymbol{z}} \left. \langle P| S^{[\square]}(\boldsymbol{z}, \boldsymbol{0}) |P\rangle \right|_{z^+=0}
\end{aligned}
\tag{6.28}
$$

is the *forward* Wilson loop correlator. We choose to parametrize this correlator as follows (the usual Hermiticity, parity, and time reversal constraints will be given in the next subsection where we will include hadron spin):

$$
\Gamma^{[\square]}(\boldsymbol{k}) = \frac{2\pi g^2 L}{M^2} \, e(\boldsymbol{k}^2),
\tag{6.29}
$$

where e is a (real) T-even Wilson loop TMD.

Similar to the GTMD case in eq. (6.16), there is only *one* independent TMD in the small-x limit. From eqs. (6.27) and (6.29) it follows that the two TMDs appearing in eq. (5.17) are related as

$$
\lim_{x \to 0} x f_1 = \lim_{x \to 0} x h_1^{\perp(1)} = e^{(1)}.
\tag{6.30}
$$

This result implies that bound (5.41) is saturated. It also indicates that the unpolarized dipole gluon distribution grows as $1/x$ towards small x, apart from subdominant modifications arising from resummation of large logarithms in $1/x$ (see e.g. [229]) and higher-twist effects. From a recent global fit called NNPDF3.0 it follows that the small-x asymptotic behavior of the collinear gluon PDF for the proton at $x = 10^{-4}$ is $1/x^{1.20\pm0.46}$ at $Q^2 = 10$ GeV^2 and $1/x^{1.382\pm0.011}$ at $Q^2 = 10^4$ GeV^2 [230].

From eqs. (6.16), (6.28), and (6.29) we infer that the Wilson loop function $\mathcal{E}$ reduces to e in the forward limit:

$$
\lim_{\boldsymbol{\Delta} \to 0} \mathcal{E}(\boldsymbol{k}^2, \boldsymbol{\Delta}^2, \boldsymbol{k} \cdot \boldsymbol{\Delta}) = e(\boldsymbol{k}^2).
\tag{6.31}
$$

In the previous subsection we argued that $\mathcal{E}$ contains both pomeron and odderon contributions. In the forward limit, only the angular-independent harmonic ($n = 0$) contained in $\mathcal{E}$ survives – this is a pomeron contribution and precisely corresponds with the T-even function e. Thus, at least for unpolarized hadrons and at leading

6

twist, there are *no* odderon contributions in the TMD case. Figures 6.2 and 6.3 summarize the main results of this section so far. They provide the relations between the various correlators and distribution functions that we have discussed. We stress that the main results, eqs. (6.17) and (6.30), only hold at leading twist and for the dipole-type gauge link structure. In different contexts, the small-x limit has also been considered for the WW-type gauge link structure, see e.g. [231]. However, it seems in that case impossible to relate the result to a Wilson loop or any other 'simple structure' that would allow for partial integration as in eq. (6.13). Hence, we think that a similar analysis for the WW case is not possible.
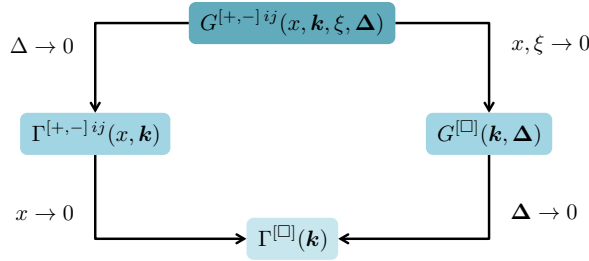
Figure 6.2: Relations (schematically) between the dipole-type TMD and GTMD correlators and the forward and off-forward Wilson loop correlators. The precise relations are given in the main text.
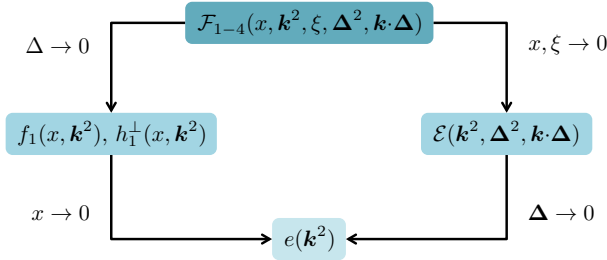
Figure 6.3: Relations (schematically) between the gluon and Wilson loop TMDs and GTMDs for unpolarized hadrons. The precise relations are given in the main text.

6.2.3 *Wilson loop TMDs for polarized hadrons*

In subsection 6.2.1 we studied the limit of small x, ξ of the GTMD correlator for *unpolarized* hadrons, and in subsection 6.2.2 we considered the forward limit of those results. However, it is easy to see that all these limits are actually completely independent of hadron polarization. In particular, eq. (6.27) *also* holds for vector and tensor polarized hadrons (this result was already obtained for transversely polarized

hadrons in [219]). The left-hand side of that equation has been parametrized in terms of TMDs in eq. (5.16) for spin-1 hadrons. Naturally, we can do the same thing for the right-hand side, i.e. we can extend the parametrization of the forward Wilson loop correlator for unpolarized hadrons in eq. (6.29) to the case of spin-1 hadrons; this will be done below.

For spin-1 hadrons, the forward Wilson loop correlator is defined as

$$\Gamma^{[\square]}(\boldsymbol{k}) \equiv \frac{2N_c}{P^+} \int \frac{d^2\boldsymbol{z}}{(2\pi)^2}\, e^{-ik\cdot z}\, \langle P; S, T|\, S^{[\square]}(\boldsymbol{z}, \boldsymbol{0})\, |P; S, T\rangle \bigg|_{z^+=0}. \tag{6.32}$$

Since this correlator has no open indices, a parametrization in terms of independent Wilson loop TMDs is rather straightforward. The relevant Hermiticity and parity constrains follow directly from eqs. (5.2) and (5.3) and are respectively given by

$$\Gamma^{[\square]}(\boldsymbol{k}, P, S, T, n) = \Gamma^{[\square]*}(\boldsymbol{k}, P, S, T, n), \tag{6.33}$$

$$\Gamma^{[\square]}(\boldsymbol{k}, P, S, T, n) = \Gamma^{[\square]}(-\boldsymbol{k}, \bar{P}, -\bar{S}, \bar{T}, \bar{n}). \tag{6.34}$$

We propose the following parametrization that respects these constraints:

$$\Gamma^{[\square]}(\boldsymbol{k}) = \frac{2\pi g^2 L}{M^2} \left[e(\boldsymbol{k}^2) + \frac{\epsilon_T^{S_T k_T}}{M}\, e_T(\boldsymbol{k}^2) + S_{LL}\, e_{LL}(\boldsymbol{k}^2) \right.$$

$$\left. + \frac{\boldsymbol{k}\cdot\boldsymbol{S}_{LT}}{M}\, e_{LT}(\boldsymbol{k}^2) + \frac{k_T^{\alpha\beta} S_{TT\alpha\beta}}{M^2}\, e_{TT}(\boldsymbol{k}^2) \right]. \tag{6.35}$$

According to the time reversal constraint

$$\Gamma^{[\square]*}(\boldsymbol{k}, P, S, T, n) = \Gamma^{[\square^\dagger]}(-\boldsymbol{k}, \bar{P}, \bar{S}, \bar{T}, \bar{n}), \tag{6.36}$$

the function e_T is T-odd while the other functions are T-even. From eqs. (5.9), (6.20), (6.27), and (6.32) it follows that in the small-x limit the T-odd contribution to the dipole-type TMD correlator is given by

$$\lim_{x\to 0} \Gamma_{(d)}^{(\text{T-odd})\,ij}(x, \boldsymbol{k}) \equiv \frac{1}{2} \lim_{x\to 0} \left[\Gamma^{[+,-]\,ij}(x, \boldsymbol{k}) - \Gamma^{[-,+]\,ij}(x, \boldsymbol{k}) \right]$$

$$= \frac{1}{4\pi g^2 L} \left(\frac{1}{2}\boldsymbol{k}^2 \delta_T^{ij} + k_T^{ij} \right) \left[\Gamma^{[\square]}(\boldsymbol{k}) - \Gamma^{[\square^\dagger]}(\boldsymbol{k}) \right]$$

$$= \frac{iN_c}{\pi g^2 P^+ L} \left(\frac{1}{2}\boldsymbol{k}^2 \delta_T^{ij} + k_T^{ij} \right)$$

$$\times \int \frac{d^2\boldsymbol{z}}{(2\pi)^2}\, e^{-ik\cdot z}\, \langle P; S, T|\, \mathcal{O}(\boldsymbol{z}, \boldsymbol{0})\, |P; S, T\rangle \bigg|_{z^+=0}. \tag{6.37}$$

Hence, analogous to eq. (6.24) for the GTMD case, the T-odd contribution features the odderon operator. Since there is only one T-odd Wilson loop TMD, this contribution is completely captured by the function e_T.

Note that in eq. (6.35) there is no Wilson loop TMD corresponding to longitudinally polarized hadrons. Thus, it follows from eqs. (5.18) and (6.27) that g_1 and $h_{1L}^\perp$ vanish in the small-x limit, or rather that they are less divergent than $1/x$ in the limit of small x. The latter is in agreement with the results in [232, 233] where large logarithms in $1/x$ were resummed for the gluon helicity function g_1. For the other hadron polarizations we can derive relations similar to the one for unpolarized hadrons, eq. (6.30). Notice that as the right-hand side of eq. (6.27) is symmetric in i, j, the g-type functions (that come with the antisymmetric structure $i\epsilon_T^{ij}$) generally vanish in the small-x limit.

Now let us consider transversely polarized hadrons. From eqs. (6.27) and (6.35) it follows that the TMDs appearing in eq. (5.19) are related as

$$\lim_{x\to 0} xf_{1T}^\perp = \lim_{x\to 0} xh_1 = -\lim_{x\to 0} xh_{1T}^{\perp(1)} = e_T^{(1)}. \qquad (6.38)$$

This is in agreement with the results in [219]. The fact that a transversely polarized hadron can give rise to an odderon contribution was already noted in [234–236], but without discussion of its operator structure. The precise operator structure was first given in [237] and was referred to as the 'spin-dependent' odderon. As discussed in the previous subsection, we have not identified a spin-*independent* odderon. However, as we argued below eq. (6.25), odderon contributions do appear for unpolarized hadrons in the GTMD case.

Finally, let us consider tensor polarized hadrons. In the small-x limit, the TMDs for longitudinal-longitudinal (LL) polarization appearing in eq. (5.20) are related as

$$\lim_{x\to 0} xf_{1LL} = \lim_{x\to 0} xh_{1LL}^{\perp(1)} = e_{LL}^{(1)}. \qquad (6.39)$$

For longitudinal-transverse (LT) polarization, see eq. (5.21), we obtain

$$\lim_{x\to 0} xf_{1LT} = \lim_{x\to 0} xh_{1LT} = -\frac{1}{2}\lim_{x\to 0} xh_{1LT}^{\perp(1)} = \frac{1}{2}e_{LT}^{(1)}, \qquad (6.40)$$

and for the TMDs corresponding to transverse-transverse (TT) polarization in eq. (5.22) we find

$$\lim_{x\to 0} xf_{1TT}^{(1)} = \frac{1}{3}\lim_{x\to 0} xh_{1TT} = -\frac{1}{2}\lim_{x\to 0} xh_{1TT}^{\perp(1)} = \frac{1}{3}\lim_{x\to 0} xh_{1TT}^{\perp\perp(2)} = \frac{1}{3}e_{TT}^{(2)}. \qquad (6.41)$$

Note that for all hadron polarizations the time reversal properties of the dipole-type TMDs and the Wilson loop TMDs correctly match.

Similarly to the gluon TMDs in section 5.2, we can also derive bounds for the Wilson loop TMDs that appear in the parametrization of the forward Wilson loop correlator given by

$$\Gamma^{[\square]}(\boldsymbol{k}, S, T) \equiv \mathrm{Tr}\left(\rho(S, T)\, N^{[\square]}(\boldsymbol{k})\right) = \rho_{s's}(S, T)\, N_{ss'}^{[\square]}(\boldsymbol{k}). \qquad (6.42)$$

Upon inverting this relation, we find that

$$
N^{[\square]} = \frac{2\pi g^2 L}{M^2}
\begin{pmatrix}
e + \frac{e_{LL}}{2} & \frac{|k|e^{-i\phi}}{\sqrt{2}M}\left(e_{LT} + ie_T\right) & \frac{k^2 e^{-2i\phi}}{M^2} e_{TT} \\
\frac{|k|e^{i\phi}}{\sqrt{2}M}\left(e_{LT} - ie_T\right) & e - e_{LL} & -\frac{|k|e^{-i\phi}}{\sqrt{2}M}\left(e_{LT} - ie_T\right) \\
\frac{k^2 e^{2i\phi}}{M^2} e_{TT} & -\frac{|k|e^{i\phi}}{\sqrt{2}M}\left(e_{LT} + ie_T\right) & e + \frac{e_{LL}}{2}
\end{pmatrix} . \quad (6.43)
$$

In analogy to eq. (5.26), we can write the elements of $N^{[\square]}$ in the following form:

$$
\begin{aligned}
N_{ss'}^{[\square]}(k) &\equiv \frac{2N_c}{P^+} \int \frac{d^2 z}{(2\pi)^2} \, e^{-ik \cdot z} \, \langle P; s| \, S^{[\square]}(z, 0) \, |P; s'\rangle \Big|_{z^+ = 0} \\
&= \frac{2}{P^+} \sum_m \text{Tr} \left(\langle P_m| \, U_{[\infty;\infty,0]}^T \, U_{[\infty,-\infty;0]}^n \, U_{[-\infty;0,\infty]}^T \, |P; s\rangle^* \right. \\
&\quad \left. \times \langle P_m| \, U_{[\infty;\infty,0]}^T \, U_{[\infty,-\infty;0]}^n \, U_{[-\infty;0,\infty]}^T \, |P; s'\rangle \right) \Big|_{z^+ = 0} \delta^{(2)}(P_m + k), \quad (6.44)
\end{aligned}
$$

where we inserted a complete set of momentum eigenstates $\{|P_m\rangle\}$ and we used that

$$
\begin{aligned}
U_{[0,z]}^{[\square]} &= U_{[-\infty,\infty;0]}^n \, U_{[\infty;0,z]}^T \, U_{[\infty,-\infty;z]}^n \, U_{[-\infty;z,0]}^T \\
&= \left(U_{[-\infty;\infty,0]}^T \, U_{[-\infty,\infty;0]}^n \, U_{[\infty;0,\infty]}^T \right) \left(U_{[-\infty;\infty,z]}^T \, U_{[-\infty,\infty;z]}^n \, U_{[\infty;z,\infty]}^T \right)^\dagger . \quad (6.45)
\end{aligned}
$$

From eq. (6.44) it follows that $N^{[\square]}$ is positive semidefinite – in other words, its eigenvalues must be ≥ 0. To establish bounds for the Wilson loop TMDs, we restrict ourselves to two-dimensional principal minors and find

$$
\frac{k^2}{2M^2}\left(e_T^2 + e_{LT}^2\right) \leq (e - e_{LL})\left(e + \frac{e_{LL}}{2}\right), \quad (6.46)
$$

$$
\frac{k^2}{M^2}|e_{TT}| \leq e + \frac{e_{LL}}{2}. \quad (6.47)
$$

Applying the small-x limit to bounds (5.27)–(5.35), one indeed recovers bounds (6.46) and (6.47). Besides these two bounds, we also have $e \geq 0$ (this follows directly from eq. (6.45)).

Let us finally also comment on the case of a spin-$\frac{1}{2}$ hadron. The parametrization of the forward Wilson loop correlator for spin-$\frac{1}{2}$ hadrons is given in terms of the two functions e and e_T. For $N^{[\square]}$ we now obtain

$$
N^{[\square]} = \frac{2\pi g^2 L}{M^2}
\begin{pmatrix}
e & \frac{i|k|e^{-i\phi}}{M} e_T \\
-\frac{i|k|e^{i\phi}}{M} e_T & e
\end{pmatrix}, \quad (6.48)
$$

from which we can derive the following upper bound for the magnitude of the spin-dependent odderon e_T:

$$
\frac{|k|}{M}|e_T| \leq e. \quad (6.49)
$$

Note that upon omitting the functions related to tensor polarization, the bounds (6.46) and (6.47) reduce to a bound that is consistent with but less strict than (6.49). Aside from diagonalizing higher-dimensional minors to sharpen bounds (6.46) and (6.47), we could also obtain (6.49) by applying the small-x limit to bounds (5.36)–(5.40) for spin-$\frac{1}{2}$ hadrons.

6.3 THE MCLERRAN-VENUGOPALAN MODEL

Probed at increasingly smaller values of x, the gluon density of hadrons and nuclei rises. This implies that the description of an incoming parton scattering off a single gluon starts to fall short – due to the high density of gluons, scattering off many gluons needs to be taken into account. Rather than describing gluons inside nuclei by PDFs, we could model the internal structure of nuclei by an effective color field called the color glass condensate (CGC). The term 'glass' refers to the state of 'frozen disorder' as seen by a probe (due to time dilation), and the term 'condensate' refers to the very high occupation numbers of the gluon field. The CGC effective theory is the most widely accepted description of the gluon content of ultrarelativistic nuclei. This theory will be employed in the remainder of this chapter, in particular to calculate the expectation values of the pomeron and odderon operators in a nuclear environment in section 6.4. Reviews on the CGC formalism can be found in e.g. [107, 238–243].

The starting point of the CGC picture is provided by the MV model [244–246] in which the large-x partons (the valence quarks) are considered as static, acting as color charge sources for the smaller-x gluons that constitute the color field. The typical size of fluctuations in the color charge density is characterized by the so-called saturation scale Q_s. As it turns out, this scale grows with increasing atomic number A and with decreasing x; more specifically $Q_s^2 \propto A^{1/3}/x^\lambda$, with $\lambda \sim 0.2$-0.3 [242, 247]. Hence, for sufficiently large nuclei (i.e. $A \gg 1$) and small values of x (typically one requires $x \lesssim 0.01$) we have $Q_s \gg \Lambda_{\text{QCD}}$ and $\alpha_s(Q_s) \ll 1$. This implies that the small-x gluon field of a large nucleus can be described by a *classical* color field – quantum corrections are suppressed by extra powers of α_s [244, 245]. This color field is given by the classical Yang-Mills equations of motion [248]:[3]

$$D_\mu F^{\mu\nu} = J^\nu. \tag{6.50}$$

For an ultrarelativistic nucleus moving along the plus direction, the source current J only has a nonzero plus component:

$$J^\nu(x^-, \boldsymbol{x}) = \delta^{\nu+}\rho(x^-, \boldsymbol{x}), \tag{6.51}$$

where ρ is the color charge density of the sources (due to current conservation this density is independent of x^+).

---

3 In the following discussion, several quantities 'live' in the adjoint representation of SU(3), e.g. we have $J^\nu \equiv J^{a\nu}T^a$.

Often CGC calculations are performed in a covariant gauge with the only nonzero field component given by $\alpha \equiv A^+$ (the minus component can be neglected in our field configuration); we will stick to this gauge in the remainder of this chapter. In this particular gauge, the solution to eq. (6.50) is given by [249]

$$\nabla_x^2 \alpha(x^-, \boldsymbol{x}) = -\rho(x^-, \boldsymbol{x}). \tag{6.52}$$

Upon inverting this relation, we obtain the color field in terms of the color sources:

$$\alpha^a(x^-, \boldsymbol{x}) = \int d^2\boldsymbol{z} \, G(\boldsymbol{x} - \boldsymbol{z}) \, \rho^a(x^-, \boldsymbol{z}), \tag{6.53}$$

where the Green's function associated to the two-dimensional Laplacian is given by

$$G(\boldsymbol{x} - \boldsymbol{z}) = \frac{1}{4\pi} \ln \left(\frac{1}{(\boldsymbol{x} - \boldsymbol{z})^2 \Lambda^2} \right). \tag{6.54}$$

Here Λ is an infrared cut-off scale, commonly taken to be Λ_{QCD}.

The result in eq. (6.53) provides the classical color field that arises from a single configuration of color sources described by the density ρ. Even though the sources (three valence quarks per nucleon) are considered static in this model, their precise distribution can vary between collisions – the density ρ is *arbitrary*. Hence, to calculate an observable quantity, we need to *average* the corresponding operator over all possible color source configurations. This can be achieved by a functional average that is weighted by a suitable weight functional. In the (original) MV model, this weight functional is taken to be Gaussian in the color sources. The expectation value of an operator $O = O[\rho]$ (also referred to as a CGC average) can be calculated as

$$\langle O \rangle_G = \int \mathcal{D}\rho \, w_G[\rho] \, O[\rho], \tag{6.55}$$

provided the weight functional w_G is properly normalized, i.e. $\int \mathcal{D}\rho \, w_G[\rho] = 1$. The Gaussian weight functional is given by

$$w_G[\rho] = \exp \left(-\int d^2\boldsymbol{x} \, \frac{\delta^{ab} \rho^a(\boldsymbol{x}) \rho^b(\boldsymbol{x})}{2\mu_A} \right), \tag{6.56}$$

with $\rho^a(\boldsymbol{x}) \equiv \int dx^- \rho^a(x^-, \boldsymbol{x})$, and $\mu_A = g^2 A/(2\pi R_A^2)$ represents the average color charge squared per unit area (R_A denotes the nuclear radius).

In general, the calculation of a functional average of an arbitrary product of sources can be expressed in terms of the average of a product of two sources,

$$\left\langle \rho^a(\boldsymbol{x}) \rho^b(\boldsymbol{y}) \right\rangle_G = \mu_A \, \delta^{ab} \delta^{(2)}(\boldsymbol{x} - \boldsymbol{y}). \tag{6.57}$$

While the average of a product of an even number of sources can be obtained by summing over all possible pairings of two sources, the average of a product of an odd

number of sources vanishes [250]. As was first argued in [222], the latter implies that the expectation value of the odderon operator in the (original) MV model vanishes in a trivial way. In the next section we will explain how the functional averaging can be modified to properly account for odderon excitations in the CGC theory.

As mentioned earlier, the MV model is based on the separation between large-x partons, the sources, and small-x partons, the gluons. Naturally, an observable should not depend on what values of x are called small or large with respect to some arbitrary reference scale. This requirement gives rise to the so-called Jalilian-Marian-Iancu-McLerran-Weigert-Leonidov-Kovner (JIMWLK) evolution equation [249, 251–256]. This is a complicated differential equation in x for the weight functional, with the Gaussian weight w_G serving as its initial condition. Evolution in x does not preserve the Gaussian form and goes beyond the MV model (and beyond the scope of this thesis). In the remainder of this chapter, all CGC averages are (implicitly) considered at some small (but nonzero) longitudinal momentum fraction x.

6.4 THE DIPOLE OPERATOR IN A NUCLEAR COLOR GLASS CONDENSATE

In section 6.5 we will study dihadron production in pA collisions through two incoming quarks from the proton that scatter off the small-x gluon field of the nucleus, modeled by a CGC. As it turns out, the cross section of that process features a product of two dipole operators. The dipole operator corresponds to the S-matrix operator for a $q\bar{q}$ pair (a 'color dipole') that scatters off a nucleus through infinitely many gluon exchanges; with the antiquark at position x and the quark at position y, the dipole operator is given by

$$S(x, y) \equiv \frac{1}{N_c} \operatorname{Tr} \left(U_x^\dagger U_y \right), \tag{6.58}$$

where

$$U_z \equiv U_{[-\infty, \infty; z]}^n = \mathcal{P} \exp \left(ig \int_{-\infty}^{\infty} d\eta^- \alpha^a (\eta^-, z) \, t^a \right). \tag{6.59}$$

The color gauge invariant version of the dipole operator is the Wilson loop operator $S^{\square}$ given in eq. (6.15). To facilitate the study of pA collisions in the CGC theory, we will calculate in this section the CGC averages of the real (C-even) and imaginary (C-odd) parts of the dipole operator. In our covariant gauge, those pomeron and odderon contributions are respectively given by

$$\mathcal{P}(x, y) \equiv \frac{1}{2N_c} \operatorname{Tr} \left(U_x^\dagger U_y + U_y^\dagger U_x \right), \tag{6.60}$$

$$\mathcal{O}(x, y) \equiv \frac{1}{2iN_c} \operatorname{Tr} \left(U_x^\dagger U_y - U_y^\dagger U_x \right). \tag{6.61}$$

In the following calculations, we will deviate from the original MV model described in section 6.3 in a number of ways. First of all, we will calculate the

expectation values of the pomeron and odderon operators for a nuclear environment where the different nucleons are distributed in transverse space according to some profile function. The latter is in contrast with the original MV model where the nuclear profile is taken to be constant – that assumption is reflected in the constant-ness of μ_A in eq. (6.56). Since for our analysis in section 6.5 it is important to allow for a nontrivial transverse structure of the nucleus, we perform the replacement $\mu_A \rightarrow \mu(\boldsymbol{u})$, with $\boldsymbol{u}$ the position of a color source with respect to the center of the nucleus. Another difference with the original MV model lies in the modeling of nucleons: for simplicity, each nucleon is modeled to contain one valence quark only. Finally, as already mentioned in the previous section, the odderon operator requires a modified weight functional. As it turns out, this modification amounts to adding an extra term to the exponent that is *cubic* in the color charge density [257]. This extra term, that goes beyond the original MV model formulation, was derived in [258] and was shown to be subleading in the large-A limit as compared to the Gaussian term. With all these modifications, the CGC average for a given operator $O = O[\rho]$ is given by

$$\langle O \rangle = \int \mathcal{D}\rho \, w[\rho] \, O[\rho]. \tag{6.62}$$

The weight functional now reads

$$w[\rho] = w_G[\rho] \, w_c[\rho], \tag{6.63}$$

with the Gaussian and cubic weights respectively given by

$$w_G[\rho] = \exp\left(-\int d^2\boldsymbol{n} \int \frac{d^2\boldsymbol{q}}{\pi R_N^2} \, \frac{\delta^{ab}\rho^a(\boldsymbol{u})\,\rho^b(\boldsymbol{u})}{2\mu(\boldsymbol{u})} \right), \tag{6.64}$$

$$w_c[\rho] = \exp\left(\int d^2\boldsymbol{n} \int \frac{d^2\boldsymbol{q}}{\pi R_N^2} \, \frac{d^{abc}\rho^a(\boldsymbol{u})\,\rho^b(\boldsymbol{u})\,\rho^c(\boldsymbol{u})}{\kappa(\boldsymbol{u})} \right), \tag{6.65}$$

where, again, the weights are properly normalized. The quantities μ and κ can be related to the nuclear profile function T as follows:

$$\mu(\boldsymbol{u}) = \frac{g^2}{2N_c} \, T(\boldsymbol{u}), \qquad \kappa(\boldsymbol{u}) = \frac{g^3}{N_c} \, T^2(\boldsymbol{u}). \tag{6.66}$$

If the nuclear profile were taken to be constant within the nucleus, the function T would simply equal the average number of nucleons per unit of area, $A/(\pi R_A^2)$. Rather than integrating over $\boldsymbol{u} = \boldsymbol{n} + \boldsymbol{q}$ in the CGC weights, we instead average over the position $\boldsymbol{q}$ of the valence quark with respect to the center of the nucleon (we denote the radius of the nucleon by R_N), as well as integrate over all positions $\boldsymbol{n}$ of nucleons with respect to the center of the nucleus (note that at this point the two ways of writing the integrations are completely equivalent as the integrands depend on $\boldsymbol{u}$ only; the latter form will, however, prove more convenient in what follows). The geometrical picture is schematically illustrated in figure 6.4.
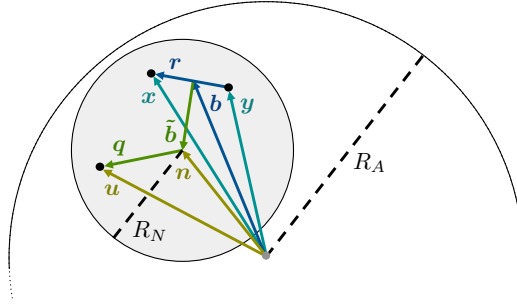
Figure 6.4: Geometrical picture of a dipole hitting a nucleon (that contains one valence quark) inside a nucleus.

6.4.1 *The pomeron operator*

In the following we will calculate the CGC averages of the real and imaginary parts of the dipole operator according to eq. (6.62). We start by providing the expectation value of the pomeron operator given in eq. (6.60). The CGC average of this operator can be obtained by considering the dipole operator with a Gaussian weight (the cubic weight does not affect the real part). The expectation value of the pomeron operator is given by the well-known result [250, 259, 260]

$$
\begin{aligned}
\langle \mathcal{P}(\boldsymbol{x}, \boldsymbol{y}) \rangle &= \langle S(\boldsymbol{x}, \boldsymbol{y}) \rangle_G \\
&= \frac{1}{N_c} \int \mathcal{D}\rho \, w_G[\rho] \, \mathrm{Tr} \left(U_{\boldsymbol{x}}^\dagger U_{\boldsymbol{y}} \right) \\
&= \exp \left(-\frac{1}{4} |\boldsymbol{r}|^2 Q_s^2(|\boldsymbol{b}|) \ln \left(\frac{1}{|\boldsymbol{r}|\Lambda} \right) \right),
\end{aligned}
\tag{6.67}
$$

where $\boldsymbol{b} \equiv (\boldsymbol{x} + \boldsymbol{y})/2$ represents the impact parameter of the dipole with respect to the center of the nucleus and $|\boldsymbol{r}| \equiv |\boldsymbol{x} - \boldsymbol{y}|$ is the dipole size (see also figure 6.4). The saturation scale Q_s is defined in terms of the nuclear profile function T as

$$
Q_s^2(|\boldsymbol{b}|) \equiv 2\pi \alpha_s^2 \frac{N_c^2 - 1}{N_c^2} T(|\boldsymbol{b}|).
\tag{6.68}
$$

6.4.2 *The odderon operator*

Next, we will calculate the expectation value of the odderon operator given in eq. (6.61). In the following we generalize the derivation of [257] by taking into account inhomogeneities in the transverse structure of the nucleus. Our final result matches with the result given in [260], which was obtained in [261] following a diagrammatic approach.

The CGC average of the odderon operator is given by

$$\langle \mathcal{O}(\boldsymbol{x}, \boldsymbol{y}) \rangle = \frac{1}{2iN_c} \int \mathcal{D}\rho \, w_c[\rho] \, w_G[\rho] \, \mathrm{Tr}\left(U_x^\dagger U_y - U_y^\dagger U_x\right)$$

$$\approx \frac{d^{abc}}{2iN_c} \int d^2\boldsymbol{n} \, \frac{1}{\kappa(|\boldsymbol{n}|)} \int \frac{d^2\boldsymbol{q}}{\pi R_N^2} \int \mathcal{D}\rho \, \rho^a(\boldsymbol{u}) \, \rho^b(\boldsymbol{u}) \, \rho^c(\boldsymbol{u}) \, w_G[\rho]$$

$$\times \, \mathrm{Tr}\left(U_x^\dagger U_y - U_y^\dagger U_x\right), \tag{6.69}$$

where we have expanded the cubic weight up to first nontrivial order. Furthermore, we have used that $T(\boldsymbol{u}) \approx T(|\boldsymbol{n}|)$, i.e. we have neglected the difference between the location of a given nucleon and the location of a valence quark inside that nucleon and we have assumed rotational symmetry. From the functional derivative acting on the Gaussian weight,

$$\delta^a w_G[\rho] \equiv \frac{\delta}{\delta \rho^a(u^-, \boldsymbol{u})} \, w_G[\rho] = -\frac{\rho^a(\boldsymbol{u})}{\mu(\boldsymbol{n})} \, w_G[\rho], \tag{6.70}$$

it follows that

$$\langle \mathcal{O}(\boldsymbol{x}, \boldsymbol{y}) \rangle = -\frac{d^{abc}}{2iN_c} \int d^2\boldsymbol{n} \, \frac{\mu^3(|\boldsymbol{n}|)}{\kappa(|\boldsymbol{n}|)} \int \frac{d^2\boldsymbol{q}}{\pi R_N^2} \int \mathcal{D}\rho \left(\delta^a \delta^b \delta^c \, w_G[\rho]\right) \mathrm{Tr}\left(U_x^\dagger U_y - U_y^\dagger U_x\right)$$

$$= \frac{g^3}{16iN_c^3} \int d^2\boldsymbol{n} \, T(|\boldsymbol{n}|) \int \frac{d^2\boldsymbol{q}}{\pi R_N^2} \left\langle d^{abc} \delta^a \delta^b \delta^c \, \mathrm{Tr}\left(U_x^\dagger U_y - U_y^\dagger U_x\right)\right\rangle_G, \tag{6.71}$$

where in the second step we used eq. (6.66) and integration by parts.

For convenience, we will now only consider the first term of eq. (6.71) explicitly. At the end, the second term can be obtained from the first term by changing the overall sign and interchanging $\boldsymbol{x}$ and $\boldsymbol{y}$. Using the symmetry properties of d, we obtain

$$d^{abc} \delta^a \delta^b \delta^c \, \mathrm{Tr}\left(U_x^\dagger U_y\right) = d^{abc} \, \mathrm{Tr}\left((\delta^a \delta^b \delta^c U_x^\dagger) U_y + U_x^\dagger (\delta^a \delta^b \delta^c U_y)\right)$$

$$+ 3(\delta^a \delta^b U_x^\dagger)(\delta^c U_y) + 3(\delta^a U_x^\dagger)(\delta^b \delta^c U_y)\right). \tag{6.72}$$

A single functional derivative acting on a Wilson line gives

$$\delta^c U_{[a^-, b^-; z]}^n = ig \int_{a^-}^{b^-} d\eta^- \, U_{[a^-, \eta^-; z]}^n \left[\delta^c \alpha(\eta^-, z)\right] U_{[\eta^-, b^-; z]}^n, \tag{6.73}$$

where, using eq. (6.53),

$$\delta^c \alpha(\eta^-, z) = \delta(\eta^- - u^-) \, G(z - \boldsymbol{u}) \, t^c. \tag{6.74}$$

Keeping in mind that

$$\int_{-\infty}^{u^-} d\eta^- \, \delta(\eta^- - u^-) = \frac{1}{2}, \tag{6.75}$$

we can now compute the first, second, and third order functional derivatives of a Wilson line:

$$\delta^c U_z = ig\,G(z-u)\,U^n_{[-\infty,u^-;z]}\,t^c\,U^n_{[u^-,\infty;z]}, \tag{6.76}$$

$$\delta^b\delta^c U_z = -\frac{g^2}{2}\,G^2(z-u)\,U^n_{[-\infty,u^-;z]}\left(t^b t^c + t^c t^b\right)U^n_{[u^-,\infty;z]}, \tag{6.77}$$

$$\delta^a\delta^b\delta^c U_z = -\frac{ig^3}{4}\,G^3(z-u)\,U^n_{[-\infty,u^-;z]}\left(t^a t^b t^c + t^a t^c t^b + t^b t^c t^a + t^c t^b t^a\right)U^n_{[u^-,\infty;z]}. \tag{6.78}$$

With these results we proceed with the computation of the (Gaussian) expectation values of the four terms in eq. (6.72):

$$d^{abc}\left\langle \mathrm{Tr}\left((\delta^a\delta^b\delta^c U^\dagger_x)U_y\right)\right\rangle_G = \frac{g^3(N_c^2-4)}{N_c}\frac{i(N_c^2-1)}{4N_c}\,G^3(x-u)$$
$$\times\left\langle \mathrm{Tr}\left(U^\dagger_x U_y\right)\right\rangle_G, \tag{6.79}$$

$$d^{abc}\left\langle \mathrm{Tr}\left(U^\dagger_x(\delta^a\delta^b\delta^c U_y)\right)\right\rangle_G = -\frac{g^3(N_c^2-4)}{N_c}\frac{i(N_c^2-1)}{4N_c}\,G^3(y-u)$$
$$\times\left\langle \mathrm{Tr}\left(U^\dagger_x U_y\right)\right\rangle_G, \tag{6.80}$$

$$d^{abc}\left\langle \mathrm{Tr}\left(3(\delta^a\delta^b U^\dagger_x)(\delta^c U_y)\right)\right\rangle_G = -\frac{g^3(N_c^2-4)}{N_c}\frac{3i}{2}\,G^2(x-u)\,G(y-u)$$
$$\times\left\langle \mathrm{Tr}\left(t^a\,U^n_{[u^-,\infty;y]}\,U^n_{[\infty,u^-;x]}\right.\right.$$
$$\left.\left.\times\,t^a\,U^n_{[u^-,-\infty;x]}\,U^n_{[-\infty,u^-;y]}\right)\right\rangle_G, \tag{6.81}$$

$$d^{abc}\left\langle \mathrm{Tr}\left(3(\delta^a U^\dagger_x)(\delta^b\delta^c U_y)\right)\right\rangle_G = \frac{g^3(N_c^2-4)}{N_c}\frac{3i}{2}\,G(x-u)\,G^2(y-u)$$
$$\times\left\langle \mathrm{Tr}\left(t^a\,U^n_{[u^-,\infty;y]}\,U^n_{[\infty,u^-;x]}\right.\right.$$
$$\left.\left.\times\,t^a\,U^n_{[u^-,-\infty;x]}\,U^n_{[-\infty,u^-;y]}\right)\right\rangle_G, \tag{6.82}$$

where we used that

$$d^{abc}t^a t^b t^c = \frac{(N_c^2-4)(N_c^2-1)}{4N_c^2}\,I, \qquad d^{abc}t^a t^b = \frac{N_c^2-4}{2N_c}\,t^c. \tag{6.83}$$

The Gaussian CGC average appearing in eqs. (6.81) and (6.82) can be simplified to [250]

$$\left\langle t^a\,U^n_{[u^-,\infty;y]}\,U^n_{[\infty,u^-;x]}\,t^a\,U^n_{[u^-,-\infty;x]}\,U^n_{[-\infty,u^-;y]}\right\rangle_G$$
$$= t^a\left\langle U^n_{[u^-,\infty;y]}\,U^n_{[\infty,u^-;x]}\right\rangle_G t^a\left\langle U^n_{[u^-,-\infty;x]}\,U^n_{[-\infty,u^-;y]}\right\rangle_G$$

$$= t^a t^a \left\langle U_x^\dagger U_y \right\rangle_G, \tag{6.84}$$

with $t^a t^a = (N_c^2 - 1)/(2N_c) I$. In the last step we used that both Gaussian averages contain $t^a t^a \propto I$ [250] and thus commute with all t^a.

Employing eqs. (6.54), (6.67), (6.79)–(6.82), and (6.84), the expectation value of the odderon operator in eq. (6.71) becomes

$$\langle \mathcal{O}(x, y) \rangle = \frac{g^6 (N_c^2 - 4)}{8i N_c^4} \int d^2 n \, T(|n|) \int \frac{d^2 q}{\pi R_N^2}$$

$$\times \operatorname{Tr} \left(\frac{i(N_c^2 - 1)}{4N_c} \left[G^3(x - u) - G^3(y - u) \right] \left\langle U_x^\dagger U_y \right\rangle_G \right.$$

$$\left. - \frac{3i}{2} \left[G^2(x - u) G(y - u) - G(x - u) G^2(y - u) \right] t^a t^a \left\langle U_x^\dagger U_y \right\rangle_G \right)$$

$$= \frac{g^6 (N_c^2 - 4)(N_c^2 - 1)}{32 N_c^5} \left\langle \operatorname{Tr} \left(U_x^\dagger U_y \right) \right\rangle_G \int d^2 n \, T(|n|)$$

$$\times \int \frac{d^2 q}{\pi R_N^2} \left[G(x - u) - G(y - u) \right]^3$$

$$= c_0 \alpha_s^3 \exp \left(-\frac{1}{4} |r|^2 Q_s^2(|b|) \ln \left(\frac{1}{|r|\Lambda} \right) \right) \int d^2 n \, T(|n|)$$

$$\times \int \frac{d^2 q}{\pi R_N^2} \ln^3 \left(\frac{|x - n - q|}{|y - n - q|} \right), \tag{6.85}$$

where the color factor c_0 is defined as[4]

$$c_0 \equiv -\frac{(N_c^2 - 4)(N_c^2 - 1)}{12 N_c^3}. \tag{6.86}$$

Following the same steps as in [260], we can perform the integration over q in eq. (6.85), which is constrained to the area of the nucleon. It was already noted in [261] that

$$\int d^2 q \, \ln^3 \left(\frac{|x - n - q|}{|y - n - q|} \right) = 0 \tag{6.87}$$

in case the integration would carry over the entire transverse plane. This result can be used to write

$$\int d^2 q \, \ln^3 \left(\frac{|x - n - q|}{|y - n - q|} \right) \theta(R_N - |q|) = -\int d^2 q \, \ln^3 \left(\frac{|x - n - q|}{|y - n - q|} \right) \theta(|q| - R_N). \tag{6.88}$$

The integral on the right-hand side can be performed if we expand the integrand in powers of $|x - n|/|q|$ and $|y - n|/|q|$. A Taylor expansion up to the second

---

4 This is the color factor given in [260]. Our calculation actually gives the slightly modified color factor $c_0' = 3c_0/N_c$. It is unclear to us where this difference comes from (for $N_c = 3$ the results are the same, of course).

derivative of the logarithm with respect to $|q|$ (this turns out to correspond to the first nontrivial order after integration), gives

$$\ln\left(|q-(x-n)|\right) = \ln(|q|) + \frac{1}{q^2}\left[\frac{(x-n)^2}{2} - q\cdot(x-n)\right] - \frac{(q\cdot(x-n))^2}{q^4}, \quad (6.89)$$

and likewise for the logarithm containing y. At this order, the integral evaluates to

$$\int d^2q \, \ln^3\left(\frac{|x-n-q|}{|y-n-q|}\right)\theta(R_N - |q|) \approx -\frac{3\pi}{4R_N^2}|r|^2\, r\cdot\tilde{b}, \quad (6.90)$$

where we used that $\phi_x \approx \phi_y \approx \phi_b$, as appropriate for a large nucleus. Furthermore, we have introduced the quantity $-\tilde{b} \equiv b - n$, which represents the impact parameter of the dipole with respect to the center of the struck nucleon (see also figure 6.4). The result in eq. (6.90) is a good approximation for $|r| < R_N$ (which means that the dipole is 'perturbatively small') and for $|\tilde{b}| \ll R_N$. Substituting the result of the integration in eq. (6.85), we obtain

$$\langle \mathcal{O}(x, y)\rangle = -\frac{3}{4R_N^4}\, c_0\alpha_s^3\, |r|^2 \exp\left(-\frac{1}{4}|r|^2 Q_s^2(|b|)\ln\left(\frac{1}{|r|\Lambda}\right)\right)$$

$$\times\, r\cdot\int d^2\tilde{b}\,\tilde{b}\, T\left(|b+\tilde{b}|\right). \quad (6.91)$$

Without any further knowledge or assumptions on the nuclear profile function, this expression cannot be simplified further.

6.5 ANGULAR CORRELATIONS IN PROTON-NUCLEUS COLLISIONS

This section provides a possible explanation for certain angular correlations that have been observed in pA collisions referred to as collectivity. After a brief introduction to collectivity, we will calculate a particular flow coefficient for dihadron production in pA collisions that originates from odderon effects. To that end, we will employ the CGC results of the previous section.

6.5.1 *Collective behavior*

Collective phenomena are common in physics and are characterized by collective behavior of a group of physical entities. A goal for theorists is to explain the observed behavior in terms of the fundamental degrees of freedom of the system. A well-known example of a collective phenomenon is collectivity in ultrarelativistic AA collisions; it refers to the observed anisotropies in the angular distributions of final-state hadrons. This phenomenon is often studied through correlations between transverse momenta of produced particle pairs. A characteristic feature of collectivity is that for relative azimuthal angles $\Delta\phi \sim 0$ and π the correlations extend over a long

range in the relative pseudorapidity $\Delta\eta$. These 'ridge-like' two-particle correlations can be characterized by Fourier harmonics in the relative azimuthal angle $\Delta\phi$ and are quantified by so-called flow coefficients. The term 'flow' is used here as collectivity in AA collisions can be explained by hydrodynamic flow models of strong interactions between final-state particles; for that reason this phenomenon is also referred to as collective flow.

Over the last decade, it has come as a surprise that the same azimuthal correlations as in AA collisions (of the same order of magnitude) have also been observed in hadron production from pp and pA collisions at various experiments at the LHC [262–283] and RHIC [284–288]. An example of measured ridge-like correlations in pA collisions is shown in figure 6.5. Whether these observations, which concern much smaller colliding systems, can be correctly described by hydrodynamic flow models remains an issue of debate. Although they can provide reasonably accurate descriptions of the data [289–297], hydrodynamics might not be appropriate for such small systems [298]. For this reason, theorists have mainly focused on other ideas, particularly explanations offered by so-called initial-state models. These models do not rely on hydrodynamic approaches, but assume that the asymmetry is created at the moment of particle production due to correlations between gluons that exist within the nuclear CGC. This category of models has received a lot of support over recent years, see e.g. [299–320]. Perhaps collectivity in high-energy pp and pA collisions should be attributed to both initial- and final-state effects [321]. The experimental and theoretical status of collectivity in small colliding systems is reviewed in [298].
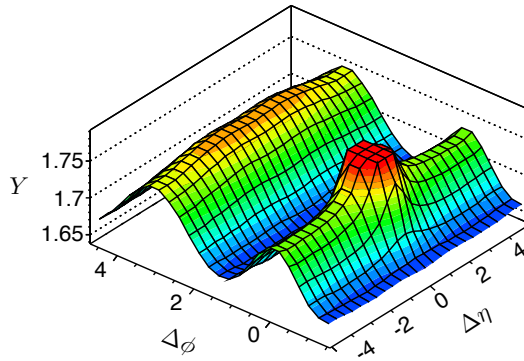
Figure 6.5: An illustration of ridge-like structures as measured by the ATLAS experiment at CERN. The two-particle correlation function $Y(\Delta\phi, \Delta\eta)$ is shown for proton-lead collisions with a CM energy of $\sqrt{s} = 5.02$ TeV [269].

In the next subsection we will study azimuthal anisotropies in the production of two hadrons in pA collisions. We will focus on generating nonzero odd-harmonic two-particle correlations entirely from initial-state effects, which has been studied less than generating even-harmonic correlations. Odd harmonics from initial-state

effects have been studied in different settings in [322–327]. The cross section for hadron production from an incoming quark that scatters off a nucleus is proportional to the Fourier transform of the expectation value of the dipole operator S. The real part (the pomeron) of S produces even harmonics, while the imaginary part (the odderon) generates odd harmonics. We extend the work of [328] and show that the odderon Wigner distribution at small x can give rise to a nonzero v_1 flow coefficient that characterizes the so-called directed flow (v_1 is associated to the first odd-harmonic contribution to the cross section). Compared to previous literature, the source of odd azimuthal correlations in our model is the relative orientation between the impact parameter and transverse momentum of the produced particle for a nuclear CGC that is inhomogeneous in the transverse plane. The same mechanism was discussed in [329] for the case of even harmonics and for single-inclusive hadron production. Although v_1 can arise from C-odd fluctuations in a rotationally noninvariant nucleus [322], we show that it can also arise from C-odd correlations without breaking of rotational symmetry.

6.5.2 *Directed flow in dihadron production*

Let us consider dihadron production from high-energy pA collisions, $pA \rightarrow H_1 H_2 X$, where the proton and nucleus are respectively moving along the minus and plus directions. At forward rapidities (in the direction of the proton), the produced hadrons probe the proton at large longitudinal momentum fractions, while the nucleus is probed at small longitudinal momentum fractions $x \ll 1$, see e.g. [199, 231, 330]. This means that in the forward-rapidity region the proton is 'dilute' and mainly consists of valence quarks, whereas the nucleus is characterized by a high gluon density that can be modeled by a CGC. Furthermore, it has been shown that at forward rapidities the cross section may be dominated by contributions from *double parton scattering* rather than from single parton scattering [331–333]. For that reason, we will study contributions from the former and consider two quarks coming from the proton that scatter (independently) off the CGC system, producing two final-state hadrons. The distribution of a pair of quarks inside the proton can be described by the so-called (collinear) double parton distribution (DPD) $F_p(x_1, x_2, \boldsymbol{b}_1 - \boldsymbol{b}_2)$, where x_1 and x_2 denote the longitudinal momentum fractions of the quarks and $\boldsymbol{b}_1$ and $\boldsymbol{b}_2$ denote their scattering positions with respect to the center of the nucleus (for more information on DPDs, see e.g. [334]). The description of 'dilute-dense' collisions by a collinear PDF on the dilute proton side and a CGC (that includes transverse momentum effects) on the dense nucleus side, is called the hybrid approach [335].

The transverse momenta of the final-state quarks are acquired through multiple rescatterings off the gluon field of the nucleus which are resummed into Wilson lines. The operator that describes the scattering of two quarks off the CGC at the level of the cross section is given by the product of two dipole operators, $S(\boldsymbol{x}_1, \boldsymbol{y}_1) S(\boldsymbol{x}_2, \boldsymbol{y}_2)$. From the physical interpretation of the dipole operator given at the beginning of

section 6.4, it follows that the picture of two quarks scattering off a nucleus is mathematically equivalent to the scattering of two color dipoles off a nucleus, see e.g. [242] for more details.

Assuming factorization, one can write the cross section for the production of two quarks with momenta k_1 and k_2 and rapidities y_1 and y_2 as (see e.g. [325]):[5]

$$\frac{d\sigma^{pA}}{dy_1 dy_2 d^2k_1 d^2k_2} \propto \int d^2\boldsymbol{b}_1 \int d^2\boldsymbol{b}_2 \, F_p(x_1,x_2,\boldsymbol{b}_1-\boldsymbol{b}_2) \int \frac{d^2\boldsymbol{r}_1}{(2\pi)^2} \, e^{-ik_1 \cdot r_1} \int \frac{d^2\boldsymbol{r}_2}{(2\pi)^2} \, e^{-ik_2 \cdot r_2}$$
$$\times \left\langle S\left(\boldsymbol{b}_1 + \tfrac{\boldsymbol{r}_1}{2}, \boldsymbol{b}_1 - \tfrac{\boldsymbol{r}_1}{2}\right) S\left(\boldsymbol{b}_2 + \tfrac{\boldsymbol{r}_2}{2}, \boldsymbol{b}_2 - \tfrac{\boldsymbol{r}_2}{2}\right)\right\rangle, \tag{6.92}$$

where $\boldsymbol{r}_1 \equiv \boldsymbol{x}_1 - \boldsymbol{y}_1$ and $\boldsymbol{r}_2 \equiv \boldsymbol{x}_2 - \boldsymbol{y}_2$ are the dipole orientations, and $\boldsymbol{b}_1 \equiv (\boldsymbol{x}_1 + \boldsymbol{y}_1)/2$ and $\boldsymbol{b}_2 \equiv (\boldsymbol{x}_2 + \boldsymbol{y}_2)/2$ are the corresponding impact parameters. We will work in the large-N_c limit and consider only the leading contributions in $1/N_c$ to the azimuthal correlations. At large N_c, the expectation value of the product of traces factorizes (see e.g. [325]):

$$\left\langle S\left(\boldsymbol{b}_1 + \tfrac{\boldsymbol{r}_1}{2}, \boldsymbol{b}_1 - \tfrac{\boldsymbol{r}_1}{2}\right) S\left(\boldsymbol{b}_2 + \tfrac{\boldsymbol{r}_2}{2}, \boldsymbol{b}_2 - \tfrac{\boldsymbol{r}_2}{2}\right)\right\rangle$$
$$\approx \left\langle S\left(\boldsymbol{b}_1 + \tfrac{\boldsymbol{r}_1}{2}, \boldsymbol{b}_1 - \tfrac{\boldsymbol{r}_1}{2}\right)\right\rangle \left\langle S\left(\boldsymbol{b}_2 + \tfrac{\boldsymbol{r}_2}{2}, \boldsymbol{b}_2 - \tfrac{\boldsymbol{r}_2}{2}\right)\right\rangle. \tag{6.93}$$

The neglected corrections have been shown to give rise to azimuthal asymmetries in [316, 317, 325]; here we demonstrate that odd azimuthal asymmetries can arise already at leading power in $1/N_c$. In the large-N_c limit, the cross section in eq. (6.92) can be written as a convolution of two dipole Wigner distributions:

$$\frac{d\sigma^{pA}}{dy_1 dy_2 \, d^2k_1 d^2k_2} \propto \int d^2\boldsymbol{b}_1 \int d^2\boldsymbol{b}_2 \, F_p(x_1, x_2, \boldsymbol{b}_1 - \boldsymbol{b}_2)$$
$$\times xW(x, \boldsymbol{b}_1, \boldsymbol{k}_1) \, xW(x, \boldsymbol{b}_2, \boldsymbol{k}_2), \tag{6.94}$$

where the dipole Wigner distribution W is defined in the CGC theory as

$$xW(x, \boldsymbol{b}, \boldsymbol{k}) \equiv \int \frac{d^2\boldsymbol{r}}{(2\pi)^2} \, e^{-ik \cdot r} \left\langle S\left(\boldsymbol{b} + \tfrac{\boldsymbol{r}}{2}, \boldsymbol{b} - \tfrac{\boldsymbol{r}}{2}\right)\right\rangle. \tag{6.95}$$

We refer to [4] for a motivation of this definition.

The source for azimuthal angular correlations (both even and odd) in the cross section of this process is the relative orientation between the transverse momenta of the produced particles and their impact parameters, which is encoded in the dipole Wigner distributions. To extract the angular correlations, we parametrize the Wigner distribution of eq. (6.95) in terms of the different harmonic contributions [221]:

$$xW(x, \boldsymbol{b}, \boldsymbol{k}) = x\mathcal{W}_0(x, \boldsymbol{b}^2, \boldsymbol{k}^2) + 2\cos(\phi_{\boldsymbol{b}} - \phi_{\boldsymbol{k}}) \, x\mathcal{W}_1(x, \boldsymbol{b}^2, \boldsymbol{k}^2)$$
$$+ 2\cos(2(\phi_{\boldsymbol{b}} - \phi_{\boldsymbol{k}})) \, x\mathcal{W}_2(x, \boldsymbol{b}^2, \boldsymbol{k}^2) + \dots \tag{6.96}$$

---

5 For simplicity we shall not consider explicitly the fragmentation of the produced partons into hadrons.

The so-called elliptic flow resulting from the elliptic Wigner distribution $\mathcal{W}_2$ has been studied in e.g. [221,328,329,336–339]. As explained in subsection 6.2.1 for the case of dipole GTMDs, the even and odd harmonics in eq. (6.96) constitute, respectively, the pomeron and odderon contributions to the dipole Wigner distribution W. In this analysis we will focus on the odd-harmonic correlations that are generated by the odderon operator. We will only consider the first odd contribution explicitly, i.e. the one associated to the function $\mathcal{W}_1$, which we will refer to as the 'odderon Wigner distribution'.

For simplicity, we assume that the dependence on $\boldsymbol{b}_1 - \boldsymbol{b}_2$ in the DPD factorizes and we take a Gaussian ansatz for the transverse density profile (see e.g. [102,340]):

$$F_p(x_1, x_2, \boldsymbol{b}_1 - \boldsymbol{b}_2) = \frac{1}{4\pi R_N^2} f_p(x_1, x_2) \, e^{-\frac{(b_1 - b_2)^2}{4R_N^2}} \, . \qquad (6.97)$$

Using eqs. (6.96) and (6.97) in the cross section (6.94), and integrating over the angles ϕ_{b_1} and ϕ_{b_2}, we obtain

$$\frac{d\sigma^{pA}}{dy_1 dy_2 \, d^2\boldsymbol{k}_1 d^2\boldsymbol{k}_2} \propto \frac{\pi}{8R_N^2} f_p(x_1, x_2) \int_0^\infty db_1^2 \int_0^\infty db_2^2 \, e^{-\frac{b_1^2 + b_2^2}{4R_N^2}}$$

$$\times \left[2I_0 \left(\frac{b_1 b_2}{2R_N^2} \right) x\mathcal{W}_0(x, \boldsymbol{b}_1^2, \boldsymbol{k}_1^2) \, x\mathcal{W}_0(x, \boldsymbol{b}_2^2, \boldsymbol{k}_2^2) \right.$$

$$+ 4\cos(\phi_{k_1} - \phi_{k_2}) \, I_1 \left(\frac{b_1 b_2}{2R_N^2} \right) x\mathcal{W}_1(x, \boldsymbol{b}_1^2, \boldsymbol{k}_1^2) \, x\mathcal{W}_1(x, \boldsymbol{b}_2^2, \boldsymbol{k}_2^2)$$

$$\left. + 4\cos(2(\phi_{k_1} - \phi_{k_2})) \, I_2 \left(\frac{b_1 b_2}{2R_N^2} \right) x\mathcal{W}_2(x, \boldsymbol{b}_1^2, \boldsymbol{k}_1^2) \, x\mathcal{W}_2(x, \boldsymbol{b}_2^2, \boldsymbol{k}_2^2) \right]$$

$$+ \dots, \qquad (6.98)$$

where $b_{1,2} \equiv |\boldsymbol{b}_{1,2}|$ and I_n is the nth modified Bessel function of the first kind. We note that the odderon Wigner distribution $\mathcal{W}_1$ of the nucleus gives rise to an odd azimuthal correlation between the transverse momenta of the produced particles.

The standard way of quantifying azimuthal correlations in particle production is through the flow coefficients v_n [341]. For two-particle correlations, we have

$$v_n(\boldsymbol{k}, \boldsymbol{k}^{\text{ref}}) \equiv \frac{V_n(\boldsymbol{k}, \boldsymbol{k}^{\text{ref}})}{\sqrt{V_n(\boldsymbol{k}^{\text{ref}}, \boldsymbol{k}^{\text{ref}})}} \, , \qquad (6.99)$$

where the coefficients V_n are obtained from the decomposition of the cross section into Fourier modes in the relative azimuthal angle between the produced particles,

$$\frac{d\sigma^{pA}}{dy_1 dy_2 \, d^2\boldsymbol{k}_1 d^2\boldsymbol{k}_2} \propto 1 + \sum_{n=1}^\infty 2V_n(\boldsymbol{k}_1, \boldsymbol{k}_2) \cos(n(\phi_{k_1} - \phi_{k_2})), \qquad (6.100)$$

and k^{ref} is a reference momentum corresponding to an experimental reference bin. From eqs. (6.98) and (6.100) it follows that the first odd coefficient, which characterizes the directed flow, is given by

$$
V_1(k_1, k_2) = \frac{\int_0^\infty db_1^2 \int_0^\infty db_2^2 \, e^{-\frac{b_1^2 + b_2^2}{4R_N^2}} \, I_1\left(\frac{b_1 b_2}{2R_N^2}\right) x\mathcal{W}_1(x, b_1^2, k_1^2) \, x\mathcal{W}_1(x, b_2^2, k_2^2)}{\int_0^\infty db_1^2 \int_0^\infty db_2^2 \, e^{-\frac{b_1^2 + b_2^2}{4R_N^2}} \, I_0\left(\frac{b_1 b_2}{2R_N^2}\right) x\mathcal{W}_0(x, b_1^2, k_1^2) \, x\mathcal{W}_0(x, b_2^2, k_2^2)} \, . \quad (6.101)
$$

In the following we will derive an explicit expression for V_1 in the CGC theory for small (but nonzero) x. We will make use of the CGC averages that we calculated in section 6.4. The expression for the angular-independent contribution, $\mathcal{W}_0$, can be obtained from the CGC average of the pomeron operator given in eq. (6.67):

$$
x\mathcal{W}_0(x, b^2, k^2) = \int \frac{d^2r}{(2\pi)^2} \, e^{-ik\cdot r} \left\langle \mathcal{P}\left(b + \tfrac{r}{2}, b - \tfrac{r}{2}\right)\right\rangle
$$

$$
= \frac{1}{2\pi} \int_0^\infty dr\, r\, J_0(k r) \, \exp\left(-\frac{1}{4}r^2 Q_s^2(b) \ln\left(\frac{1}{r\Lambda}\right)\right), \quad (6.102)
$$

where $k \equiv |k|$, $r \equiv |r|$, and J_n is the nth Bessel function of the first kind.

Next, we turn to the CGC calculation of the odderon Wigner function $\mathcal{W}_1$. This function can be obtained from the expectation value of the odderon operator given in eq. (6.91), which comes with the nuclear profile function $T = T\left(|b + \tilde{b}|\right)$. Since we consider angular asymmetries generated by the inhomogeneity of the nucleus in the transverse plane, which is naturally larger at the edge of the nucleus than at its center, our main focus is on *peripheral* collisions. As $\tilde{b} \equiv |\tilde{b}|$ is confined to the area of the nucleon, those type of collisions are characterized by $\tilde{b} \ll b$. Hence, we can expand the nuclear profile function in powers of $\tilde{b}/b$:

$$
T\left(|b + \tilde{b}|\right) = \left[1 + \tilde{b}^i \frac{\partial}{\partial b^i} + \frac{1}{2}\tilde{b}^i\tilde{b}^j \frac{\partial^2}{\partial b^i \partial b^j} + \frac{1}{3!}\tilde{b}^i\tilde{b}^j\tilde{b}^k \frac{\partial^3}{\partial b^i \partial b^j \partial b^k} + \dots\right] T(b). \quad (6.103)
$$

Plugging this expansion (up to second order in $\tilde{b}/b$) in eq. (6.91) and performing the integration over $\tilde{b}$, we obtain:

$$
\left\langle \mathcal{O}\left(b + \tfrac{r}{2}, b - \tfrac{r}{2}\right)\right\rangle = -\frac{3\pi}{16} c_0 \alpha_s^3 r^2 \exp\left(-\frac{1}{4}r^2 Q_s^2(b) \ln\left(\frac{1}{r\Lambda}\right)\right) r^i \frac{\partial}{\partial b^i} T(b)
$$

$$
= -\frac{3\pi}{16} c_0 \alpha_s^3 r^3 \exp\left(-\frac{1}{4}r^2 Q_s^2(b) \ln\left(\frac{1}{r\Lambda}\right)\right)
$$

$$
\times \cos(\phi_b - \phi_r)\, T'(b). \quad (6.104)
$$

Note that only the linear term in the expansion survives the integration; it has given rise to a $\cos(\phi_b - \phi_r)$ correlation.[6] From the Fourier transform of eq. (6.104),

$$\int \frac{d^2 r}{(2\pi)^2} e^{-i\mathbf{k}\cdot\mathbf{r}} \left\langle \mathcal{O}\left(\mathbf{b} + \tfrac{\mathbf{r}}{2}, \mathbf{b} - \tfrac{\mathbf{r}}{2}\right) \right\rangle = \frac{3}{32} i c_0 \alpha_s^3 \cos(\phi_b - \phi_k)\, T'(b) \int_0^\infty dr\, r^4 J_1(kr)$$

$$\times \exp\left(-\frac{1}{4} r^2 Q_s^2(b) \ln\left(\frac{1}{r\Lambda}\right)\right), \qquad (6.105)$$

combined with eq. (6.96), it follows that the odderon Wigner distribution is given by

$$x\mathcal{W}_1(x, \mathbf{b}^2, \mathbf{k}^2) = -\frac{3}{64} c_0 \alpha_s^3\, T'(b) \int_0^\infty dr\, r^4 J_1(kr)$$

$$\times \exp\left(-\frac{1}{4} r^2 Q_s^2(b) \ln\left(\frac{1}{r\Lambda}\right)\right). \qquad (6.106)$$

From this expression we infer that the C-odd interaction described by $\mathcal{W}_1$ vanishes for a nucleus that is homogeneous in the transverse plane as then $T'(b) = 0$ (this is *not* the case for the C-even exchange described by $\mathcal{W}_0$).

With the results in eqs. (6.102) and (6.106) we can calculate the directed flow coefficient v_1, defined in eq. (6.99). To numerically estimate its size in our CGC model, we need to assume a certain nuclear profile. For a rotationally symmetric nucleus, the nuclear profile function T can be written as follows:

$$T(b) = \int_{-\infty}^\infty dz\, \rho_A\left(\sqrt{\mathbf{b}^2 + z^2}\right), \qquad (6.107)$$

where ρ_A is the density of nucleons in the nucleus. For a large nucleus one typically takes a Woods-Saxon profile [342],

$$\rho_A(r) = \frac{N_A}{1 + e^{(r - R_A)/\delta}}, \qquad (6.108)$$

with $R_A = (1.12\,\text{fm})A^{1/3}$ the nuclear radius, $\delta = 0.54$ fm the width of the 'nuclear edge' [329], and N_A is a normalization factor such that $\int d^3\vec{r}\, \rho_A(\vec{r}) = A$, the number of nucleons inside the nucleus. We will calculate v_1 for lead and copper nuclei with, respectively, $A = 208$ and $A = 63$, and we will use the following approximation for the nuclear profile function:

$$T(b) \approx \beta \left(1 - \frac{1}{1 + e^{-(b/b_0 - \gamma)}}\right), \qquad (6.109)$$

with $\beta = 2.13\,\text{fm}^{-2}$ and $\gamma = 5.4$ for lead, and $\beta = 1.38\,\text{fm}^{-2}$ and $\gamma = 3.5$ for copper, and $b_0 = 1$ fm. This form approximates the Woods-Saxon distribution sufficiently

---

6 The quadratic term in the expansion gives rise to a $\cos(2(\phi_b - \phi_r))$ correlation in the *real* part of the dipole operator and gives rise to elliptic flow [329]. The third derivative of the profile function, combined with higher-order corrections to eq. (6.91), can give rise to a nonzero v_3 flow coefficient.

well for our purposes. To save computation time, we take $b_{1,2} \gtrsim R_N \sim 1$ fm and we only consider the leading terms in the asymptotic expansions of the Bessel functions $I_n \left(\frac{b_1 b_2}{2 R_N^2} \right)$.

In figure 6.6 we have plotted v_1 as a function of k for $k^{\text{ref}} = 0.80$ GeV (the behavior and magnitude of v_1 are similar for other choices of k^{ref}). We have taken $\alpha_s = 0.3$ and $\Lambda = \Lambda_{\text{QCD}} = 0.24$ GeV [23]. The two curves for lead and copper start at $k = 0.14$ GeV (unphysical fluctuations are present at lower values of k, where v_1 should go to zero as $k \to 0$). A word of caution is required, however, for small values of k: we expect the calculation to break down for small momenta $k < \Lambda_{\text{QCD}}$ as our result only applies to perturbatively small dipoles for which $r < R_N \sim 1/\Lambda_{\text{QCD}}$.
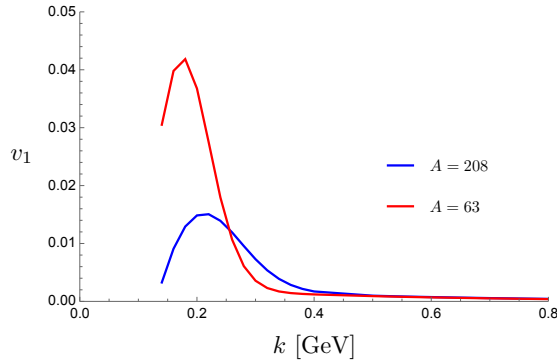
Figure 6.6: The directed flow $v_1(k)$ for a lead nucleus with $A = 208$ (blue curve) and for a copper nucleus with $A = 63$ (red curve). For the reference momentum we have taken $k^{\text{ref}} = 0.80$ GeV.

Our numerical estimate for lead with $A = 208$ can be compared to a v_1 measurement by the ATLAS experiment at CERN [269] (see also figure 6.5), keeping in mind that their observable does not correspond exactly to ours. We conclude that the sign change observed in [269] is not present in our result and that the behavior at large k is different.[7] However, our goal here is *not* to describe the data, but rather to illustrate that our simple model can produce a nonzero v_1 that is of the right order of magnitude. Indeed, we find that the magnitude of v_1 for $k \sim 0.2$ GeV is of the same order as the maximal value observed in the data for $k \lesssim 1$ GeV [269]. We emphasize that we have not made use of any fit parameters in our analysis.

With regards to the dependence on A, we note that v_1 decreases as A increases. This observation is not surprising given that we have computed $\mathcal{W}_1$ in the MV model *extended* by a cubic weight. As already mentioned in section 6.4, the cubic weight is subleading in $1/A$ compared to the Gaussian weight (which has been used to

---

7 Note that, as opposed to the analysis in [269], we have *not* included the sign function $\text{sgn}(k^{\text{ref}} - k^0)$ in the definition of v_1 (the sign function is defined to be negative for $k^{\text{ref}} < k^0 = 1.5$ GeV and is positive otherwise).

calculate $\mathcal{W}_0$) and therefore our observable is suppressed in the large-A limit. Note that the precise dependence of our result on A is not straightforward to extract from our expressions due to the nontrivial dependence on Q_s.

Finally, let us comment on particle production at more central rapidities. We have derived the directed flow coefficient for the production of two hadrons by two incoming *quarks* from the proton that scatter off a nuclear CGC. This description is valid at forward rapidity, such that the proton consists of valence quarks only. Moving away from the forward-rapidity region, one also needs to account for the presence of sea quarks and antiquarks inside the proton. The scattering of an antiquark off a nucleus is given by the complex conjugate of the quark's dipole operator, which implies that $\mathcal{W}_1^q = -\mathcal{W}_1^{\bar{q}}$. Hence, if one assumes that away from the valence region the proton consists of an equal number of quarks and antiquarks with exactly the same DPDs for qq, $\bar{q}\bar{q}$, and $q\bar{q}$ pairs, then the sum over all parton pairs that could scatter off a nucleus would not give any odd harmonics.[8] However, it is known that for values of $x \lesssim 10^{-1}$, the DPDs *differ* for different pairs of partons with an x-dependent width in the Gaussian model [343, 344]. In general, the proton DPD takes the following form:

$$F_p(x_1, x_2, \boldsymbol{b}_1 - \boldsymbol{b}_2) = \frac{1}{4\pi h_{ab}(x_1, x_2)} f_{ab}(x_1, x_2) \, e^{-\frac{(\boldsymbol{b}_1 - \boldsymbol{b}_2)^2}{4 h_{ab}(x_1, x_2)}}, \qquad (6.110)$$

where the labels a, b refer to quarks or antiquarks. The precise form of h_{ab}, including numerical values for the relevant parameters, can be found in [344]. Because the proton DPDs differ for different combinations of quarks and antiquarks, our mechanism would *also* produce nonzero odd azimuthal correlations away from the forward-rapidity region.

---

8 The scattering of gluons off a nuclear CGC does not contribute to odd harmonics in our model as the *adjoint* dipole operator is purely real.

7

SUMMARY AND CONCLUSIONS

In this thesis we have studied multidimensional quark and gluon PDFs as probes of the internal structure of hadrons. In high-energetic collisions involving hadrons, multidimensional PDFs are often related to asymmetries in the angular distributions of the produced particles. Thus, measuring a certain angular harmonic contribution to a cross section can imply the presence of certain partonic correlations inside hadrons. More knowledge on the rich dynamics of the hadronic constituents can bring us closer to understanding the origins of for example hadron masses and spin angular momenta. Moreover, it can allow for more accurate descriptions of processes in which QCD effects play an important role.

In chapter 2 we introduced the quark-quark correlator as the Fourier transform of a bilocal hadronic matrix element containing two quark fields. This quantity forms the starting point for any study on multidimensional PDFs. Correlators can depend on different sets of kinematical variables, depending on which projection or limit is considered. The most general correlator that we have studied is the GTMD correlator that features an off-forward matrix element; it depends on the variables x, k, and Δ. In the forward limit ($\Delta \to 0$), the GTMD correlator reduces to the TMD correlator which depends on x and k. Upon integration over the partonic transverse momentum k, the TMD correlator reduces to the collinear correlator that depends only on the longitudinal momentum fraction x.

Since correlators encode nonperturbative physics describing the dynamics of quarks and gluons inside hadrons, they cannot be calculated using perturbation theory. Instead, they are typically parametrized in terms of a priori unknown functions, the PDFs, that can be fitted to data. In general, correlators satisfy certain field theoretical constraints that can limit the number of allowed structures in such parametrizations or impose constraints on the PDFs. Both parton and hadron spin effects give rise to a wealth of different TMDs and play a central role in TMD phenomenology. For example, the quark-quark TMD correlator, constrained by Hermiticity, parity, and time reversal requirements, can be parametrized (at leading twist) in terms of two, six, and ten quark TMDs for unpolarized, vector polarized, and tensor polarized hadrons, respectively.

An important ingredient of any hadronic correlator is the gauge link, a path ordered exponential of the gluon field that ensures color gauge invariance. The presence of a gauge link in correlators has important consequences for TMDs, GTMDs, and Wigner distributions that all depend on the partonic transverse momentum. For those functions, in contrast to for example collinear PDFs, the integration paths of the associated gauge links are not unique but depend on the process under consid-

eration. Consequently, those functions themselves become nonuniversal. The most well-known example of the nonuniversality of TMDs is the expected sign change for T-odd quark TMDs between DY and SIDIS, which results from the fact that these processes come with past-pointing ($[-]$) and future-pointing ($[+]$) gauge links, respectively. Experimental verification of this sign flip is ongoing and forms a vital test of the TMD formalism.

In part 1 we looked in detail at the DY process, which is for example essential for the description of pp collisions at the LHC. We considered the DY cross section differential in the transverse momentum of the produced lepton pair (in general, the less inclusive a process, the more sensitive it is to multidimensional PDFs). For small transverse momenta, the DY cross section satisfies a TMD factorization theorem, which means that it can be factorized in terms of a perturbatively calculable hard scattering factor and two TMDs. Factorization formulas provide the foundation for QCD phenomenology. TMD factorization for DY scattering was established by Collins, Soper, and Sterman (CSS) at leading power in the small parameter Λ/Q and to all orders in perturbation theory. In chapter 3 we reviewed various elements of their proof.

The CSS proof is based on a momentum regions analysis. Graphs can receive leading-power contributions from loop momenta with hard, collinear, soft, or Glauber scalings. The latter type of scalings represents virtual modes characterized by $|\ell^+\ell^-| \ll \ell^2 \ll Q^2$. To reconstruct a full graph from its leading-power contributions, the leading regions need to be summed over and overlap contributions can be accounted for by the Collins subtraction method. The DY graphs can be organized in terms of subgraphs that each characterize a particular momentum scaling. The goal of the factorization procedure is to prove that those subgraphs, that are a priori connected through soft and collinear gluon exchanges, become disconnected (at leading power) upon summing over all contributions. By virtue of Ward identities, the hard, collinear, and soft subgraphs factorize upon summing over all possible soft and collinear gluon attachments; the Grammer-Yennie approximation plays an essential role in this. The latter approximation, however, is not valid for Glauber modes. To establish factorization, it needs to be shown that Glauber effects ultimately cancel out or can be captured by other region contributions. Through suitable contour deformations, some (but not all) Glauber exchanges can be absorbed into soft or collinear regions. As it turns out, upon summing over all possible final-state cuts all Glauber modes become untrapped, allowing for contour deformations into different momentum regions; the cancellation of final-state poles that is responsible for this is related to unitarity of the theory. This means that all Glauber partons can eventually be absorbed into soft or collinear regions where Grammer-Yennie approximations can be applied. Hence, upon summing over all graphs, cuts, and leading regions, the DY cross section factorizes into two TMD parton densities (these arise from combining soft and collinear subgraphs) and a hard scattering factor.

Over recent years, there has been debate about possible color-entanglement effects in DY that arise from double T-odd contributions. These effects would go beyond

the expected sign flip between DY and SIDIS and would imply a loophole in the CSS proof for T-odd TMDs that involve polarization. An example of a double T-odd contribution to the DY cross section is the $\cos(2\phi)$ angular modulation associated to two BM functions (the dBM effect). As color-entanglement effects could lead to modified color factors in the (factorized) cross section, or even to a breakdown of factorization, we have reexamined this issue in detail.

By largely following the steps of the CSS factorization proof, we investigated in chapter 4 by an explicit model calculation whether color-entanglement effects persist in the dBM contribution to the cross section after properly summing over all relevant graphs, cuts, and leading regions. This study was performed at the lowest order in perturbation theory for which color-entanglement effects have been argued to be present – this is $\mathcal{O}(\alpha_s^2)$, corresponding to the level of two gluon exchanges. At this order we identified various graphs that contribute to the dBM term and determined four common leading regions of loop momenta involving collinear and Glauber scalings only. Upon summing over all contributions, we find that color-entanglement effects completely cancel on a region-by-region basis; the final result conforms with the predictions of the CSS factorization theorem. Furthermore, the Glauber region turns out to play an important role: depending on the choice of rapidity regulators, it is possible to assign the full contribution to the dBM term at the considered order to the region in which the two gluons have Glauber scalings.

Similarly to the CSS proof, the two mechanisms that lie at the heart of our findings are the non-Abelian Ward identity and the unitarity cancellation of final-state poles. Furthermore, following the summation over all possible final-state cuts, the Glauber gluons can be absorbed into other region contributions, consistent with the CSS works. More specifically, we find that the Glauber momenta can be deformed into the collinear regions. Although our model calculation was performed at $\mathcal{O}(\alpha_s^2)$, we have no reasons to doubt that the same physical principles also drive the disentanglement of different color structures at higher orders in perturbation theory. Finally, we expect our results to go beyond our model setup and to apply also to DY in QCD.

In part II of this thesis we studied multidimensional gluon distributions. Compared to the quark case, the extraction of gluon functions remains challenging. In this respect the possible construction of the EIC is very promising, as it could lead to unraveling the gluon content of (polarized) hadrons by measuring a wide variety of gluon observables in the small-x region where gluons dominate over quarks. Today, no gluon TMDs other than the unpolarized one have been fitted to data yet.

To facilitate the study of gluon observables, we provided in chapter 5 an instructive parametrization of the gluon-gluon TMD correlator in terms of leading-twist gluon TMDs that appear in the description of spin-0, spin-$\frac{1}{2}$, and spin-1 hadrons. The ones related to tensor polarization are new. As from the experimental side the gluon functions are largely unknown, we have also derived model-independent inequalities for combinations of TMDs that follow from the operator structure of the correlator. These bounds may be used to estimate maximal contributions of certain functions to observables.

7

In chapter 6 we focused on gluon distributions at small x. We only looked at one particular gauge link structure, namely the dipole-type structure $[+, -]$. We have shown that for this gauge link structure the gluon-gluon GTMD correlator is related to the correlator of a Wilson loop ($\square$) in the limits of small x and ξ. For unpolarized hadrons, the off-forward Wilson loop correlator can be parametrized in terms of a single GTMD. The fact that the gluon-gluon GTMD correlator is parametrized in terms of four GTMDs implies that at vanishing x and ξ these functions all become related to each other. Hence, one expects that when hadrons are probed at very high energies the picture of gluon GTMDs becomes very simple. By taking the forward limit of these results, we arrive at similar conclusions for dipole-type TMDs: for each type of hadron polarization we find that in the small-x limit dipole-type TMDs either vanish or become proportional to each other. The nonvanishing gluon TMDs and GTMDs are expected to behave as $1/x$, or a slightly modified power after resummation of large logarithms in $1/x$ from α_s corrections.

For transversely polarized hadrons we identified a C-odd TMD contribution (e_T), which has been referred to as the spin-dependent odderon. We also derived an upper bound for this function. In contrast to the TMD case, the dipole-type GTMDs also contain odderon contributions in the unpolarized situation; these contributions originate from correlations between the off-forwardness and the gluon momentum in the transverse plane. The same is true for dipole Wigner distributions, with the off-forwardness replaced by its Fourier conjugate, the impact parameter $\boldsymbol{b}$.

Finally, we considered partonic correlations within the nuclear wave function as sources for the observed collective behavior in ultrarelativistic pA collisions. The high density of gluons in the nuclei can be modeled by a CGC, which is an effective color field generated by static color sources (the valence quarks). In the MV model, the expectation value of an operator in the CGC theory can be obtained by performing a functional average over all possible configurations of color sources in the nucleus. For dihadron production through two incoming quarks from the proton that scatter off a nuclear CGC, the cross section is proportional to a convolution of two dipole Wigner distributions. These distributions can give rise to both even and odd harmonics in the two-particle angular correlations, the latter being due to odderon effects. Within a simple CGC model (with a cubic weight functional) and in the large-N_c limit, we have calculated the directed flow coefficient v_1 that quantifies the contribution of the first odd harmonic with respect to the angular-independent contribution. For peripheral collisions involving a lead nucleus described by a Woods-Saxon-like profile, we find a v_1 at the percent level, which is the same order of magnitude as is observed experimentally at the LHC. The shape of v_1, however, is not well-described by our model as it is probably not realistic enough due to many simplifying assumptions.

A

CONVENTIONS

Below we list the most important conventions that are used throughout this thesis:

- We adopt natural units where $\hbar = c = 1$. With this choice, mass is the only dimension that is left: $[\text{mass}] = [\text{energy}] = [\text{length}]^{-1} = [\text{time}]^{-1}$, where the square brackets denote the mass dimension.

- Spacetime components are labeled by Greek indices and the roman indices $i, j, k, l = 1, 2$ refer to transverse components only. For Dirac indices we use the Greek letters $\alpha, \beta = 1, 2, 3, 4$, and the roman indices $a, b, c = 1, 2, \ldots, 8$ label the generators of the $SU(3)$ Lie algebra, unless explicitly stated otherwise.

- For the Minkowski metric g we adopt the 'mostly minus' convention, i.e. its nonzero components read $g^{00} = -g^{11} = -g^{22} = -g^{33} = 1$.

- A four-vector a can be represented in terms of its light-cone components as $a = (a^+, a^-, \boldsymbol{a})$, where $a^\pm \equiv (a^0 \pm a^3)/\sqrt{2}$ and $\boldsymbol{a} \equiv (a^1, a^2)$. A dot product between two four-vectors a and b is given by $a \cdot b = a^+ b^- + a^- b^+ - \boldsymbol{a} \cdot \boldsymbol{b}$.

- The lightlike vectors $\bar{n}$ and n are defined as $\bar{n} \equiv (1, 0, \boldsymbol{0})$ and $n \equiv (0, 1, \boldsymbol{0})$, such that $n \cdot \bar{n} = 1$. Any four-vector a can be decomposed in terms of $\bar{n}$ and n as $a^\mu = a^+ \bar{n}^\mu + a^- n^\mu + a_T^\mu$, where we have defined $a_T \equiv (0, 0, \boldsymbol{a})$ (note that $a_T^2 = -\boldsymbol{a}^2$).

- Given two four-vectors a and b, their symmetrized product is defined as $a^{\{\mu} b^{\nu\}} \equiv a^\mu b^\nu + a^\nu b^\mu$ and their antisymmetrized product as $a^{[\mu} b^{\nu]} \equiv a^\mu b^\nu - a^\nu b^\mu$.

- For a four-vector a and a rank-two tensor b we define $\bar{a}^\mu \equiv \delta^{\mu\nu} a_\nu$ and $\bar{b}^{\mu\nu} \equiv \delta^{\mu\rho} \delta^{\nu\sigma} b_{\rho\sigma}$.

- The Levi-Civita symbol ϵ is fixed by taking $\epsilon^{0123} = \epsilon^{-+12} = 1$. Furthermore, we employ the notation $\epsilon^{abcd} \equiv \epsilon^{\mu\nu\rho\sigma} a_\mu b_\nu c_\rho d_\sigma$.

- The metric tensor in transverse space is defined as $g_T^{\mu\nu} \equiv g^{\mu\nu} - \bar{n}^{\{\mu} n^{\nu\}}$ and has nonzero components $g_T^{11} = g_T^{22} = -1$ (note that $g_T^{ij} = -\delta_T^{ij}$).

- The Levi-Civita symbol in the transverse plane is given by $\epsilon_T^{\mu\nu} \equiv \epsilon^{\mu\nu-+}$ with nonzero components $\epsilon_T^{12} = -\epsilon_T^{21} = 1$. We also define $\tilde{a}_T^\nu \equiv \epsilon_T^{\mu\nu} a_{T\mu}$.

- The Dirac matrices satisfy the anticommutation relation $\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}I$ (with I the identity matrix) and the Hermiticity condition $(\gamma^\mu)^\dagger = \gamma^0 \gamma^\mu \gamma^0$. We define a fifth matrix $\gamma^5 \equiv i\gamma^0 \gamma^1 \gamma^2 \gamma^3$ and the antisymmetric structure $\sigma^{\mu\nu} \equiv \frac{i}{2}[\gamma^\mu, \gamma^\nu]$.

A

SYMMETRIC TRACELESS TENSORS

In this appendix we list the completely symmetric and traceless tensors $k_T^{i_1\cdots i_n}$ that are built from the partonic momentum k_T. Up to rank $n = 4$, these tensors are given by

$$k_T^{ij} \equiv k_T^i k_T^j + \frac{1}{2}k^2 g_T^{ij}, \tag{B.1}$$

$$k_T^{ijk} \equiv k_T^i k_T^j k_T^k + \frac{1}{4}k^2 \left(g_T^{ij} k_T^k + g_T^{ik} k_T^j + g_T^{jk} k_T^i \right), \tag{B.2}$$

$$k_T^{ijkl} \equiv k_T^i k_T^j k_T^k k_T^l + \frac{1}{6}k^2 \left(g_T^{ij} k_T^{kl} + g_T^{ik} k_T^{jl} + g_T^{il} k_T^{jk} + g_T^{jk} k_T^{il} + g_T^{jl} k_T^{ik} + g_T^{kl} k_T^{ij} \right)$$
$$- \frac{1}{8}k^4 \left(g_T^{ij} g_T^{kl} + g_T^{ik} g_T^{jl} + g_T^{il} g_T^{jk} \right), \tag{B.3}$$

satisfying

$$g_{Tij}\, k_T^{ij} = g_{Tij}\, k_T^{ijk} = g_{Tij}\, k_T^{ijkl} = 0. \tag{B.4}$$

Products of symmetric traceless tensors can be decomposed as follows:

$$k_T^i k_T^j = k_T^{ij} - \frac{1}{2}k^2 g_T^{ij}, \tag{B.5}$$

$$k_T^i k_T^{jk} = k_T^{ijk} - \frac{1}{4}k^2 \left(g_T^{ij} k_T^k + g_T^{ik} k_T^j - g_T^{jk} k_T^i \right), \tag{B.6}$$

$$k_T^{ij} k_T^{kl} = k_T^{ijkl} - \frac{1}{6}k^2 \left(g_T^{ik} k_T^{jl} + g_T^{il} k_T^{jk} + g_T^{jk} k_T^{il} + g_T^{jl} k_T^{ik} - 2g_T^{ij} k_T^{kl} - 2g_T^{kl} k_T^{ij} \right)$$
$$+ \frac{1}{8}k^4 \left(g_T^{ik} g_T^{jl} + g_T^{il} g_T^{jk} - g_T^{ij} g_T^{kl} \right). \tag{B.7}$$

The symmetric traceless tensor $k_T^{i_1\cdots i_n}$ of rank $n \geq 1$ only has two independent components, which allows for the following decomposition:
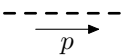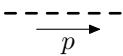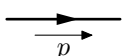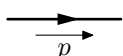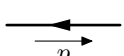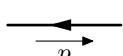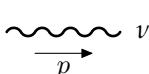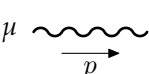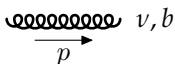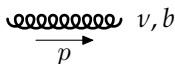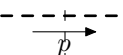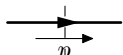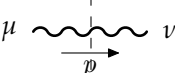
$$k_T^{i_1\cdots i_n} \;\to\; \frac{|\mathbf{k}|^n}{2^{n-1}}\, e^{\pm in\varphi}, \tag{B.8}$$

in terms of the two real numbers $|\mathbf{k}|$ and φ, the polar coordinates of the transverse vector k_T.

C

FEYNMAN RULES FOR SPECTATOR MODEL

In this appendix we provide the Feynman rules for the scalar spectator model that is employed in chapter 4, including those for eikonal lines and vertices. In table C.1 we list the relevant rules for the model, adopting the Feynman gauge; they are given for both amplitudes and conjugate amplitudes.

Amplitude		Conjugate amplitude	
(scalar propagator)	$\dfrac{i}{p^2 - m_s^2 + i\epsilon}$	(scalar propagator)	$-\dfrac{i}{p^2 - m_s^2 - i\epsilon}$
(fermion propagator)	$\dfrac{i(\not{p} + m_f)}{p^2 - m_f^2 + i\epsilon}$	(fermion propagator)	$-\dfrac{i(\not{p} + m_f)}{p^2 - m_f^2 - i\epsilon}$
(antifermion propagator)	$-\dfrac{i(\not{p} - m_f)}{p^2 - m_f^2 + i\epsilon}$	(antifermion propagator)	$\dfrac{i(\not{p} - m_f)}{p^2 - m_f^2 - i\epsilon}$
$\mu \;\;\;\; \nu$ (photon)	$-\dfrac{i g^{\mu\nu}}{p^2 + i\epsilon}$	$\mu \;\;\;\; \nu$ (photon)	$\dfrac{i g^{\mu\nu}}{p^2 - i\epsilon}$
$\mu,a \;\;\;\; \nu,b$ (gluon)	$-\dfrac{i\delta^{ab} g^{\mu\nu}}{p^2 + i\epsilon}$	$\mu,a \;\;\;\; \nu,b$ (gluon)	$\dfrac{i\delta^{ab} g^{\mu\nu}}{p^2 - i\epsilon}$
		(cut scalar)	$2\pi\,\theta(p^0)\,\delta(p^2 - m_s^2)$
		(cut fermion)	$2\pi\,\theta(p^0)\,\delta(p^2 - m_f^2) \times (\not{p} + m_f)$
		(cut antifermion)	$-2\pi\,\theta(p^0)\,\delta(p^2 - m_f^2) \times (\not{p} - m_f)$
		$\mu \;\;\;\; \nu$ (cut photon)	$-2\pi\,\theta(p^0)\,\delta(p^2)\,g^{\mu\nu}$

μ, a ⟶ p ⟶ ν, b		$-2\pi\,\theta(p^0)\,\delta(p^2)\,\delta^{ab}\,g^{\mu\nu}$

(amplitude)		(conjugate amplitude)	
μ	$-ie_q e\gamma^\mu$	μ	$ie_q e\gamma^\mu$
p_1, μ,a, p_2	$-igt^a(p_1+p_2)^\mu$	p_1, μ,a, p_2	$igt^a(p_1+p_2)^\mu$
p_1, μ,a, p_2	$igt^a(p_1+p_2)^\mu$	p_1, μ,a, p_2	$-igt^a(p_1+p_2)^\mu$
μ,a	$-igt^a\gamma^\mu$	μ,a	$igt^a\gamma^\mu$
μ,a, p_1; p_3; ρ,c; ν,b, p_2	$-gf^{abc}V^{\mu\nu\rho}$	μ,a, p_1; p_3; ρ,c; ν,b, p_2	$gf^{abc}V^{\mu\nu\rho}$
μ,a; ν,b	$ig^2\{t^a,t^b\}g^{\mu\nu}$	μ,a; ν,b	$-ig^2\{t^a,t^b\}g^{\mu\nu}$

Table C.1: The (relevant) Feynman rules for the scalar spectator model. The dashed line separating the columns represents the final-state cut. In the calculations of chapter 4 we set the quark mass to zero ($m_f = 0$).

The rules for conjugate amplitudes are easily obtained from those for amplitudes through complex conjugation. There is one exception to this, though: the three-gluon vertex changes sign when it crosses the final-state cut [345].[1] Note that table C.1 also

---

[1] This rule in fact applies to all momentum-dependent vertex factors that change sign when all momenta are reversed.

provides the rules for cutting lines. For cutting a line with mass m and momentum p going to the right, we replace the denominator of its propagator by an on-shell delta function (see e.g. [346]):

$$\frac{1}{p^2 - m^2 + i\epsilon} \ \rightarrow \ -2\pi i \, \theta(p^0) \, \delta(p^2 - m^2), \qquad (C.1)$$

where the theta function ensures a positive energy flow from left to right through the cut. The numerator remains unchanged.

The sign of the scalar-scalar-gluon coupling depends on the direction of color flow through the scalar lines with respect to the direction of momentum flow. To indicate the direction of color flow, we use blue arrows on the scalar lines. Furthermore, for convenience we define the following object for the three-gluon vertex:

$$V^{\mu\nu\rho} \equiv g^{\mu\nu}(p_1 - p_2)^\rho + g^{\nu\rho}(p_2 - p_3)^\mu + g^{\rho\mu}(p_3 - p_1)^\nu. \qquad (C.2)$$

All listed vertex factors would come with an opposite sign if in the covariant derivatives the interaction terms with the vector boson fields would be defined with a minus rather than with a plus sign [347].

In table C.2 we list the relevant eikonal Feynman rules. We make use of an auxiliary vector v that depends on whether the fermion is left-moving or right-moving. In the former case we can take $v = n$ and in the latter case we can choose $v = \bar{n}$. Throughout the thesis we use the following notation for eikonal lines [102, 114]: the circles at the ends of an eikonal line indicate the direction of momentum flow (from the full to the empty circle) of the original fermion, and the arrow on the line denotes the direction of color flow (and thus also the direction of fermion number flow).
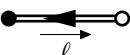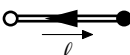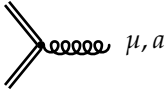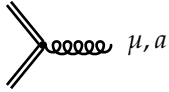
Amplitude		Conjugate amplitude	
	$\dfrac{i}{\ell \cdot v + i\epsilon}$		$-\dfrac{i}{\ell \cdot v - i\epsilon}$
	$-\dfrac{i}{\ell \cdot v + i\epsilon}$		$\dfrac{i}{\ell \cdot v - i\epsilon}$
μ, a	$-igt^a v^\mu$	μ, a	$igt^a v^\mu$

Table C.2: The (relevant) Feynman rules for eikonal lines and vertices. The dashed line separating the columns represents the final-state cut.

D

DEFINITIONS OF GLUON TMDS

In this appendix we provide the definitions of the leading-twist gluon TMDs in eq. (5.16) in terms of the coefficient functions A_i that appear in eqs. (5.11)–(5.13). The TMD correlator in eq. (5.15) is related to the unintegrated correlator as

$$\Gamma^{ij}(x, \mathbf{k}) \equiv \frac{2}{P^+} \int dk^- \, \Gamma^{+i;+j}(k) = \frac{M^2}{(P^+)^2} \int [d\sigma d\tau] \, \Gamma^{+i;+j}(k), \qquad (D.1)$$

where we have introduced the shorthand notation

$$[d\sigma d\tau] \equiv d\sigma d\tau \, \delta \left(\tau - x\sigma + x^2 + \frac{\mathbf{k}^2}{M^2} \right), \qquad (D.2)$$

with the dimensionless invariants σ and τ given by

$$\sigma \equiv \frac{2k \cdot P}{M^2}, \qquad \tau \equiv \frac{k^2}{M^2}. \qquad (D.3)$$

Making use of eq. (D.1), the gluon TMDs are related to the coefficient functions as follows:

$$x f_1(x, \mathbf{k}^2) \equiv M^2 \int [d\sigma d\tau] \left(A_2 + 2x A_4 + x^2 A_3 + \frac{\mathbf{k}^2}{2M^2} A_6 \right), \qquad (D.4)$$

$$x h_1^\perp(x, \mathbf{k}^2) \equiv M^2 \int [d\sigma d\tau] \, A_6, \qquad (D.5)$$

$$x g_1(x, \mathbf{k}^2) \equiv 2M^2 \int [d\sigma d\tau] \left\{ A_8 + A_9 + x \left(A_{10} + A_{11} \right) \right.$$

$$+ \left(\frac{\sigma}{2} - x \right) \left[A_{12} + x \left(A_{14} + A_{15} \right) + x^2 A_{13} \right]$$

$$\left. + \frac{\mathbf{k}^2}{2M^2} \left[A_{19} + A_{23} + \left(\frac{\sigma}{2} - x \right) A_{25} \right] \right\}, \qquad (D.6)$$

$$x h_{1L}^\perp(x, \mathbf{k}^2) \equiv -2M^2 \int [d\sigma d\tau] \left[A_{18} + A_{22} + \left(\frac{\sigma}{2} - x \right) A_{24} \right], \qquad (D.7)$$

$$x f_{1T}^\perp(x, \mathbf{k}^2) \equiv M^2 \int [d\sigma d\tau] \left[A_{16} - A_{20} + x \left(A_{18} - A_{22} \right) \right], \qquad (D.8)$$

$$x g_{1T}(x, \mathbf{k}^2) \equiv -M^2 \int [d\sigma d\tau] \left[2A_{12} + A_{17} + A_{21} + 2x \left(A_{14} + A_{15} \right) \right.$$

$$+ x \left(A_{19} + A_{23}\right) + 2x^2 A_{13} + \frac{k^2}{M^2} A_{25} \Bigg], \qquad (D.9)$$

$$x h_1(x, \boldsymbol{k}^2) \equiv 2M^2 \int [d\sigma d\tau] \left[A_{16} + A_{20} + x \left(A_{18} + A_{22}\right) + \frac{k^2}{2M^2} A_{24} \right], \qquad (D.10)$$

$$x h_{1T}^{\perp}(x, \boldsymbol{k}^2) \equiv 2M^2 \int [d\sigma d\tau] \, A_{24}, \qquad (D.11)$$

$$x f_{1LL}(x, \boldsymbol{k}^2) \equiv \frac{M^2}{3} \int [d\sigma d\tau] \Bigg\{ A_{27} - 2A_{34} + 2x A_{28} + x^2 A_{26}$$

$$+ 2(\sigma - 2x) \left(A_{37} + x A_{35}\right) + \frac{(\sigma - 2x)^2}{2} \left(A_{40} + 2x A_{42} + x^2 A_{41} \right)$$

$$- \frac{k^2}{M^2} \Bigg[A_{26} - A_{32} + A_{40} + 2x A_{42} + x^2 A_{41} + (\sigma - 3x) A_{30}$$

$$- \left(\frac{(\sigma - 2x)^2}{4} - \frac{k^2}{2M^2} \right) A_{44} \Bigg] \Bigg\}, \qquad (D.12)$$

$$x h_{1LL}^{\perp}(x, \boldsymbol{k}^2) \equiv -\frac{2M^2}{3} \int [d\sigma d\tau] \Bigg[A_{26} - A_{32} + (\sigma - 3x) A_{30}$$

$$- \left(\frac{(\sigma - 2x)^2}{4} - \frac{k^2}{2M^2} \right) A_{44} \Bigg], \qquad (D.13)$$

$$x f_{1LT}(x, \boldsymbol{k}^2) \equiv -M^2 \int [d\sigma d\tau] \Bigg\{ A_{37} + x A_{35} + \left(\frac{\sigma}{2} - x \right) \left(A_{40} + 2x A_{42} + x^2 A_{41} \right)$$

$$- \frac{k^2}{4M^2} \left[A_{30} + \left(x - \frac{\sigma}{2} \right) A_{44} \right] \Bigg\}, \qquad (D.14)$$

$$x g_{1LT}(x, \boldsymbol{k}^2) \equiv -\frac{M^2}{2} \int [d\sigma d\tau] \left[A_{29} + \left(x - \frac{\sigma}{2} \right) \left(A_{33} + x A_{31} \right) \right], \qquad (D.15)$$

$$x h_{1LT}(x, \boldsymbol{k}^2) \equiv \frac{M^2}{2} \int [d\sigma d\tau] \Bigg\{ A_{28} + x A_{26} + \left(\frac{\sigma}{2} - x \right) \left(A_{32} + x A_{30} \right)$$

$$+ \frac{k^2}{2M^2} \left[A_{30} + \left(x - \frac{\sigma}{2} \right) A_{44} \right] \Bigg\}, \qquad (D.16)$$

$$x h_{1LT}^{\perp}(x, \boldsymbol{k}^2) \equiv -M^2 \int [d\sigma d\tau] \left[A_{30} + \left(x - \frac{\sigma}{2} \right) A_{44} \right], \qquad (D.17)$$

$$x f_{1TT}(x, \boldsymbol{k}^2) \equiv \frac{M^2}{2} \int [d\sigma d\tau] \left(A_{40} + 2x A_{42} + x^2 A_{41} + \frac{k^2}{6M^2} A_{44} \right), \qquad (D.18)$$

$$x g_{1TT}(x, \boldsymbol{k}^2) \equiv \frac{M^2}{2} \int [d\sigma d\tau] \left(A_{33} + x A_{31} \right), \qquad (D.19)$$

$$xh_{1TT}(x, \boldsymbol{k}^2) \equiv -\frac{M^2}{2} \int [d\sigma d\tau] \left[A_{27} + 2xA_{28} + x^2 A_{26} \right.$$

$$\left. + \frac{\boldsymbol{k}^2}{M^2} \left(A_{32} + xA_{30} \right) - \frac{\boldsymbol{k}^4}{4M^4} A_{44} \right], \tag{D.20}$$

$$xh_{1TT}^{\perp}(x, \boldsymbol{k}^2) \equiv \frac{M^2}{2} \int [d\sigma d\tau] \left(A_{32} + xA_{30} - \frac{\boldsymbol{k}^2}{3M^2} A_{44} \right), \tag{D.21}$$

$$xh_{1TT}^{\perp\perp}(x, \boldsymbol{k}^2) \equiv \frac{M^2}{2} \int [d\sigma d\tau] \, A_{44} . \tag{D.22}$$

E

CORRELATORS IN COORDINATE SPACE

In subsections 2.3.2 and 5.1.2 the quark-quark and gluon-gluon TMD correlators are parametrized in terms of leading-twist quark and gluon TMDs, respectively. Those parametrizations are given in momentum space (k_T-space). In this appendix we provide parametrizations of those correlators in coordinate space (z_T-space), based on our work in [1,67].

To facilitate the translation into z_T-space, let us denote a generic TMD correlator by Ω and a generic TMD function by f. Their counterparts in z_T-space are related by a Fourier transformation [348]:

$$\tilde{\Omega}(x,z) \equiv \int d^2\mathbf{k} \, e^{i\mathbf{k}\cdot\mathbf{z}} \, \Omega(x,\mathbf{k}), \tag{E.1}$$

$$\tilde{f}(x,z^2) \equiv \int d^2\mathbf{k} \, e^{i\mathbf{k}\cdot\mathbf{z}} \, f(x,\mathbf{k}^2). \tag{E.2}$$

In momentum space, the parametrization of the correlator in terms of TMDs f_j with rank $n = n(j) \geq 0$ takes the form[1]

$$\Omega(x,\mathbf{k}) = \sum_j C_j \frac{k_T^{i_1\ldots i_n}}{M^n} f_j(x,\mathbf{k}^2), \tag{E.3}$$

where C_j is a coefficient independent of k that contains information on both the hadron and parton polarization. Using eqs. (B.8), (E.1), and (E.3), the correlator in z_T-space is given by

$$
\begin{aligned}
\tilde{\Omega}(x,z) &= \sum_j C_j \int d^2\mathbf{k} \, e^{i\mathbf{k}\cdot\mathbf{z}} \frac{k_T^{i_1\ldots i_n}}{M^n} f_j(x,\mathbf{k}^2) \\
&= \sum_j C_j \frac{z_T^{i_1\ldots i_n}}{M^n} \int_0^\infty d|\mathbf{k}| \, |\mathbf{k}| \left(\frac{|\mathbf{k}|}{|\mathbf{z}|}\right)^n (2\pi i^n) J_n(|\mathbf{k}||\mathbf{z}|) f_j(x,\mathbf{k}^2) \\
&= \sum_j \frac{i^n}{n!} C_j M^n z_T^{i_1\ldots i_n} \tilde{f}_j^{(n)}(x,z^2),
\end{aligned} \tag{E.4}
$$

where on the second line the Bessel function of the first kind $J_k(u)$ arose from the integral identity

$$\int_0^{2\pi} d\alpha \, e^{ik\alpha} e^{iu\cos(\alpha-\beta)} = 2\pi i^k J_k(u) \, e^{ik\beta}. \tag{E.5}$$

---

1 For simplicity, possible Lorentz indices on Ω or C_j are omitted.

Following the conventions in [348], we define the function $\tilde{f}^{(n)}$ as

$$\tilde{f}^{(n)}(x, z^2) \equiv n! \left(-\frac{2}{M^2} \frac{\partial}{\partial z^2} \right)^n \tilde{f}(x, z^2)$$

$$= \frac{2\pi n!}{M^{2n}} \int_0^\infty d|\boldsymbol{k}|\, |\boldsymbol{k}| \left(\frac{|\boldsymbol{k}|}{|\boldsymbol{z}|} \right)^n J_n(|\boldsymbol{k}||\boldsymbol{z}|)\, f(x, \boldsymbol{k}^2), \tag{E.6}$$

where we used eq. (E.2), as well as the recurrence relation

$$\left(\frac{1}{u} \frac{d}{du} \right)^m \left[\frac{J_k(u)}{u^k} \right] = (-1)^m \frac{J_{k+m}(u)}{z^{k+m}}, \tag{E.7}$$

with $u = |\boldsymbol{k}||\boldsymbol{z}|$ (with $|\boldsymbol{k}|$ fixed), $m = n$, and $k = 0$.

From eqs. (E.4) and (E.6) it follows that in the parametrization of the correlator in z_T-space the nth derivative $\tilde{f}^{(n)}$ with respect to $\boldsymbol{z}^2$ appears rather than the function $\tilde{f} = \tilde{f}^{(0)}$ itself. Furthermore, we infer from eq. (E.6) that for definite-rank TMDs there is a one-to-one correspondence between the functions in momentum and coordinate space. The motivation for this particular definition of $\tilde{f}^{(n)}$ in eq. (E.6) becomes obvious once we set $\boldsymbol{z} = \boldsymbol{0}$, which is equivalent to integration over transverse momentum in k_T-space. Using the limit

$$\lim_{u \to 0} \frac{J_k(u)}{u^k} = \frac{1}{2^k k!}, \tag{E.8}$$

we see that

$$\lim_{|\boldsymbol{z}| \to 0} \tilde{f}^{(n)}(x, \boldsymbol{z}^2) = \int d^2\boldsymbol{k} \left(\frac{\boldsymbol{k}^2}{2M^2} \right)^n f(x, \boldsymbol{k}^2), \tag{E.9}$$

which is precisely the conventional nth moment $f^{(n)}(x)$ of the TMD. Hence, by construction, the derivatives in z_T-space are directly related to moments in k_T-space.

E.1 THE QUARK-QUARK TMD CORRELATOR

We can use eq. (E.4) to translate the quark-quark TMD correlator in eq. (2.33) to z_T-space. In [348] this was already done for spin-$\frac{1}{2}$ hadrons. For the spin-1 case we have

$$\tilde{\Phi}(x, z) = \tilde{\Phi}_U(x, z) + \tilde{\Phi}_L(x, z) + \tilde{\Phi}_T(x, z)$$
$$+ \tilde{\Phi}_{LL}(x, z) + \tilde{\Phi}_{LT}(x, z) + \tilde{\Phi}_{TT}(x, z), \tag{E.10}$$

where

$$\tilde{\Phi}_U(x, z) = \frac{1}{2} \left[\slashed{n}\, \tilde{f}_1(x, z^2) + iM\, \sigma_{\mu\nu} z_T^\mu \bar{n}^\nu\, \tilde{h}_1^{\perp (1)}(x, z^2) \right], \tag{E.11}$$

$$\tilde{\Phi}_L(x, z) = \frac{1}{2} \left[\gamma^5 \slashed{n}\, S_L\, \tilde{g}_1(x, z^2) - M\, \sigma_{\mu\nu} \gamma^5 \bar{n}^\mu z_T^\nu\, S_L\, \tilde{h}_{1L}^{\perp (1)}(x, z^2) \right], \tag{E.12}$$

138

$$\tilde{\Phi}_T(x,z) = \frac{1}{2}\left[iM\,\slashed{\hbar}\,\epsilon_T^{S_T z_T}\,\tilde{f}_{1T}^{\perp(1)}(x,z^2) + iM\gamma^5\slashed{\hbar}\,z\cdot S_T\,\tilde{g}_{1T}^{(1)}(x,z^2)\right.$$

$$\left. + i\sigma_{\mu\nu}\gamma^5\bar{n}^\mu S_T^\nu\,\tilde{h}_1(x,z^2) + \frac{iM^2\sigma_{\mu\nu}\gamma^5\bar{n}^\mu z_T^{\nu\rho}S_{T\rho}}{2}\,\tilde{h}_{1T}^{\perp(2)}(x,z^2)\right], \qquad \text{(E.13)}$$

$$\tilde{\Phi}_{LL}(x,z) = \frac{1}{2}\left[\slashed{\hbar}\,S_{LL}\,\tilde{f}_{1LL}(x,z^2) + iM\,\sigma_{\mu\nu}z_T^\mu\bar{n}^\nu S_{LL}\,\tilde{h}_{1LL}^{\perp(1)}(x,z^2)\right], \qquad \text{(E.14)}$$

$$\tilde{\Phi}_{LT}(x,z) = \frac{1}{2}\left[iM\,\slashed{\hbar}\,z\cdot S_{LT}\,\tilde{f}_{1LT}^{(1)}(x,z^2) + iM\gamma^5\slashed{\hbar}\,\epsilon_T^{S_{LT}z_T}\,\tilde{g}_{1LT}^{(1)}(x,z^2)\right.$$

$$\left. + \sigma_{\mu\nu}\bar{n}^\nu S_{LT}^\mu\,\tilde{h}_{1LT}(x,z^2) + \frac{M^2\sigma_{\mu\nu}\bar{n}^\nu z_T^{\mu\rho}S_{LT\rho}}{2}\,\tilde{h}_{1LT}^{\perp(2)}(x,z^2)\right], \qquad \text{(E.15)}$$

$$\tilde{\Phi}_{TT}(x,z) = \frac{1}{2}\left[-\frac{M^2\slashed{\hbar}\,z_T^{\mu\nu}S_{TT\mu\nu}}{2}\,\tilde{f}_{1TT}^{(2)}(x,z^2) + \frac{M^2\gamma^5\slashed{\hbar}\,\epsilon_{T\mu\nu}z_T^{\mu\rho}S_{TT\rho}^\nu}{2}\,\tilde{g}_{1TT}^{(2)}(x,z^2)\right.$$

$$- iM\,\sigma_{\mu\nu}\bar{n}^\nu z_T^\rho S_{TT\rho}^\mu\,\tilde{h}_{1TT}^{(1)}(x,z^2)$$

$$\left. - \frac{iM^3\sigma_{\mu\nu}\bar{n}^\nu z_T^{\mu\rho\sigma}S_{TT\rho\sigma}}{6}\,\tilde{h}_{1TT}^{\perp(3)}(x,z^2)\right]. \qquad \text{(E.16)}$$

The quark TMDs in z_T-space are one-to-one related to their k_T-space counterparts through eq. (E.6).

E.2 THE GLUON-GLUON TMD CORRELATOR

We can use eq. (E.4) to translate the gluon-gluon TMD correlator in eq. (5.16) to z_T-space. It is given by

$$\tilde{\Gamma}^{ij}(x,z) = \tilde{\Gamma}_U^{ij}(x,z) + \tilde{\Gamma}_L^{ij}(x,z) + \tilde{\Gamma}_T^{ij}(x,z)$$
$$+ \tilde{\Gamma}_{LL}^{ij}(x,z) + \tilde{\Gamma}_{LT}^{ij}(x,z) + \tilde{\Gamma}_{TT}^{ij}(x,z), \qquad \text{(E.17)}$$

where

$$\tilde{\Gamma}_U^{ij}(x,z) = \frac{xP^+}{2}\left[\delta_T^{ij}\,\tilde{f}_1(x,z^2) - \frac{M^2 z_T^{ij}}{2}\,\tilde{h}_1^{\perp(2)}(x,z^2)\right], \qquad \text{(E.18)}$$

$$\tilde{\Gamma}_L^{ij}(x,z) = \frac{xP^+}{2}\left[i\epsilon_T^{ij}S_L\,\tilde{g}_1(x,z^2) - \frac{M^2\epsilon_{T\,\alpha}^{\{i}z_T^{j\}\alpha}S_L}{4}\,\tilde{h}_{1L}^{\perp(2)}(x,z^2)\right], \qquad \text{(E.19)}$$

$$\tilde{\Gamma}_T^{ij}(x,z) = \frac{xP^+}{2}\left[iM\,\delta_T^{ij}\,\epsilon_T^{S_T z_T}\,\tilde{f}_{1T}^{\perp(1)}(x,z^2) - M\,\epsilon_T^{ij}\,z\cdot S_T\,\tilde{g}_{1T}^{(1)}(x,z^2)\right.$$

$$- \frac{iM\left(\epsilon_T^{z_T\{i}S_T^{j\}} + \epsilon_T^{S_T\{i}z_T^{j\}}\right)}{4} \tilde{h}_1^{(1)}(x, z^2)$$

$$+ \frac{iM^3\epsilon_{T\ \alpha}^{\{i}z_T^{j\}\alpha S_T}}{12} \tilde{h}_{1T}^{\perp(3)}(x, z^2)\Bigg], \tag{E.20}$$

$$\tilde{\Gamma}_{LL}^{ij}(x, z) = \frac{xP^+}{2} \left[\delta_T^{ij}S_{LL}\, \tilde{f}_{1LL}(x, z^2) - \frac{M^2 z_T^{ij}S_{LL}}{2} \tilde{h}_{1LL}^{\perp(2)}(x, z^2)\right], \tag{E.21}$$

$$\tilde{\Gamma}_{LT}^{ij}(x, z) = \frac{xP^+}{2} \left[iM\,\delta_T^{ij}\, z\cdot S_{LT}\, \tilde{f}_{1LT}^{(1)}(x, z^2) - M\,\epsilon_T^{ij}\epsilon_T^{S_{LT}z_T}\, \tilde{g}_{1LT}^{(1)}(x, z^2)\right.$$

$$\left. + iM\, S_{LT}^{\{i}z_T^{j\}}\, \tilde{h}_{1LT}^{(1)}(x, z^2) - \frac{iM^3 z_T^{ij\alpha}S_{LT\alpha}}{6} \tilde{h}_{1LT}^{\perp(3)}(x, z^2)\right], \tag{E.22}$$

$$\tilde{\Gamma}_{TT}^{ij}(x, z) = \frac{xP^+}{2} \left[-\frac{M^2\delta_T^{ij}z_T^{\alpha\beta}S_{TT\alpha\beta}}{2} \tilde{f}_{1TT}^{(2)}(x, z^2) - \frac{iM^2\epsilon_T^{ij}\epsilon_{T\gamma}^{\beta}z_T^{\gamma\alpha}S_{TT\alpha\beta}}{2} \tilde{g}_{1TT}^{(2)}(x, z^2)\right.$$

$$\left. + S_{TT}^{ij}\, \tilde{h}_{1TT}(x, z^2) - \frac{M^2 S_{TT\alpha}^{\{i}z_T^{j\}\alpha}}{2} \tilde{h}_{1TT}^{\perp(2)}(x, z^2)\right.$$

$$\left. + \frac{M^4 z_T^{ij\alpha\beta}S_{TT\alpha\beta}}{24} \tilde{h}_{1TT}^{\perp\perp(4)}(x, z^2)\right]. \tag{E.23}$$

The gluon TMDs in z_T-space are one-to-one related to their k_T-space counterparts through eq. (E.6).

E

In high school one is taught that a proton consists of two up quarks and one down quark. However, the veracity of that claim entirely depends on the resolution that is used to look at the proton. The internal structure of the proton can only be unraveled by studying interactions of the proton with other particles. The higher the energy of the particle that bumps into the proton (a higher energy implies a smaller wavelength), the smaller the length scale that can be studied. In collisions at low energies we see the proton as a whole; it behaves the same as a point particle. As the collision energy increases, the size of the proton plays an increasingly important role. The substructure then appears to be described by three point particles, the quarks. This was first observed in 1968. However, when the proton is involved in a very high-energetic collision, it appears to consist of a 'soup' of innumerable quarks, antiquarks, and gluons, see figure 1. At high resolution, the proton is an immensely chaotic collection of elementary particles that move around with almost the speed of light and in which quark-antiquark pairs are constantly formed and disappear according to Einstein's well-known formula $E = mc^2$. Contrary to what the picture suggests, there is no physical edge that limits the proton – the density of the particle soup gradually decreases as the distance to the center increases.
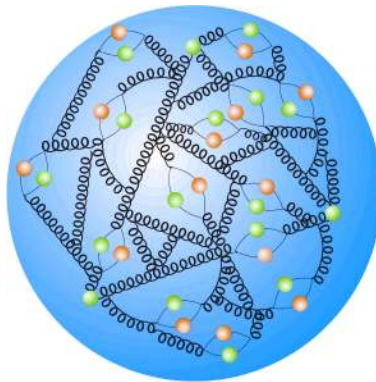
Figure 1: A schematic representation of the internal structure of the proton seen at high energy. In addition to three quarks, one can find many antiquarks and gluons inside the proton. (Source: the DESY website.)

The particles in the proton are called *partons* and their distribution is described by *parton distribution functions* (PDFs). The so-called *collinear* PDFs $f_i(x, Q)$ say something about the probability to find a parton of type i (an up quark, gluon, etc.) with a fraction x of the total proton momentum (in the direction in which

the proton moves) at a resolution scale Q inside the proton. Since PDFs depend on a resolution scale, the inner structure of the proton is not a static but rather a dynamic property. You could compare this with how you look when you are lit (and therefore interact) with photons of different wavelengths. When visible light is used, the image is clearly different from an X-ray image in which electromagnetic radiation with a much smaller wavelength is used. That what is 'constant' in the proton, which defines the proton, is that the number of up quarks minus the number of anti-up quarks equals 2 and that the number of down quarks minus the number of antidown quarks equals 1.

The internal structure of the proton is a consequence of the *strong interaction*. The strong interaction is responsible for the strong nuclear force that holds atomic nuclei together and is one of the four fundamental interactions in nature (the other three are the electromagnetic and weak interactions, and gravity). At the level of a proton, the strength of the strong interaction is respectively 10^2 and 10^6 times greater than that of the electromagnetic and weak interactions, hence the indication 'strong'. The strong interaction is described by a theory called *quantum chromodynamics* (QCD). This name refers to the Greek word 'chroma' that means color. QCD, namely, describes the dynamics of elementary particles that have a *color charge*. Color charge is somewhat analogous to electric charge. Electrically charged particles, such as electrons, interact with each other through the exchange of photons. In addition to an electric charge, quarks also have a color charge, allowing them to interact through the strong interaction. The transmitter of the strong interaction is the gluon, the QCD version of the photon.

An important difference between photons and gluons is that the photon is electrically neutral while the gluon has a color charge. The latter implies that gluons can also engage in the strong interaction with each other. A consequence of this property is that in contrast to the other fundamental interactions the strength of the strong interaction, which is characterized by the coupling parameter α_s, increases as the distance between two color charges increases. Since at some point it becomes energetically more favorable to create a quark-antiquark pair than to further separate the color charges, partons only occur in bound states called *hadrons* (protons and neutrons are examples of hadrons). This phenomenon is called *confinement*. The counterpart of confinement is *asymptotic freedom*: at very small length scales, smaller than the proton radius of about 1 fm $= 10^{-15}$ m, the strong interaction is very weak; so weak that the partons can move almost freely (in that case α_s is very close to zero).

The partons form the (fundamental) degrees of freedom of QCD at high energies, whereas at low energies the (effective) degrees of freedom are formed by the hadrons. We still do not understand how certain features of hadrons can be explained in terms of the partons. For example, it is unclear where exactly the proton mass of 938 MeV $= 1.67 \times 10^{-27}$ kg comes from. The Higgs mechanism only explains the masses of the elementary quarks that are at most a few MeVs. The remainder of the proton mass, which amounts to about 98%, comes from the binding energy in the proton. This energy corresponds to the kinetic and potential energy stored in

the 'sea' of quarks, antiquarks, and gluons and its size is mainly determined by the strong interaction. When you are standing on a scale, the number before the decimal point roughly corresponds to the total binding energy of all protons and neutrons in your body ($m = E/c^2$); the number after the decimal point comes from your interaction with the Higgs field (and the fact that you do not sink through the scale is the result of the electromagnetic interaction).

Why can we not just calculate the properties of hadrons? After all, we have a theory that describes the strong interaction, namely QCD. It turns out that what can and cannot be calculated with QCD depends on the energy scale involved in the process under consideration. Thanks to asymptotic freedom, high-energy interactions between partons (these interactions are weak) can be calculated using *perturbation theory*. This means that the probability that a certain process occurs, which is characterized by the so-called *scattering cross section*, can be expanded in powers of α_s. Since for sufficiently high energies the (dimensionless) parameter α_s is much smaller than 1, it is usually sufficient to calculate only a few terms in this expansion. Contributions to these terms are graphically represented by the well-known Feynman diagrams. However, at low energies α_s can become greater than 1. In that case, higher-order terms yield larger contributions than the lowest-order terms – the expansion does not converge and has thus become meaningless. Since the internal structure of hadrons is shaped by low-energy interactions between partons (the partons are strongly bound) it cannot be calculated with perturbative techniques. This means that we for example cannot calculate the mass of the proton using Feynman diagrams.

As noted earlier, the internal structure of hadrons is described by PDFs. These functions contain information about the dynamics of quarks and gluons inside hadrons and can depend on various kinematic variables. Since PDFs cannot be calculated using perturbative QCD, they must be extracted from collision experiments. In this thesis we study measurable quantities, also called *observables*, that depend on different types of PDFs. By measuring these observables we can learn more about what exactly is happening inside hadrons. Moreover, a greater understanding of the internal structure of the proton for example allows for more accurate descriptions of high-energy collisions between protons in the Large Hadron Collider. The latter can be of great significance for the search for new elementary particles.

An example of an observable that 'feels' the internal structure of hadrons is the cross section of the so-called *Drell-Yan* (DY) process. DY scattering is a common scenario in high-energy collisions between hadrons; during such a collision, a quark from one hadron and an antiquark from the other interact with each other, producing an electrically charged lepton-antilepton pair (an example of which is an electron-positron pair), see figure 2. Since the lepton momenta can be measured in particle detectors, the collision can be reconstructed. For the reconstruction it is important to know the probabilities for finding quarks and antiquarks with certain momenta inside hadrons – this is exactly the information contained in the PDFs. As the interaction between the quark and antiquark takes place at high energy (this is also

called the 'hard' process) it can be calculated perturbatively. Ultimately, the total cross section of DY scattering is given by the product of the hard scattering cross section with two PDFs (one for the quark and one for the antiquark). The possibility to write a cross section as a product means that the process can be divided into separate contributions that do not 'talk' to each other. This is called *factorization*. Factorization is essential for the experimental determination of PDFs and forms the foundation of hadron physics.
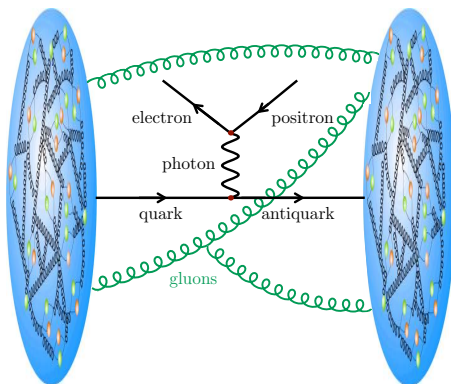

Figure 2: An example of a DY process: a quark and an antiquark from two colliding protons produce a photon that subsequently decays into an electron-positron pair. An arbitrary number of gluons is exchanged between the strongly Lorentz-contracted protons.

However, the property of factorization should not be taken for granted. In order to calculate the total cross section for the DY process, the superposition principle of quantum mechanics requires summing over all possible interactions that could possibly take place, including those between the hadrons. Since for example arbitrarily many low-energy ('soft') gluons can be exchanged between the hadrons, it is a priori unclear that the contributions from the two hadrons can be completely disentangled. In other words, it is unclear whether factorization is permitted. That factorization does hold was proven in the 1980s by Collins, Soper, and Sterman (CSS) for collinear PDFs. The fact that interference effects between the two hadrons are indeed negligible is intuitively easy to understand. After all, as the hadrons move towards each other at almost the speed of light (hence they are relativistic objects), there is length contraction along the direction of movement as well as time dilation of the interactions inside the hadrons. As a result, the hadrons can be considered as two 'frozen pancakes' that hardly 'feel' each other prior to the collision. Factorization implies that collinear PDFs are *universal*; they pertain to individual hadrons and are completely independent of any hard scattering processes. Hence, collinear PDFs can be measured in one process and then used for the description of hadrons in other processes.

As it turns out, not all types of PDFs are universal. In this thesis we look at *multidimensional* PDFs that in addition to the longitudinal component x also depend on the transverse parton momentum $\boldsymbol{k}$. These PDFs contain more information about the internal structure of hadrons than the one-dimensional collinear PDFs and may differ from process to process. The process dependence of multidimensional PDFs can be calculated and isolated in a so-called *gauge link* that depends on the direction in which color charges move in the hard process. An example of such a nonuniversal multidimensional PDF is the so-called *Boer-Mulders* (BM) function $h_{1,i}^{\perp q}(x, \boldsymbol{k}, Q)$. This (three-dimensional) function quantifies the net transverse spin polarization of a quark or antiquark of type i in an unpolarized hadron. In the DY process, the BM effect generates a very specific asymmetry in the directions of the lepton momenta in the transverse plane. Although this asymmetry in the DY cross section has been measured by various experiments, more data (among other things) is needed for an accurate extraction of the BM function.

In part I of this thesis we look at color effects that may influence the factorization property of the BM asymmetry in the DY process. These effects originate from 'color-entangled' gluons that are exchanged between the hadrons and can, at worst, give rise to a violation of factorization. Since the BM effect is actively being investigated by various experimental collaborations, it is essential to know how exactly it manifests itself in the cross section. By largely following the steps of the CSS proof, we find that the contribution of the BM effect to the DY cross section factorizes in the 'standard way'. This means that the BM asymmetry can be written as a product of a hard scattering cross section with quark and antiquark BM functions, without any complications.

Part II deals with multidimensional distributions of gluons inside hadrons. At very high collision energies, the internal structure of hadrons is dominated by gluons; this corresponds to the regime of small values of the momentum component x. We introduce several new multidimensional gluon distributions for spin-1 hadrons that contain interesting information about correlations between transverse gluon momenta and spin components. Because gluon PDFs are still largely unknown, we also derive limits for maximal values that certain functions may take. For very small values of x, different gluon PDFs appear to be related to each other. This results in a simpler description of the distributions of gluons inside hadrons at high energies. Finally, we illustrate with a model calculation that certain asymmetries observed in the angular distributions of two produced hadrons in collisions between protons and lead nuclei may originate from quantum correlations that exist between gluons inside the lead nuclei.

Nowadays we know quite well what the proton looks like from the inside wearing 'one-dimensional glasses'. Through a combination of experimental and theoretical efforts since the 1970s, the one-dimensional collinear PDFs have been determined fairly precisely (for not too small values of x that is). Since the last turn of the century, the research field has expanded to include measurements on the multidimensional

inner structure of hadrons; a few multidimensional PDFs have already been extracted. In the coming years we expect to learn more about interesting quantum effects in hadrons, an example of which is the BM effect that we investigate in part I of this thesis. Moreover, as measurements involve various different processes, we could also test some QCD predictions with regards to the process-dependent gauge links. For the research at small values of x, the future Electron-Ion Collider (EIC) could be of great importance. This particle accelerator could collide electrons with polarized protons and light atomic nuclei at very high energy and thus shed new light on the gluonic structure of hadrons. With the EIC, several theoretical predictions from part II of this thesis could be experimentally tested. In short, the research into the properties of the particle soup inside hadrons still has a long road ahead. However, we find ourselves in interesting times, gradually unraveling the rich multidimensional inner structure of hadrons through which we learn more about the most important building blocks of our (visible) universe.

Op de middelbare school leer je dat een proton bestaat uit twee upquarks en één downquark. Het waarheidsgehalte van die bewering hangt echter helemaal af van de resolutie waarmee je naar het proton kijkt. De interne structuur van het proton kan alleen worden ontrafeld door interacties te bestuderen van het proton met andere deeltjes. Hoe hoger de energie van het deeltje dat op het proton invalt (een hogere energie betekent een kleinere golflengte), des te kleiner is de lengteschaal die kan worden bestudeerd. In een botsing bij lage energie zien we het proton in z'n geheel; het gedraagt zich hetzelfde als een puntdeeltje. Naarmate de botsingsenergie wordt opgevoerd, gaat de afmeting van het proton een steeds grotere rol spelen. De substructuur blijkt dan te kunnen worden beschreven door drie puntdeeltjes, de quarks. Dit werd voor het eerst waargenomen in 1968. Echter, wanneer het proton betrokken is bij een zeer hoog-energetische botsing, dan blijkt het te bestaan uit een 'soep' van ontelbaar veel quarks, antiquarks en gluonen, zie figuur 1. Bij hoge resolutie is het proton een immens chaotische collectie van elementaire deeltjes die rondrazen met bijna de lichtsnelheid en waarin voortdurend quark-antiquark paren worden gevormd en weer verdwijnen volgens Einsteins bekende formule $E = mc^2$. In tegenstelling tot wat het plaatje suggereert is er geen fysieke rand die het proton begrenst – de dichtheid van de deeltjessoep neemt geleidelijk af naarmate de afstand tot het centrum groter wordt.
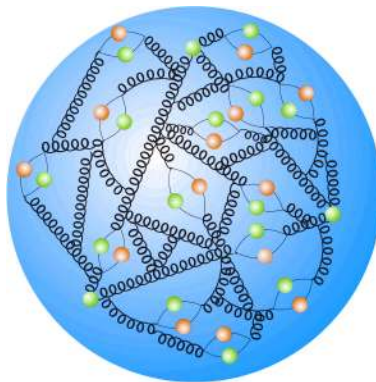
Figuur 1: Een schematische weergave van de interne structuur van het proton gezien bij hoge energie. Naast drie quarks kun je in het proton vele antiquarks en gluonen aantreffen. (Bron: de DESY website.)

De deeltjes in het proton worden *partonen* genoemd en hun verdeling wordt beschreven door *partondistributiefuncties* (PDFs). De zogenaamde *collineaire* PDFs

$f_i(x, Q)$ zeggen iets over de waarschijnlijkheid om bij een resolutie Q een parton van het type i (een upquark, gluon etc.) met een fractie x van de totale protonimpuls (in de richting waarin het proton beweegt) in het proton aan te treffen. Aangezien PDFs afhangen van een resolutieschaal, is de inwendige structuur van het proton niet een statische maar een dynamische eigenschap. Je zou dit kunnen vergelijken met hoe jijzelf eruit ziet als je wordt beschenen (en dus interacties aangaat) met fotonen van verschillende golflengtes. Bij gebruik van zichtbaar licht is het plaatje duidelijk anders dan op een röntgenfoto waarbij gebruik is gemaakt van elektromagnetische straling met een veel kleinere golflengte. De 'constante' in het proton, datgene wat het proton definieert, is dat het aantal upquarks minus het aantal anti-upquarks gelijk is aan 2 en dat het aantal downquarks minus het aantal antidownquarks gelijk is aan 1.

De interne structuur van het proton is een gevolg van de *sterke interactie*. De sterke interactie is verantwoordelijk voor de sterke kernkracht die atoomkernen bijeenhoudt en is een van de vier fundamentale interacties in de natuur (de andere drie zijn de elektromagnetische en zwakke interactie en gravitatie). Op het niveau van een proton is de kracht van de sterke interactie respectievelijk 10^2 en 10^6 keer groter dan die van de elektromagnetische en zwakke interactie, vandaar de aanduiding 'sterk'. De sterke interactie wordt beschreven door een theorie die *kwantumchromodynamica* (QCD) heet. Deze naam refereert aan het Griekse woord 'chroma' dat kleur betekent. QCD beschrijft namelijk de dynamica van elementaire deeltjes die een *kleurlading* hebben. Kleurlading is enigszins analoog aan elektrische lading. Elektrisch geladen deeltjes, zoals elektronen, interageren met elkaar door de uitwisseling van fotonen. Quarks hebben naast een elektrische lading ook een kleurlading. Hierdoor kunnen ze ook interageren via de sterke interactie. De overbrenger van de sterke interactie is het gluon, de QCD-versie van het foton.

Een belangrijk verschil tussen fotonen en gluonen is dat het foton elektrisch neutraal is terwijl het gluon een kleurlading heeft. Dit laatste impliceert dat gluonen ook met zichzelf de sterke interactie kunnen aangaan. Deze eigenschap heeft tot gevolg dat – in tegenstelling tot de andere fundamentele interacties – de sterkte van de sterke interactie, die gekarakteriseerd wordt door de koppelingsparameter α_s, toeneemt naarmate de afstand tussen twee kleurladingen groter wordt. Omdat het op een gegeven moment energetisch gunstiger is om een quark-antiquark paar te creëren dan om de kleurladingen verder uit elkaar te brengen, komen partonen alleen voor in gebonden toestanden genaamd *hadronen* (protonen en neutronen zijn voorbeelden van hadronen). Dit fenomeen wordt *confinement* genoemd. De tegenhanger van confinement is *asymptotische vrijheid*: op zeer kleine lengteschalen, kleiner dan de protonstraal van ca. 1 fm $= 10^{-15}$ m, is de sterke interactie erg zwak; zo zwak zelfs dat de partonen zich nagenoeg vrij kunnen bewegen (α_s is dan bijna gelijk aan nul).

De partonen vormen de (fundamentele) vrijheidsgraden van QCD bij hoge energieën terwijl bij lage energieën de (effectieve) vrijheidsgraden worden gevormd door de hadronen. We begrijpen nog altijd niet goed hoe bepaalde eigenschappen van

hadronen kunnen worden verklaard in termen van de partonen. Zo is het bijvoorbeeld onduidelijk waar de protonmassa van 938 MeV $= 1,67 \times 10^{-27}$ kg precies vandaan komt. Het Higgsmechanisme verklaart enkel de massa's van de elementaire quarks die hooguit een paar MeV bedragen. De rest van de protonmassa, ca. 98%, komt van de bindingsenergie in het proton. Deze energie correspondeert met de kinetische en potentiële energie die zit opgeslagen in de 'zee' van quarks, antiquarks en gluonen en de grootte ervan wordt voornamelijk bepaald door de sterke interactie. Wanneer je op de weegschaal staat dan komt het getal vóór de komma ruwweg overeen met de totale bindingsenergie van alle protonen en neutronen in je lichaam ($m = E/c^2$); het getal achter de komma komt van je interactie met het Higgsveld (en het feit dat je niet door de weegschaal zakt is het gevolg van de elektromagnetische wisselwerking).

Waarom kunnen we de eigenschappen van hadronen niet gewoon uitrekenen? We hebben immers een theorie die de sterke interactie beschrijft, namelijk QCD. Het blijkt dat wat wel en wat niet met QCD kan worden uitgerekend afhangt van de energieschaal die gemoeid gaat met het proces dat we bekijken. Dankzij asymptotische vrijheid kunnen hoog-energetische interacties tussen partonen (deze interacties zijn zwak) worden uitgerekend met behulp van *perturbatietheorie*. Dit houdt in dat de waarschijnlijkheid dat een bepaald proces zich voordoet, gekarakteriseerd door de zogenaamde *botsingsdoorsnede*, kan worden geëxpandeerd in machten van α_s. Omdat bij voldoende hoge energieën de (dimensieloze) parameter α_s veel kleiner is dan 1, volstaat het meestal om slechts enkele termen in deze expansie uit te rekenen. Bijdrages aan deze termen worden grafisch weergegeven met de bekende Feynmandiagrammen. Echter, bij lage energieën kan α_s groter worden dan 1. In dat geval leveren hogere-orde termen grotere bijdrages dan de laagste-orde termen – de expansie convergeert niet en is daarom betekenisloos geworden. Aangezien de interne structuur van hadronen wordt vormgegeven door laag-energetische interacties tussen partonen (de partonen zijn immers sterk gebonden) kan deze niet worden doorgerekend met perturbatieve technieken. Dit betekent dat we bijvoorbeeld de massa van het proton niet met behulp van Feynmandiagrammen kunnen uitrekenen.

Zoals al eerder opgemerkt wordt de interne structuur van hadronen beschreven door PDFs. Deze functies bevatten informatie over de dynamica van quarks en gluonen in hadronen en kunnen van verschillende kinematische variabelen afhangen. Aangezien PDFs niet kunnen worden uitgerekend met behulp van perturbatieve QCD, moeten ze uit metingen aan botsingsexperimenten worden gehaald. In dit proefschrift bestuderen we meetbare grootheden, ook wel *observabelen* genoemd, die afhangen van verschillende typen PDFs. Door het meten van die observabelen kunnen we meer te weten komen over wat er zich precies afspeelt in hadronen. Bovendien stelt een groter begrip van de interne structuur van het proton ons in staat om bijvoorbeeld hoog-energetische botsingen tussen protonen in de Large Hadron Collider nauwkeuriger te kunnen beschrijven. Dit laatste kan van grote betekenis zijn voor de zoektocht naar nieuwe elementaire deeltjes.

Een voorbeeld van een observabele die de interne structuur van hadronen 'voelt' is de botsingsdoorsnede van het zogenaamde *Drell-Yan* (DY) proces. DY verstrooiing is een veelvoorkomend scenario in hoog-energetische botsingen tussen hadronen; tijdens zo'n botsing gaan een quark uit het ene en een antiquark uit het andere hadron een interactie met elkaar aan en produceren vervolgens een elektrisch geladen lepton-antilepton paar (een voorbeeld hiervan is een elektron-positron paar), zie figuur 2. Doordat de impulsen van de leptonen gemeten kunnen worden in deeltjesdetectoren kan de botsing worden gereconstrueerd. Bij de reconstructie is het van belang met welke kans het quark en het antiquark met een bepaalde impuls in de hadronen voorkomen – dit is precies de informatie die bevat is in de PDFs. Aangezien de interactie tussen het quark en antiquark plaatsvindt bij hoge energie (dit wordt ook wel het 'harde' proces genoemd) kan deze perturbatief worden uitgerekend. De totale botsingsdoorsnede van DY verstrooiing wordt uiteindelijk gegeven door het product van de harde botsingsdoorsnede met twee PDFs (eentje voor het quark en eentje voor het antiquark). De mogelijkheid om een botsingsdoorsnede te schrijven als een product betekent dat het proces kan worden opgesplitst in afzonderlijke bijdrages die niet met elkaar 'praten'. Dit wordt *factorisatie* genoemd. Factorisatie is essentieel voor de experimentele bepaling van PDFs en vormt het fundament van de hadronfysica.
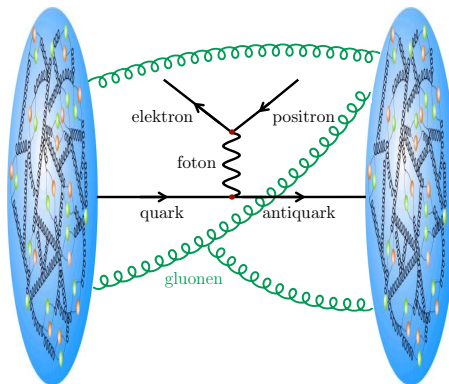
Figuur 2: Een voorbeeld van een DY proces: een quark en antiquark uit twee botsende protonen produceren een foton dat vervolgens vervalt in een elektron-positron paar. De sterk Lorentz-gecontraheerde protonen wisselen een willekeurig aantal gluonen met elkaar uit.

De eigenschap van factorisatie is echter niet vanzelfsprekend. Om de totale botsingsdoorsnede van het DY proces uit te rekenen, moet er volgens het super-positieprincipe in de kwantummechanica gesommeerd worden over alle mogelijke interacties die zouden kunnen plaatsvinden, ook die tussen de hadronen. Aangezien er bijvoorbeeld willekeurig veel laag-energetische ('zachte') gluonen tussen de ha-dronen kunnen worden uitgewisseld, is het a priori onduidelijk dat de bijdrages van de twee hadronen volledig opgesplitst kunnen worden. Met andere woorden, het is

maar de vraag of dat factorisatie is toegestaan. Dat dit wel zo is, werd in de jaren 1980 bewezen door Collins, Soper en Sterman (CSS) voor collineaire PDFs. Het feit dat interferentie-effecten tussen de twee hadronen inderdaad verwaarloosbaar zijn is intuïtief goed te begrijpen. Immers, omdat de hadronen met bijna de lichtsnelheid op elkaar afbewegen (het zijn dus relativistische objecten), vindt er lengtecontractie plaats in de bewegingsrichting alsook tijddilatatie van de interacties in de hadronen. Hierdoor kunnen de hadronen worden beschouwd als twee 'bevroren pannenkoeken' die elkaar nauwelijks 'voelen' voorafgaand aan de botsing. Factorisatie impliceert dat collineaire PDFs *universeel* zijn; ze horen bij individuele hadronen en zijn volledig onafhankelijk van harde processen. Collineaire PDFs kunnen dus in een bepaald proces worden gemeten en vervolgens gebruikt worden voor de beschrijving van hadronen in andere processen.

Het blijkt dat universaliteit geen eigenschap is van alle soorten PDFs. In dit proefschrift kijken we naar *multidimensionale* PDFs die naast de longitudinale component x ook afhangen van de transversale partonimpuls k. Deze PDFs bevatten meer informatie over de interne structuur van hadronen dan de eendimensionale collineaire PDFs en kunnen verschillen van proces tot proces. De procesafhankelijkheid van multidimensionale PDFs kan worden uitgerekend en geïsoleerd in een zogenaamde *ijklink* die afhangt van de richting waarin kleurladingen zich verplaatsen in het harde proces. Een voorbeeld van zo'n niet-universele, multidimensionale PDF is de zogenaamde *Boer-Mulders* (BM) functie $h_{1,i}^{\perp q}(x, \boldsymbol{k}, Q)$. Deze (driedimensionale) functie kwantificeert de netto transversale spinpolarisatie van een quark of antiquark van het type i in een ongepolariseerd hadron. Het BM effect genereert in het DY proces een heel specifieke asymmetrie in de richtingen van de leptonimpulsen in het transversale vlak. Hoewel deze asymmetrie in de DY botsingsdoorsnede door verschillende experimenten is gemeten, is voor een nauwkeurige extractie van de BM functie (onder andere) meer data nodig.

In deel ɪ van dit proefschrift kijken we naar kleureffecten die mogelijk de factorisatie-eigenschap van de BM asymmetrie in het DY proces beïnvloeden. Deze effecten zijn afkomstig van 'kleurverstrengelde' gluonen die tussen de hadronen worden uitgewisseld en kunnen in het ergste geval aanleiding geven tot een schending van factorisatie. Aangezien het BM effect door verschillende experimentele collaboraties actief wordt onderzocht, is het belangrijk om te weten hoe het zich precies manifesteert in de botsingsdoorsnede. Door grotendeels de stappen in het CSS bewijs te volgen, vinden we dat de bijdrage van het BM effect aan de DY botsingsdoorsnede op de 'standaardwijze' factoriseert. Dit houdt in dat de BM asymmetrie geschreven kan worden als een product van een harde botsingsdoorsnede met quark en antiquark BM functies, zonder extra complicaties.

Deel ɪɪ gaat over multidimensionale distributies van gluonen in hadronen. Bij heel hoge botsingsenergieën wordt de interne structuur van hadronen gedomineerd door gluonen; dit correspondeert met het regime van kleine waardes van de impulscomponent x. We introduceren verschillende nieuwe multidimensionale gluondistributies

voor spin-1 hadronen die interessante informatie bevatten over correlaties tussen transversale gluonimpulsen en spincomponenten. Omdat gluon PDFs nog grotendeels onbekend zijn, leiden we ook limieten af voor maximale waardes die bepaalde functies mogen aannemen. Voor heel kleine waardes van x blijken verschillende gluon PDFs aan elkaar gerelateerd te zijn. Dit resulteert in een simpelere beschrijving van de distributies van gluonen in hadronen bij hoge energieën. Tot slot illustreren we met een modelberekening dat bepaalde asymmetrieën die zijn waargenomen in de hoekverdeling van twee geproduceerde hadronen in botsingen tussen protonen en loodkernen mogelijk hun oorsprong vinden in kwantumcorrelaties die bestaan tussen gluonen in de loodkernen.

Tegenwoordig weten we redelijk goed hoe het proton er met een 'eendimensionale bril' op vanbinnen uitziet. Door een combinatie van experimenteel en theoretisch onderzoek sinds de jaren 1970 zijn namelijk de eendimensionale collineaire PDFs – voor niet al te kleine waardes van x – vrij precies bekend. Sinds de laatste eeuwwisseling heeft het onderzoeksveld zich uitgebreid naar metingen aan de multidimensionale interne structuur van hadronen en inmiddels zijn er ook een paar multidimensionale PDFs bekend. We verwachten de komende jaren meer te zullen leren over interessante kwantumeffecten in hadronen zoals het BM effect dat we onderzoeken in deel I van dit proefschrift. Aangezien er metingen worden verricht aan verschillende botsingsprocessen, zouden we tevens enkele QCD voorspellingen met betrekking tot de procesafhankelijke ijklinks kunnen testen. Voor het onderzoek bij kleine waardes van x zou de toekomstige Electron-Ion Collider (EIC) van groot belang kunnen zijn. Deze deeltjesversneller zou bij heel hoge energie elektronen op gepolariseerde protonen en lichte atoomkernen kunnen laten botsen en daarmee nieuw licht kunnen werpen op de gluonstructuur van hadronen. Met de EIC zouden verschillende theoretische voorspellingen uit deel II van dit proefschrift experimenteel getoetst kunnen worden. Kortom, het onderzoek naar de eigenschappen van de deeltjessoep in hadronen heeft nog een lange weg te gaan. Echter, we bevinden ons in een interessante tijd waarin we stapje voor stapje de rijke multidimensionale interne structuur van hadronen ontrafelen en daarmee meer leren over de belangrijkste bouwstenen van ons (zichtbare) universum.

ACADEMIC BIOGRAPHY

In 2008, Tom enrolled in the Bachelor's program of Physics & Astronomy at Radboud University Nijmegen. After obtaining his B.Sc. degree in 2011, he continued with a Master's specialization in Theoretical Physics at the same university. In 2012, he spent a summer at CERN in Geneva, Switzerland and contributed to the search for dark matter particles at the ATLAS experiment. A year later, he participated in the International Student Exchange Program and spent one semester at New Mexico State University in Las Cruces, United States. During the final year of his Master's program in Nijmegen he conducted research at the department of Theoretical High-Energy Physics on renormalization group invariant quantities in supersymmetric models under the supervision of Prof. dr. W.J.P. Beenakker. Finally, in 2014, he obtained his M.Sc. degree cum laude.

After his graduation, Tom decided to pursue a Ph.D. degree at the National Institute for Subatomic Physics in Amsterdam under the supervision of Prof. dr. D. Boer and Prof. dr. P.J.G. Mulders. This thesis is the result of four years of theoretical research on the multidimensional inner structure of hadrons, which is shaped by the strong interaction. Besides research, Tom was involved as a teaching assistant for various courses at VU University Amsterdam and he was a member of the PhD National Student Council.

ACKNOWLEDGMENTS

This work would not have been possible without the help and support of various people, both at a personal and a professional level.

First and foremost, I would like to thank my supervisors Daniël Boer and Piet Mulders. Daniël and Piet, it was a great pleasure to be part of your groups. You gave me the opportunity to work on interesting projects and encouraged me to attend multiple national and international schools, conferences, and workshops, which helped me greatly to develop myself as a scientist. I also truly appreciate all the trust and freedom that I was granted right from the start and the useful advice you gave me throughout the years. Daniël, I always very much enjoyed my visits to Groningen where I would learn a lot from your intuitive explanations of complicated concepts. Your structured way of working was a delight for me and your detailed reading of this thesis was very helpful. Piet, your enthusiasm is contagious and continues to inspire me. Your door was always open and no matter how busy you were, you would always make time for me. And for questions on the Boer-Mulders effect, who better to ask than either of you!

I also highly benefited from three 'mentors' during my PhD: Tomas Kasemets, Jonathan Gaunt, and Elena Petreska. Tomas, Jo, and Elena, it was a great pleasure to work together with you and this thesis would not have been remotely the same if it weren't for you. Tomas, as my longest-overlapping office mate you assisted me in various ways. I appreciate your help in getting me started in the field of TMDs and your guidance in the DY calculations. Jo, although Glaubers are "a pain in the . . . ", I believe they do give that extra spark to the results in our paper. I learned a lot from discussions with you about factorization and your feedback on part I of this thesis was very useful. Elena, you introduced me to the wonderful world of small-x physics for which I am very grateful. Your expertise in the CGC project was vital, as well as your feedback on part II of this thesis. All three of you are great scientists and I wish you the best of luck in your future careers.

Furthermore, I enjoyed very interesting and fruitful collaborations with Andrea Signori, Sabrina Cotogno, and Ya-Jin Zhou, and further benefited from discussions with Asmita Mukherjee, Cristian Pisano, Florent Scarpa, Jian Zhou, Maarten Buffing, and Miguel García Echevarría. I am indebted to Michael Engelhardt and Matthias Burkardt for their hospitality at New Mexico State University.

I also would like to thank Andreas Schäfer, Ignazio Scimemi, and Rob Timmermans for their careful examination of this manuscript as members of the assessment committee.

The Nikhef theory group provided a very pleasant and stimulating work environment. Thank you Eric, Robert, Giulio, Ruben, Jordy, Bert, Gillian, Pedro, Jake, Eleftheria, Ben, Kristof, Jorinde, Lorenzo, Juan, Elisa, Michael W., Mathias, Evangelos,

BIBLIOGRAPHY

[1] D. Boer, S. Cotogno, T. van Daal, P. J. Mulders, A. Signori and Y.-J. Zhou, *Gluon and Wilson loop TMDs for hadrons of spin* $\leq$ 1, *JHEP* **10** (2016) 013, [1607.01654].

[2] D. Boer, T. van Daal, J. R. Gaunt, T. Kasemets and P. J. Mulders, *Colour unwound – disentangling colours for azimuthal asymmetries in Drell-Yan scattering*, *SciPost Phys.* **3** (2017) 040, [1709.04935].

[3] S. Cotogno, T. van Daal and P. J. Mulders, *Positivity bounds on gluon TMDs for hadrons of spin* $\leq$ 1, *JHEP* **11** (2017) 185, [1709.07827].

[4] D. Boer, T. van Daal, P. J. Mulders and E. Petreska, *Directed flow from C-odd gluon correlations at small x*, *JHEP* **07** (2018) 140, [1805.05219].

[5] M. Gell-Mann, *A Schematic Model of Baryons and Mesons*, *Phys. Lett.* **8** (1964) 214–215.

[6] G. Zweig, *An SU*(3) *model for strong interaction symmetry and its breaking. Version 1*, 1964.

[7] G. Zweig, *An SU*(3) *model for strong interaction symmetry and its breaking. Version 2*, in *Developments in the Quark Theory of Hadrons, Volume 1. 1964 - 1978* (D. Lichtenberg and S. P. Rosen, eds.), pp. 22–101. 1964.

[8] E. D. Bloom et al., *High-Energy Inelastic ep Scattering at 6-Degrees and 10-Degrees*, *Phys. Rev. Lett.* **23** (1969) 930–934.

[9] M. Breidenbach, J. I. Friedman, H. W. Kendall, E. D. Bloom, D. H. Coward, H. C. DeStaebler et al., *Observed Behavior of Highly Inelastic Electron-Proton Scattering*, *Phys. Rev. Lett.* **23** (1969) 935–939.

[10] M. Kobayashi and T. Maskawa, *CP Violation in the Renormalizable Theory of Weak Interaction*, *Prog. Theor. Phys.* **49** (1973) 652–657.

[11] CDF collaboration, F. Abe et al., *Observation of top quark production in* $p\bar{p}$ *collisions*, *Phys. Rev. Lett.* **74** (1995) 2626–2631, [hep-ex/9503002].

[12] TASSO collaboration, R. Brandelik et al., *Evidence for Planar Events in* e^+e^- *Annihilation at High-Energies*, *Phys. Lett.* **86B** (1979) 243–249.

[13] D. P. Barber et al., *Discovery of Three Jet Events and a Test of Quantum Chromodynamics at PETRA Energies*, *Phys. Rev. Lett.* **43** (1979) 830.

[14] PLUTO collaboration, C. Berger et al., *Evidence for Gluon Bremsstrahlung in e^+e^- Annihilations at High-Energies*, *Phys. Lett.* **86B** (1979) 418–425.

[15] JADE collaboration, W. Bartel et al., *Observation of Planar Three Jet Events in e^+e^- Annihilation and Evidence for Gluon Bremsstrahlung*, *Phys. Lett.* **91B** (1980) 142–147.

[16] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Nonabelian Gauge Theories*, *Phys. Rev. Lett.* **30** (1973) 1343–1346.

[17] H. D. Politzer, *Reliable Perturbative Results for Strong Interactions?*, *Phys. Rev. Lett.* **30** (1973) 1346–1349.

[18] R. P. Feynman, *Very high-energy collisions of hadrons*, *Phys. Rev. Lett.* **23** (1969) 1415–1417.

[19] H. Fritzsch, M. Gell-Mann and H. Leutwyler, *Advantages of the Color Octet Gluon Picture*, *Phys. Lett.* **47B** (1973) 365–368.

[20] M. Gell-Mann, *Symmetries of baryons and mesons*, *Phys. Rev.* **125** (1962) 1067–1084.

[21] P. A. Baikov, K. G. Chetyrkin and J. H. Kühn, *Five-Loop Running of the QCD coupling constant*, *Phys. Rev. Lett.* **118** (2017) 082002, [1606.08659].

[22] F. Herzog, B. Ruijl, T. Ueda, J. A. M. Vermaseren and A. Vogt, *The five-loop beta function of Yang-Mills theory with fermions*, *JHEP* **02** (2017) 090, [1701.01404].

[23] Particle Data Group collaboration, C. Patrignani et al., *Review of Particle Physics*, *Chin. Phys.* **C40** (2016) 100001.

[24] A. Zee, *Study of the renormalization group for small coupling constants*, *Phys. Rev.* **D7** (1973) 3630–3636.

[25] S. R. Coleman and E. J. Weinberg, *Radiative Corrections as the Origin of Spontaneous Symmetry Breaking*, *Phys. Rev.* **D7** (1973) 1888–1910.

[26] A. Signori, *Flavor and Evolution Effects in TMD Phenomenology*. PhD thesis, Vrije U., Amsterdam, 2016.

[27] V. N. Gribov and L. N. Lipatov, *Deep inelastic ep scattering in perturbation theory*, *Sov. J. Nucl. Phys.* **15** (1972) 438–450.

[28] Y. L. Dokshitzer, *Calculation of the Structure Functions for Deep Inelastic Scattering and e^+e^- Annihilation by Perturbation Theory in Quantum Chromodynamics.*, *Sov. Phys. JETP* **46** (1977) 641–653.

[29] G. Altarelli and G. Parisi, *Asymptotic Freedom in Parton Language*, *Nucl. Phys.* **B126** (1977) 298–318.

[30] J. Collins, *Foundations of perturbative QCD*. Cambridge University Press, 2013.

[31] K. G. Wilson, *Confinement of Quarks*, *Phys. Rev.* **D10** (1974) 2445–2459.

[32] A. V. Efremov and A. V. Radyushkin, *Field Theoretic Treatment of High Momentum Transfer Processes. 3. Gauge Theories*, *Theor. Math. Phys.* **44** (1981) 774.

[33] M. G. A. Buffing, *Color and TMD Universality in Hadronic Interactions*. PhD thesis, NIKHEF, Amsterdam, 2015-09-02.

[34] J. C. Collins and D. E. Soper, *Parton Distribution and Decay Functions*, *Nucl. Phys.* **B194** (1982) 445–492.

[35] R. D. Tangerman and P. J. Mulders, *Intrinsic transverse momentum and the polarized Drell-Yan process*, *Phys. Rev.* **D51** (1995) 3357–3372, [hep-ph/9403227].

[36] P. J. Mulders and R. D. Tangerman, *The Complete tree level result up to order* $1/Q$ *for polarized deep inelastic leptoproduction*, *Nucl. Phys.* **B461** (1996) 197–237, [hep-ph/9510301].

[37] NNPDF collaboration, R. D. Ball et al., *Parton distributions from high-precision collider data*, *Eur. Phys. J.* **C77** (2017) 663, [1706.00428].

[38] M. Burkardt and B. Pasquini, *Modelling the nucleon structure*, *Eur. Phys. J.* **A52** (2016) 161, [1510.02567].

[39] M. Diehl, *Introduction to GPDs and TMDs*, *Eur. Phys. J.* **A52** (2016) 149, [1512.01328].

[40] A. Bacchetta, *Where do we stand with a 3-D picture of the proton?*, *Eur. Phys. J.* **A52** (2016) 163.

[41] M. G. Echevarría, A. Idilbi, K. Kanazawa, C. Lorcé, A. Metz, B. Pasquini et al., *Proper definition and evolution of generalized transverse momentum dependent distributions*, *Phys. Lett.* **B759** (2016) 336–341, [1602.06953].

[42] M. Diehl, *Generalized parton distributions*, *Phys. Rept.* **388** (2003) 41–277, [hep-ph/0307382].

[43] S. Meißner, A. Metz and M. Schlegel, *Generalized parton correlation functions for a spin-$\frac{1}{2}$ hadron*, *JHEP* **08** (2009) 056, [0906.5323].

[44] S. Bhattacharya, A. Metz and J. Zhou, *Generalized TMDs and the exclusive double Drell-Yan process*, *Phys. Lett.* **B771** (2017) 396–400, [1702.04387].

[45] X. Ji, *Viewing the proton through 'color' filters*, *Phys. Rev. Lett.* **91** (2003) 062001, [hep-ph/0304037].

[46] A. V. Belitsky, X. Ji and F. Yuan, *Quark imaging in the proton via quantum phase space distributions*, *Phys. Rev.* **D69** (2004) 074014, [hep-ph/0307383].

[47] C. Lorcé, B. Pasquini and M. Vanderhaeghen, *Unified framework for generalized and transverse-momentum dependent parton distributions within a 3Q light-cone picture of the nucleon*, *JHEP* **05** (2011) 041, [1102.4704].

[48] C. Lorcé, B. Pasquini, X. Xiong and F. Yuan, *The quark orbital angular momentum from Wigner distributions and light-cone wave functions*, *Phys. Rev.* **D85** (2012) 114006, [1111.4827].

[49] A. Mukherjee, S. Nair and V. K. Ojha, *Quark Wigner Distributions and Orbital Angular Momentum in Light-front Dressed Quark Model*, *Phys. Rev.* **D90** (2014) 014024, [1403.6233].

[50] T. Liu, *Quark orbital motions from Wigner distributions*, 1406.7709.

[51] T. Liu and B.-Q. Ma, *Quark Wigner distributions in a light-cone spectator model*, *Phys. Rev.* **D91** (2015) 034019, [1501.07690].

[52] C. Lorcé and B. Pasquini, *Multipole decomposition of the nucleon transverse phase space*, *Phys. Rev.* **D93** (2016) 034040, [1512.06744].

[53] J. More, A. Mukherjee and S. Nair, *Quark Wigner Distributions Using Light-Front Wave Functions*, *Phys. Rev.* **D95** (2017) 074039, [1701.00339].

[54] D. Chakrabarti, T. Maji, C. Mondal and A. Mukherjee, *Quark Wigner distributions and spin-spin correlations*, *Phys. Rev.* **D95** (2017) 074028, [1701.08551].

[55] S. Kaur and H. Dahiya, *Wigner distributions and GTMDs in a proton using light-front quark-diquark model*, 1810.09099.

[56] M. Diehl, *Generalized parton distributions with helicity flip*, *Eur. Phys. J.* **C19** (2001) 485–492, [hep-ph/0101335].

[57] X. Ji, *Gauge-Invariant Decomposition of Nucleon Spin*, *Phys. Rev. Lett.* **78** (1997) 610–613, [hep-ph/9603249].

[58] J. P. Ralston and D. E. Soper, *Production of Dimuons from High-Energy Polarized Proton-Proton Collisions*, *Nucl. Phys.* **B152** (1979) 109.

[59] V. Barone, F. Bradamante and A. Martin, *Transverse-spin and transverse-momentum effects in high-energy processes*, *Prog. Part. Nucl. Phys.* **65** (2010) 267–333, [1011.0909].

[60] A. Bacchetta and P. J. Mulders, *Deep inelastic leptoproduction of spin-one hadrons*, *Phys. Rev.* **D62** (2000) 114004, [hep-ph/0007120].

[61] A. Bacchetta, *Probing the Transverse Spin of Quarks in Deep Inelastic Scattering*. PhD thesis, Ph.D. thesis at VU University (Amsterdam), 2002.

[62] M. G. Echevarría, A. Idilbi and I. Scimemi, *Factorization Theorem For Drell-Yan At Low q_T And Transverse Momentum Distributions On-The-Light-Cone*, *JHEP* **07** (2012) 002, [1111.4996].

[63] D. Boer and P. J. Mulders, *Time reversal odd distribution functions in leptoproduction*, *Phys. Rev.* **D57** (1998) 5780–5786, [hep-ph/9711485].

[64] P. Hoodbhoy, R. L. Jaffe and A. Manohar, *Novel Effects in Deep Inelastic Scattering from Spin 1 Hadrons*, *Nucl. Phys.* **B312** (1989) 571–588.

[65] K. Goeke, A. Metz, P. V. Pobylitsa and M. V. Polyakov, *Lorentz invariance relations among parton distributions revisited*, *Phys. Lett.* **B567** (2003) 27–30, [hep-ph/0302028].

[66] K. Goeke, A. Metz and M. Schlegel, *Parameterization of the quark-quark correlator of a spin-$\frac{1}{2}$ hadron*, *Phys. Lett.* **B618** (2005) 90–96, [hep-ph/0504130].

[67] T. van Daal, *Quark and gluon TMD correlators in momentum and coordinate space*, in *22nd International Symposium on Spin Physics (SPIN 2016) Urbana, IL, USA, September 25-30, 2016*, 2016. 1612.06585.

[68] A. Bacchetta, M. Boglione, A. Henneman and P. J. Mulders, *Bounds on transverse momentum dependent distribution and fragmentation functions*, *Phys. Rev. Lett.* **85** (2000) 712–715, [hep-ph/9912490].

[69] A. Bacchetta and P. J. Mulders, *Positivity bounds on spin one distribution and fragmentation functions*, *Phys. Lett.* **B518** (2001) 85–93, [hep-ph/0104176].

[70] A. V. Belitsky, X. Ji and F. Yuan, *Final state interactions and gauge invariant parton distributions*, *Nucl. Phys.* **B656** (2003) 165–198, [hep-ph/0208038].

[71] D. Boer, P. J. Mulders and F. Pijlman, *Universality of T odd effects in single spin and azimuthal asymmetries*, *Nucl. Phys.* **B667** (2003) 201–241, [hep-ph/0303034].

[72] C. J. Bomhof, P. J. Mulders and F. Pijlman, *The Construction of gauge-links in arbitrary hard processes*, *Eur. Phys. J.* **C47** (2006) 147–162, [hep-ph/0601171].

[73] D. W. Sivers, *Single Spin Production Asymmetries from the Hard Scattering of Point-Like Constituents*, *Phys. Rev.* **D41** (1990) 83.

[74] D. W. Sivers, *Hard scattering scaling laws for single spin production asymmetries*, *Phys. Rev.* **D43** (1991) 261–263.

[75] S. J. Brodsky, D. S. Hwang and I. Schmidt, *Final state interactions and single spin asymmetries in semi-inclusive deep inelastic scattering*, *Phys. Lett.* **B530** (2002) 99–107, [hep-ph/0201296].

[76] HERMES collaboration, A. Airapetian et al., *Single-spin asymmetries in semi-inclusive deep-inelastic scattering on a transversely polarized hydrogen target*, *Phys. Rev. Lett.* **94** (2005) 012002, [hep-ex/0408013].

[77] J. C. Collins, *Leading twist single transverse-spin asymmetries: Drell-Yan and deep inelastic scattering*, *Phys. Lett.* **B536** (2002) 43–48, [hep-ph/0204004].

[78] M. G. A. Buffing, A. Mukherjee and P. J. Mulders, *Generalized Universality of Higher Transverse Moments of Quark TMD Correlators*, *Phys. Rev.* **D86** (2012) 074030, [1207.3221].

[79] M. G. A. Buffing, A. Mukherjee and P. J. Mulders, *Generalized Universality of Definite Rank Gluon Transverse Momentum Dependent Correlators*, *Phys. Rev.* **D88** (2013) 054027, [1306.5897].

[80] D. Boer, M. G. A. Buffing and P. J. Mulders, *Operator analysis of p_T-widths of TMDs*, *JHEP* **08** (2015) 053, [1503.03760].

[81] HERMES collaboration, A. Airapetian et al., *Observation of the Naive-T-odd Sivers Effect in Deep-Inelastic Scattering*, *Phys. Rev. Lett.* **103** (2009) 152002, [0906.3918].

[82] COMPASS collaboration, E. S. Ageev et al., *A New measurement of the Collins and Sivers asymmetries on a transversely polarised deuteron target*, *Nucl. Phys.* **B765** (2007) 31–70, [hep-ex/0610068].

[83] COMPASS collaboration, M. Alekseev et al., *Collins and Sivers asymmetries for pions and kaons in muon-deuteron DIS*, *Phys. Lett.* **B673** (2009) 127–135, [0802.2160].

[84] COMPASS collaboration, C. Adolph et al., *II – Experimental investigation of transverse spin asymmetries in μ-p SIDIS processes: Sivers asymmetries*, *Phys. Lett.* **B717** (2012) 383–389, [1205.5122].

[85] COMPASS collaboration, C. Adolph et al., *Collins and Sivers asymmetries in muonproduction of pions and kaons off transversely polarised protons*, *Phys. Lett.* **B744** (2015) 250–259, [1408.4405].

[86] COMPASS collaboration, C. Adolph et al., *Sivers asymmetry extracted in SIDIS at the hard scales of the Drell-Yan process at COMPASS*, *Phys. Lett.* **B770** (2017) 138–145, [1609.07374].

[87] Jefferson Lab Hall A collaboration, X. Qian et al., *Single Spin Asymmetries in Charged Pion Production from Semi-Inclusive Deep Inelastic Scattering on a Transversely Polarized ^{3}He Target*, *Phys. Rev. Lett.* **107** (2011) 072003, [1106.0363].

[88] STAR collaboration, L. Adamczyk et al., *Measurement of the transverse single-spin asymmetry in $p^{\uparrow}p \to W^{\pm}/Z^0$ at RHIC*, *Phys. Rev. Lett.* **116** (2016) 132301, [1511.06003].

[89] COMPASS collaboration, M. Aghasyan et al., *First measurement of transverse-spin-dependent azimuthal asymmetries in the Drell-Yan process*, *Phys. Rev. Lett.* **119** (2017) 112002, [1704.00488].

[90] A. Bacchetta, F. Delcarro, C. Pisano, M. Radici and A. Signori, *Extraction of partonic transverse momentum distributions from semi-inclusive deep-inelastic scattering, Drell-Yan and Z-boson production*, *JHEP* **06** (2017) 081, [1703.10157].

[91] A. Accardi et al., *Electron Ion Collider: The Next QCD Frontier*, *Eur. Phys. J.* **A52** (2016) 268, [1212.1701].

[92] R. Ent, *TMDs and GPDs at a future Electron-Ion Collider*, *Eur. Phys. J.* **A52** (2016) 162.

[93] D. Boer et al., *Gluons and the quark sea at high energies: Distributions, polarization, tomography*, 1108.1713.

[94] C. A. Aidala, S. D. Bass, D. Hasch and G. K. Mallot, *The Spin Structure of the Nucleon*, *Rev. Mod. Phys.* **85** (2013) 655–691, [1209.2803].

[95] R. Ángeles Martínez et al., *Transverse Momentum Dependent (TMD) parton distribution functions: status and prospects*, *Acta Phys. Polon.* **B46** (2015) 2501–2534, [1507.05267].

[96] M. Boglione and A. Prokudin, *Phenomenology of transverse spin: past, present and future*, *Eur. Phys. J.* **A52** (2016) 154, [1511.06924].

[97] J. C. Collins, D. E. Soper and G. F. Sterman, *Factorization for Short Distance Hadron-Hadron Scattering*, *Nucl. Phys.* **B261** (1985) 104–142.

[98] J. C. Collins, D. E. Soper and G. F. Sterman, *Soft Gluons and Factorization*, *Nucl. Phys.* **B308** (1988) 833–856.

[99] T. Becher and M. Neubert, *Drell-Yan Production at Small q_T, Transverse Parton Distributions and the Collinear Anomaly*, *Eur. Phys. J.* **C71** (2011) 1665, [1007.4005].

[100] M. G. Echevarría, A. Idilbi and I. Scimemi, *Soft and Collinear Factorization and Transverse Momentum Dependent Parton Distribution Functions*, *Phys. Lett.* **B726** (2013) 795–801, [1211.1947].

[101] M. G. A. Buffing and P. J. Mulders, *Color entanglement for azimuthal asymmetries in the Drell-Yan process*, *Phys. Rev. Lett.* **112** (2014) 092002, [1309.4681].

[102] M. Diehl, D. Ostermeier and A. Schäfer, *Elements of a theory for multiparton interactions in QCD*, *JHEP* **03** (2012) 089, [1111.0910].

[103] M. Diehl, J. R. Gaunt, D. Ostermeier, P. Plößl and A. Schäfer, *Cancellation of Glauber gluon exchange in the double Drell-Yan process*, *JHEP* **01** (2016) 076, [1510.08696].

[104] J. R. Gaunt, *Glauber Gluons and Multiple Parton Interactions*, *JHEP* **07** (2014) 110, [1405.2080].

[105] J. C. Collins, D. E. Soper and G. F. Sterman, *Factorization of Hard Processes in QCD*, *Adv. Ser. Direct. High Energy Phys.* **5** (1989) 1–91, [hep-ph/0409313].

[106] G. F. Sterman, *Partons, factorization and resummation*, in *QCD and beyond. Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics, TASI-95, Boulder, USA, June 4-30, 1995*, pp. 327–408, 1995. hep-ph/9606312.

[107] F. Dominguez, *Unintegrated Gluon Distributions at Small-x*. PhD thesis, Columbia U., 2011.

[108] T. C. Rogers, *An overview of transverse-momentum-dependent factorization and evolution*, *Eur. Phys. J.* **A52** (2016) 153, [1509.04766].

[109] D. Bonocore, *Next-to-Soft Factorization and Unitarity in Drell-Yan Processes*. PhD thesis, Amsterdam U., 2016.

[110] J. Qiu and G. F. Sterman, *Power corrections in hadronic scattering. 1. Leading $1/Q^2$ corrections to the Drell-Yan cross-section*, *Nucl. Phys.* **B353** (1991) 105–136.

[111] J. Qiu and G. F. Sterman, *Power corrections to hadronic scattering. 2. Factorization*, *Nucl. Phys.* **B353** (1991) 137–164.

[112] J. C. Collins and D. E. Soper, *Back-To-Back Jets in QCD*, *Nucl. Phys.* **B193** (1981) 381.

[113] X. Ji, J.-P. Ma and F. Yuan, *QCD factorization for semi-inclusive deep-inelastic scattering at low transverse momentum*, *Phys. Rev.* **D71** (2005) 034005, [hep-ph/0404183].

[114] M. G. A. Buffing, M. Diehl and T. Kasemets, *Transverse momentum in double parton scattering: factorisation, evolution and matching*, *JHEP* **01** (2018) 044, [1708.03528].

[115] S. D. Drell and T.-M. Yan, *Massive Lepton Pair Production in Hadron-Hadron Collisions at High-Energies*, *Phys. Rev. Lett.* **25** (1970) 316–320.

[116] D. Boer, *Investigating the origins of transverse spin asymmetries at RHIC*, *Phys. Rev.* **D60** (1999) 014012, [hep-ph/9902255].

[117] D. Boer, S. J. Brodsky and D. S. Hwang, *Initial state interactions in the unpolarized Drell-Yan process*, *Phys. Rev.* **D67** (2003) 054003, [hep-ph/0211110].

[118] J. C. Collins and D. E. Soper, *Angular Distribution of Dileptons in High-Energy Hadron Collisions*, *Phys. Rev.* **D16** (1977) 2219.

[119] S. B. Libby and G. F. Sterman, *Mass Divergences in Two Particle Inelastic Scattering*, *Phys. Rev.* **D18** (1978) 4737.

[120] G. F. Sterman, *Mass Divergences in Annihilation Processes. 1. Origin and Nature of Divergences in Cut Vacuum Polarization Diagrams*, *Phys. Rev.* **D17** (1978) 2773.

[121] L. D. Landau, *On analytic properties of vertex parts in quantum field theory*, *Nucl. Phys.* **13** (1959) 181–192.

[122] S. Coleman and R. E. Norton, *Singularities in the physical region*, *Nuovo Cim.* **38** (1965) 438–442.

[123] I. Z. Rothstein and I. W. Stewart, *An Effective Field Theory for Forward Scattering and Factorization Violation*, *JHEP* **08** (2016) 025, [1601.04695].

[124] A. V. Manohar and I. W. Stewart, *The Zero-Bin and Mode Factorization in Quantum Field Theory*, *Phys. Rev.* **D76** (2007) 074002, [hep-ph/0605001].

[125] J.-Y. Chiu, A. Jain, D. Neill and I. Z. Rothstein, *The Rapidity Renormalization Group*, *Phys. Rev. Lett.* **108** (2012) 151601, [1104.0881].

[126] J.-Y. Chiu, A. Jain, D. Neill and I. Z. Rothstein, *A Formalism for the Systematic Treatment of Rapidity Logarithms in Quantum Field Theory*, *JHEP* **05** (2012) 084, [1202.0814].

[127] T. Becher, A. Broggio and A. Ferroglia, *Introduction to Soft-Collinear Effective Theory*, *Lect. Notes Phys.* **896** (2015) pp.1–206, [1410.1892].

[128] G. Grammer, Jr. and D. R. Yennie, *Improved treatment for the infrared divergence problem in quantum electrodynamics*, *Phys. Rev.* **D8** (1973) 4332–4344.

[129] M. Luo, J. Qiu and G. F. Sterman, *Anomalous nuclear enhancement in deeply inelastic scattering and photoproduction*, *Phys. Rev.* **D50** (1994) 1951–1971.

[130] D. Boer and P. J. Mulders, *Color gauge invariance in the Drell-Yan process*, *Nucl. Phys.* **B569** (2000) 505–526, [hep-ph/9906223].

[131] R. E. Cutkosky, *Singularities and discontinuities of Feynman amplitudes*, *J. Math. Phys.* **1** (1960) 429–433.

[132] M. J. G. Veltman, *Diagrammatica: The Path to Feynman rules*, Cambridge Lect. Notes Phys. **4** (1994) 1–284.

[133] R. J. Rietkerk, *Unitarity Methods and On-Shell Particles in Scattering Amplitudes*. PhD thesis, Amsterdam U., 2016.

[134] R. Ángeles Martínez, *Coulomb gluons and the ordering variable*. PhD thesis, The University of Manchester, 2015-12-27.

[135] J. C. Collins and A. Metz, *Universality of soft and collinear factors in hard-scattering factorization*, *Phys. Rev. Lett.* **93** (2004) 252001, [hep-ph/0408249].

[136] A. Vladimirov, *Structure of rapidity divergences in multi-parton scattering soft factors*, *JHEP* **04** (2018) 045, [1707.07606].

[137] J. Collins and J. Qiu, k_T *factorization is violated in production of high-transverse-momentum particles in hadron-hadron collisions*, *Phys. Rev.* **D75** (2007) 114014, [0705.2141].

[138] J. Collins, *2-soft-gluon exchange and factorization breaking*, 0708.4410.

[139] T. C. Rogers and P. J. Mulders, *No Generalized TMD-Factorization in Hadro-Production of High Transverse Momentum Hadrons*, *Phys. Rev.* **D81** (2010) 094006, [1001.2977].

[140] J. C. Collins, *Fragmentation of transversely polarized quarks probed in transverse momentum distributions*, *Nucl. Phys.* **B396** (1993) 161–182, [hep-ph/9208213].

[141] B. U. Musch, P. Hägler, M. Engelhardt, J. W. Negele and A. Schäfer, *Sivers and Boer-Mulders observables from lattice QCD*, *Phys. Rev.* **D85** (2012) 094510, [1111.4249].

[142] M. Engelhardt, P. Hägler, B. Musch, J. Negele and A. Schäfer, *Lattice QCD study of the Boer-Mulders effect in a pion*, *Phys. Rev.* **D93** (2016) 054501, [1506.07826].

[143] L. P. Gamberg, G. R. Goldstein and K. A. Oganessyan, *Novel transversity properties in semi-inclusive deep inelastic scattering*, *Phys. Rev.* **D67** (2003) 071504, [hep-ph/0301018].

[144] Z. Lu and B.-Q. Ma, *Non-zero transversity distribution of the pion in a quark-spectator-antiquark model*, *Phys. Rev.* **D70** (2004) 094044, [hep-ph/0411043].

[145] L. P. Gamberg and G. R. Goldstein, *T-odd effects in unpolarized Drell-Yan scattering*, *Phys. Lett.* **B650** (2007) 362–368, [hep-ph/0506127].

[146] Z. Lu and B.-Q. Ma, *Azimuthal asymmetry in unpolarized πN Drell-Yan process*, *Phys. Lett.* **B615** (2005) 200–206, [hep-ph/0504184].

[147] Z. Lu, B.-Q. Ma and I. Schmidt, *Flavor separation of the Boer-Mulders function from unpolarized $\pi^- p$ and $\pi^- D$ Drell-Yan processes*, *Phys. Lett.* **B639** (2006) 494–498, [hep-ph/0702006].

[148] L. P. Gamberg, G. R. Goldstein and M. Schlegel, *Transverse Quark Spin Effects and the Flavor Dependence of the Boer-Mulders Function*, *Phys. Rev.* **D77** (2008) 094016, [0708.0324].

[149] A. Bacchetta, F. Conti and M. Radici, *Transverse-momentum distributions in a diquark spectator model*, *Phys. Rev.* **D78** (2008) 074010, [0807.0323].

[150] A. Courtoy, S. Scopetta and V. Vento, *Analyzing the Boer-Mulders function within different quark models*, *Phys. Rev.* **D80** (2009) 074032, [0909.1404].

[151] B. Pasquini and F. Yuan, *Sivers and Boer-Mulders functions in Light-Cone Quark Models*, *Phys. Rev.* **D81** (2010) 114013, [1001.5398].

[152] B. Pasquini and P. Schweitzer, *Pion transverse momentum dependent parton distributions in a light-front constituent approach, and the Boer-Mulders effect in the pion-induced Drell-Yan process*, *Phys. Rev.* **D90** (2014) 014050, [1406.2056].

[153] Z. Wang, X. Wang and Z. Lu, *Boer-Mulders function of pion meson and q_T-weighted* cos 2ϕ *asymmetry in the unpolarized* $\pi^- p$ *Drell-Yan at COMPASS*, *Phys. Rev.* **D95** (2017) 094004, [1702.03637].

[154] X. Wang, W. Mao and Z. Lu, *Boer-Mulders effect in the unpolarized pion induced Drell–Yan process at COMPASS within TMD factorization*, *Eur. Phys. J.* **C78** (2018) 643, [1805.03017].

[155] NA10 collaboration, S. Falciano et al., *Angular Distributions of Muon Pairs Produced by 194-GeV/c Negative Pions*, *Z. Phys.* **C31** (1986) 513.

[156] NA10 collaboration, M. Guanziroli et al., *Angular Distributions of Muon Pairs Produced by Negative Pions on Deuterium and Tungsten*, *Z. Phys.* **C37** (1988) 545.

[157] J. S. Conway et al., *Experimental Study of Muon Pairs Produced by 252-GeV Pions on Tungsten*, *Phys. Rev.* **D39** (1989) 92–122.

[158] NuSea collaboration, L. Y. Zhu et al., *Measurement of Angular Distributions of Drell-Yan Dimuons in p-d Interactions at 800-GeV/c*, *Phys. Rev. Lett.* **99** (2007) 082301, [hep-ex/0609005].

[159] NuSea collaboration, L. Y. Zhu et al., *Measurement of Angular Distributions of Drell-Yan Dimuons in p-p Interactions at 800-GeV/c*, *Phys. Rev. Lett.* **102** (2009) 182001, [0811.4589].

[160] J.-C. Peng, *TMDs and Drell-Yan Experiments at Fermilab and J-PARC*, in *Transversity 2008: 2nd International Workshop on Transverse Polarization Phenomena in Hard Processes Ferrara, Italy, May 28-31, 2008*. 0811.4596.

[161] E906/SeaQuest collaboration, K. Nakano, *Measurement of Boer-Mulders Function via Drell-Yan Process by SeaQuest Experiment at Fermilab*, *Int. J. Mod. Phys. Conf. Ser.* **40** (2016) 1660041.

[162] B. Zhang, Z. Lu, B.-Q. Ma and I. Schmidt, *Extracting Boer-Mulders functions from p-D Drell-Yan processes*, *Phys. Rev.* **D77** (2008) 054011, [0803.1692].

[163] Z. Lu and I. Schmidt, *Updating Boer-Mulders functions from unpolarized pd and pp Drell-Yan data*, *Phys. Rev.* **D81** (2010) 034023, [0912.2031].

[164] V. Barone, S. Melis and A. Prokudin, *Azimuthal asymmetries in unpolarized Drell-Yan processes and the Boer-Mulders distributions of antiquarks*, *Phys. Rev.* **D82** (2010) 114025, [1009.3423].

[165] M. Lambertsen and W. Vogelsang, *Drell-Yan lepton angular distributions in perturbative QCD*, *Phys. Rev.* **D93** (2016) 114013, [1605.02625].

[166] V. Barone, S. Melis and A. Prokudin, *The Boer-Mulders effect in unpolarized SIDIS: An Analysis of the COMPASS and HERMES data on the* cos 2ϕ *asymmetry*, *Phys. Rev.* **D81** (2010) 114026, [0912.5194].

[167] V. Barone, M. Boglione, J. O. Gonzalez Hernandez and S. Melis, *Phenomenological analysis of azimuthal asymmetries in unpolarized semi-inclusive deep inelastic scattering*, *Phys. Rev.* **D91** (2015) 074019, [1502.04214].

[168] HERMES collaboration, A. Airapetian et al., *Multiplicities of charged pions and kaons from semi-inclusive deep-inelastic scattering by the proton and the deuteron*, *Phys. Rev.* **D87** (2013) 074029, [1212.5407].

[169] HERMES collaboration, A. Airapetian et al., *Azimuthal distributions of charged hadrons, pions, and kaons produced in deep-inelastic scattering off unpolarized protons and deuterons*, *Phys. Rev.* **D87** (2013) 012010, [1204.4161].

[170] COMPASS collaboration, C. Adolph et al., *Hadron Transverse Momentum Distributions in Muon Deep Inelastic Scattering at* 160 *GeV/c*, *Eur. Phys. J.* **C73** (2013) 2531, [1305.7317].

[171] COMPASS collaboration, C. Adolph et al., *Measurement of azimuthal hadron asymmetries in semi-inclusive deep inelastic scattering off unpolarised nucleons*, *Nucl. Phys.* **B886** (2014) 1046–1077, [1401.6284].

[172] Z. Lu, *Spin physics through unpolarized processes*, *Front. Phys.(Beijing)* **11** (2016) 111204.

[173] R. Jakob, P. J. Mulders and J. Rodrigues, *Modeling quark distribution and fragmentation functions*, *Nucl. Phys.* **A626** (1997) 937–965, [hep-ph/9704335].

[174] A. Bacchetta, A. Schäfer and J.-J. Yang, *Sivers function in a spectator model with axial vector diquarks*, *Phys. Lett.* **B578** (2004) 109–118, [hep-ph/0309246].

[175] S. Meißner, A. Metz and K. Goeke, *Relations between generalized and transverse momentum dependent parton distributions*, *Phys. Rev.* **D76** (2007) 034002, [hep-ph/0703176].

[176] T. Kasemets, *Double parton scattering – a tale of two partons*. PhD thesis, Hamburg U., 2013.

[177] G. Passarino and M. J. G. Veltman, *One Loop Corrections for e^+e^- Annihilation Into $\mu^+\mu^-$ in the Weinberg Model*, *Nucl. Phys.* **B160** (1979) 151–207.

[178] G. R. Goldstein and L. Gamberg, *Transversity and meson photoproduction*, in *High energy physics. Proceedings, 31st International Conference, ICHEP 2002, Amsterdam, Netherlands, July 25-31, 2002*, pp. 452–454. hep-ph/0209085.

[179] T. Kasemets and M. Diehl, *Angular correlations in the double Drell-Yan process*, *JHEP* **01** (2013) 121, [1210.5434].

[180] S. Arnold, A. Metz and M. Schlegel, *Dilepton production from polarized hadron-hadron collisions*, *Phys. Rev.* **D79** (2009) 034005, [0809.2262].

[181] D. Boer, W. J. den Dunnen and A. Kotzinian, *Double Sivers effect asymmetries and their impact on transversity measurements at RHIC*, *Phys. Rev.* **D83** (2011) 114032, [1103.0908].

[182] P. J. Mulders and J. Rodrigues, *Transverse momentum dependence in gluon distribution and fragmentation functions*, *Phys. Rev.* **D63** (2001) 094021, [hep-ph/0009343].

[183] R. L. Jaffe and A. Manohar, *Nuclear Gluonometry*, *Phys. Lett.* **B223** (1989) 218–224.

[184] X. Artru and M. Mekhfi, *Transversely Polarized Parton Densities, their Evolution and their Measurement*, *Z. Phys.* **C45** (1990) 669.

[185] C. J. Bomhof and P. J. Mulders, *Non-universality of transverse momentum dependent parton distribution functions*, *Nucl. Phys.* **B795** (2008) 409–427, [0709.1390].

[186] S. Cotogno, *Parametrization of the Transverse Momentum Dependent Light-Front Correlator for Gluons*, *Few Body Syst.* **58** (2017) 92.

[187] J. Soffer, *Positivity constraints for spin dependent parton distributions*, *Phys. Rev. Lett.* **74** (1995) 1292–1294, [hep-ph/9409254].

[188] W. Vogelsang, *Next-to-leading order evolution of transversity distributions and Soffer's inequality*, *Phys. Rev.* **D57** (1998) 1886–1894, [hep-ph/9706511].

[189] C. Bourrely, J. Soffer and O. V. Teryaev, *The Q^2 evolution of Soffer inequality*, *Phys. Lett.* **B420** (1998) 375–381, [hep-ph/9710224].

[190] O. Martin, A. Schäfer, M. Stratmann and W. Vogelsang, *Soffer's inequality and the transversely polarized Drell-Yan process at next-to-leading order*, *Phys. Rev.* **D57** (1998) 3084–3090, [hep-ph/9710300].

[191] M. G. Echevarría, A. Idilbi and I. Scimemi, *Unified treatment of the QCD evolution of all (un-)polarized transverse momentum dependent functions: Collins function as a study case*, *Phys. Rev.* **D90** (2014) 014003, [1402.0869].

[192] M. G. Echevarría, T. Kasemets, P. J. Mulders and C. Pisano, *QCD evolution of (un)polarized gluon TMDPDFs and the Higgs q_T-distribution*, *JHEP* **07** (2015) 158, [1502.05354].

[193] D. Boer, W. J. den Dunnen, C. Pisano and M. Schlegel, *Determining the Higgs spin and parity in the diphoton decay channel*, *Phys. Rev. Lett.* **111** (2013) 032002, [1304.2654].

[194] C. Pisano, D. Boer, S. J. Brodsky, M. G. A. Buffing and P. J. Mulders, *Linear polarization of gluons and photons in unpolarized collider experiments*, *JHEP* **10** (2013) 024, [1307.3417].

[195] D. Boer, P. J. Mulders, C. Pisano and J. Zhou, *Asymmetries in Heavy Quark Pair and Dijet Production at an EIC*, *JHEP* **08** (2016) 001, [1605.07934].

[196] D. Boer, C. Lorcé, C. Pisano and J. Zhou, *The gluon Sivers distribution: status and future prospects*, *Adv. High Energy Phys.* **2015** (2015) 371396, [1504.04332].

[197] D. Kharzeev, Y. V. Kovchegov and K. Tuchin, *Cronin effect and high-p_T suppression in pA collisions*, *Phys. Rev.* **D68** (2003) 094013, [hep-ph/0307037].

[198] F. Dominguez, B.-W. Xiao and F. Yuan, *k_T-factorization for Hard Processes in Nuclei*, *Phys. Rev. Lett.* **106** (2011) 022301, [1009.2141].

[199] F. Dominguez, C. Marquet, B.-W. Xiao and F. Yuan, *Universality of Unintegrated Gluon Distributions at small x*, *Phys. Rev.* **D83** (2011) 105005, [1101.0715].

[200] P. Kotko, K. Kutak, C. Marquet, E. Petreska, S. Sapeta and A. van Hameren, *Improved TMD factorization for forward dijet production in dilute-dense hadronic collisions*, *JHEP* **09** (2015) 106, [1503.03421].

[201] S. J. Brodsky, F. Fleuret, C. Hadjidakis and J.-P. Lansberg, *Physics Opportunities of a Fixed-Target Experiment using the LHC Beams*, *Phys. Rept.* **522** (2013) 239–255, [1202.6585].

[202] J.-P. Lansberg, C. Pisano, F. Scarpa and M. Schlegel, *Pinning down the linearly-polarised gluons inside unpolarised protons using quarkonium-pair production at the LHC*, *Phys. Lett.* **B784** (2018) 217–222, [1710.01684].

[203] I. Schmidt, J. Soffer and J.-J. Yang, *Transverse single spin asymmetries in photon production*, *Phys. Lett.* **B612** (2005) 258–262, [hep-ph/0503127].

[204] A. Bacchetta, C. Bomhof, U. D'Alesio, P. J. Mulders and F. Murgia, *The Sivers single-spin asymmetry in photon-jet production*, *Phys. Rev. Lett.* **99** (2007) 212002, [hep-ph/0703153].

[205] W. J. den Dunnen, J.-P. Lansberg, C. Pisano and M. Schlegel, *Accessing the Transverse Dynamics and Polarization of Gluons inside the Proton at the LHC*, *Phys. Rev. Lett.* **112** (2014) 212001, [1401.7611].

[206] J.-P. Lansberg et al., *Spin physics and TMD studies at A Fixed-Target ExpeRiment at the LHC (AFTER@LHC)*, *EPJ Web Conf.* **85** (2015) 02038, [1410.1962].

[207] J.-P. Lansberg and H.-S. Shao, *Double-quarkonium production at a fixed-target experiment at the LHC (AFTER@LHC)*, *Nucl. Phys.* **B900** (2015) 273–294, [1504.06531].

[208] S. Abeyratne et al., *Science Requirements and Conceptual Design for a Polarized Medium Energy Electron-Ion Collider at Jefferson Lab*, 1209.0757.

[209] S. Abeyratne et al., *MEIC Design Summary*, 1504.07961.

[210] J. Ball et al., *On the large COMPASS polarized deuteron target*, *Czech. J. Phys.* **56** (2006) F295–F305.

[211] W. Detmold, R. Jaffe, J. Maxwell, R. Milner, D. Crabb, D. Day et al., *Letter of intent*, *Jefferson Lab* **LOI12-14-001** (2014) .

[212] D. Boer, S. J. Brodsky, P. J. Mulders and C. Pisano, *Direct Probes of Linearly Polarized Gluons inside Unpolarized Hadrons*, *Phys. Rev. Lett.* **106** (2011) 132001, [1011.4225].

[213] J. Qiu, M. Schlegel and W. Vogelsang, *Probing Gluonic Spin-Orbit Correlations in Photon Pair Production*, *Phys. Rev. Lett.* **107** (2011) 062001, [1103.3861].

[214] D. Boer and C. Pisano, *Polarized gluon studies with charmonium and bottomonium at LHCb and AFTER*, *Phys. Rev.* **D86** (2012) 094007, [1208.3642].

[215] D. Boer, *Linearly polarized gluon effects in unpolarized collisions*, *PoS* **QCDEV2015** (2015) 023, [1510.05915].

[216] D. Boer, *Gluon TMD studies at EIC*, *EPJ Web Conf.* **112** (2016) 02003, [1601.01813].

[217] J.-P. Lansberg, C. Pisano and M. Schlegel, *Associated production of a dilepton and a $\Upsilon(J/\psi)$ at the LHC as a probe of gluon transverse momentum dependent distributions*, *Nucl. Phys.* **B920** (2017) 192–210, [1702.00305].

[218] R. M. Godbole, A. Kaushik, A. Misra and S. Padval, *Probing the Gluon Sivers Function through direct photon production at RHIC*, 1810.07113.

[219] D. Boer, M. G. Echevarría, P. Mulders and J. Zhou, *Single spin asymmetries from a single Wilson loop*, *Phys. Rev. Lett.* **116** (2016) 122001, [1511.03485].

[220] C. Lorcé and B. Pasquini, *Structure analysis of the generalized correlator of quark and gluon for a spin-$\frac{1}{2}$ target*, *JHEP* **09** (2013) 138, [1307.4497].

[221] Y. Hatta, B.-W. Xiao and F. Yuan, *Probing the Small-x Gluon Tomography in Correlated Hard Diffractive Dijet Production in Deep Inelastic Scattering*, *Phys. Rev. Lett.* **116** (2016) 202301, [1601.01585].

[222] Y. Hatta, E. Iancu, K. Itakura and L. McLerran, *Odderon in the color glass condensate*, *Nucl. Phys.* **A760** (2005) 172–207, [hep-ph/0501171].

[223] L. Łukaszuk and B. Nicolescu, *A Possible interpretation of pp rising total cross-sections*, *Lett. Nuovo Cim.* **8** (1973) 405–413.

[224] E. M. Levin and M. G. Ryskin, *High-energy hadron collisions in QCD*, *Phys. Rept.* **189** (1990) 267–382.

[225] E. Martynov and B. Nicolescu, *Did TOTEM experiment discover the Odderon?*, *Phys. Lett.* **B778** (2018) 414–418, [1711.03288].

[226] TOTEM collaboration, G. Antchev et al., *First determination of the ρ parameter at $\sqrt{s} = 13$ TeV – probing the existence of a colourless three-gluon bound state*, *Submitted to: Phys. Rev.* (2017) .

[227] T. Csörgő, R. Pasechnik and A. Ster, *Odderon and proton substructure from a model-independent Lévy imaging of elastic pp and pp̄ collisions*, 1807.02897.

[228] E. Martynov and B. Nicolescu, *Odderon effects in the differential cross-sections at Tevatron and LHC energies*, 1808.08580.

[229] R. D. Ball, V. Bertone, M. Bonvini, S. Marzani, J. Rojo and L. Rottoli, *Parton distributions with small-x resummation: evidence for BFKL dynamics in HERA data*, *Eur. Phys. J.* **C78** (2018) 321, [1710.05935].

[230] R. D. Ball, E. R. Nocera and J. Rojo, *The asymptotic behaviour of parton distributions at small and large x*, *Eur. Phys. J.* **C76** (2016) 383, [1604.00024].

[231] C. Marquet, E. Petreska and C. Roiesnel, *Transverse-momentum-dependent gluon distributions from JIMWLK evolution*, *JHEP* **10** (2016) 065, [1608.02577].

[232] Y. V. Kovchegov, D. Pitonyak and M. D. Sievert, *Small-x Asymptotics of the Gluon Helicity Distribution*, *JHEP* **10** (2017) 198, [1706.04236].

[233] Y. V. Kovchegov and M. D. Sievert, *Small-x Helicity Evolution: an Operator Treatment*, 1808.09010.

[234] M. G. Ryskin, *Odderon and Polarization Phenomena in QCD*, *Sov. J. Nucl. Phys.* **46** (1987) 337–342.

[235] N. H. Buttimore, B. Z. Kopeliovich, E. Leader, J. Soffer and T. L. Trueman, *The spin dependence of high-energy proton scattering*, *Phys. Rev.* **D59** (1999) 114010, [hep-ph/9901339].

[236] E. Leader and T. L. Trueman, *The Odderon and spin dependence of high-energy proton-proton scattering*, *Phys. Rev.* **D61** (2000) 077504, [hep-ph/9908221].

[237] J. Zhou, *Transverse single spin asymmetries at small x and the anomalous magnetic moment*, *Phys. Rev.* **D89** (2014) 074050, [1308.5912].

[238] E. Iancu, A. Leonidov and L. D. McLerran, *The Color glass condensate: An Introduction*, in *QCD perspectives on hot and dense matter. Proceedings, NATO Advanced Study Institute, Summer School, Cargese, France, August 6-18, 2001*, pp. 73–145. hep-ph/0202270.

[239] E. Iancu and R. Venugopalan, *The Color glass condensate and high-energy scattering in QCD*, in *Quark-gluon plasma 4* (R. C. Hwa and X.-N. Wang, eds.), pp. 249–3363. 2003. hep-ph/0303204.

[240] E. Wessels, *Signatures of gluon saturation in high energy scattering*. PhD thesis, Vrije U., 2009.

[241] F. Gelis, E. Iancu, J. Jalilian-Marian and R. Venugopalan, *The Color Glass Condensate*, *Ann. Rev. Nucl. Part. Sci.* **60** (2010) 463–489, [1002.0333].

[242] Y. V. Kovchegov and E. Levin, *Quantum chromodynamics at high energy*, vol. 33. Cambridge University Press, 2012.

[243] P. Taels, *Quantum Chromodynamics at small Bjorken-x*. PhD thesis, Antwerp U., 2017. 1711.03928.

[244] L. D. McLerran and R. Venugopalan, *Gluon distribution functions for very large nuclei at small transverse momentum*, *Phys. Rev.* **D49** (1994) 3352–3355, [hep-ph/9311205].

[245] L. D. McLerran and R. Venugopalan, *Computing quark and gluon distribution functions for very large nuclei*, *Phys. Rev.* **D49** (1994) 2233–2241, [hep-ph/9309289].

[246] L. D. McLerran and R. Venugopalan, *Green's functions in the color field of a large nucleus*, *Phys. Rev.* **D50** (1994) 2225–2233, [hep-ph/9402335].

[247] Y. V. Kovchegov, *Brief Review of Saturation Physics*, *Acta Phys. Polon.* **B45** (2014) 2241–2256, [1410.7722].

[248] C.-N. Yang and R. L. Mills, *Conservation of Isotopic Spin and Isotopic Gauge Invariance*, *Phys. Rev.* **96** (1954) 191–195.

[249] E. Iancu, A. Leonidov and L. D. McLerran, *Nonlinear gluon evolution in the color glass condensate. 1.*, *Nucl. Phys.* **A692** (2001) 583–645, [hep-ph/0011241].

[250] F. Gelis and A. Peshier, *Probing colored glass via qq̄ photoproduction*, *Nucl. Phys.* **A697** (2002) 879–901, [hep-ph/0107142].

[251] J. Jalilian-Marian, A. Kovner, A. Leonidov and H. Weigert, *The BFKL equation from the Wilson renormalization group*, *Nucl. Phys.* **B504** (1997) 415–431, [hep-ph/9701284].

[252] J. Jalilian-Marian, A. Kovner, A. Leonidov and H. Weigert, *The Wilson renormalization group for low x physics: Towards the high density regime*, *Phys. Rev.* **D59** (1998) 014014, [hep-ph/9706377].

[253] J. Jalilian-Marian, A. Kovner and H. Weigert, *The Wilson renormalization group for low x physics: Gluon evolution at finite parton density*, *Phys. Rev.* **D59** (1998) 014015, [hep-ph/9709432].

[254] H. Weigert, *Unitarity at small Bjorken x*, *Nucl. Phys.* **A703** (2002) 823–860, [hep-ph/0004044].

[255] E. Iancu, A. Leonidov and L. D. McLerran, *The Renormalization group equation for the color glass condensate*, *Phys. Lett.* **B510** (2001) 133–144, [hep-ph/0102009].

[256] E. Ferreiro, E. Iancu, A. Leonidov and L. D. McLerran, *Nonlinear gluon evolution in the color glass condensate. 2.*, *Nucl. Phys.* **A703** (2002) 489–538, [hep-ph/0109115].

[257] S. Jeon and R. Venugopalan, *A Classical Odderon in QCD at high energies*, *Phys. Rev.* **D71** (2005) 125003, [hep-ph/0503219].

[258] S. Jeon and R. Venugopalan, *Random walks of partons in $SU(N_c)$ and classical representations of color charges in QCD at small x*, *Phys. Rev.* **D70** (2004) 105012, [hep-ph/0406169].

[259] A. H. Mueller, *Small x Behavior and Parton Saturation: A QCD Model*, *Nucl. Phys.* **B335** (1990) 115–137.

[260] Y. V. Kovchegov and M. D. Sievert, *A New Mechanism for Generating a Single Transverse Spin Asymmetry*, *Phys. Rev.* **D86** (2012) 034028, [1201.5890].

[261] Y. V. Kovchegov, L. Szymanowski and S. Wallon, *Perturbative odderon in the dipole model*, *Phys. Lett.* **B586** (2004) 267–281, [hep-ph/0309281].

[262] ALICE collaboration, B. Abelev et al., *Long-range angular correlations on the near and away side in p-Pb collisions at $\sqrt{s_{NN}} = 5.02$ TeV*, *Phys. Lett.* **B719** (2013) 29–41, [1212.2001].

[263] ALICE collaboration, B. B. Abelev et al., *Long-range angular correlations of π, K and p in p-Pb collisions at* $\sqrt{s_{NN}}$ = 5.02 *TeV*, *Phys. Lett.* **B726** (2013) 164–177, [1307.3237].

[264] ALICE collaboration, B. B. Abelev et al., *Multiparticle azimuthal correlations in p-Pb and Pb-Pb collisions at the CERN Large Hadron Collider*, *Phys. Rev.* **C90** (2014) 054901, [1406.2474].

[265] ALICE collaboration, J. Adam et al., *Forward-central two-particle correlations in p-Pb collisions at* $\sqrt{s_{NN}}$ = 5.02 *TeV*, *Phys. Lett.* **B753** (2016) 126–139, [1506.08032].

[266] ALICE collaboration, S. Acharya et al., *Azimuthal anisotropy of heavy-flavour decay electrons in p-Pb collisions at* $\sqrt{s_{NN}}$ = 5.02 *TeV*, 1805.04367.

[267] ATLAS collaboration, G. Aad et al., *Observation of Associated Near-Side and Away-Side Long-Range Correlations in* $\sqrt{s_{NN}}$ = 5.02 *TeV Proton-Lead Collisions with the ATLAS Detector*, *Phys. Rev. Lett.* **110** (2013) 182302, [1212.5198].

[268] ATLAS collaboration, G. Aad et al., *Measurement with the ATLAS detector of multi-particle azimuthal correlations in p-Pb collisions at* $\sqrt{s_{NN}}$ = 5.02 *TeV*, *Phys. Lett.* **B725** (2013) 60–78, [1303.2084].

[269] ATLAS collaboration, G. Aad et al., *Measurement of long-range pseudorapidity correlations and azimuthal harmonics in* $\sqrt{s_{NN}}$ = 5.02 *TeV proton-lead collisions with the ATLAS detector*, *Phys. Rev.* **C90** (2014) 044906, [1409.1792].

[270] ATLAS collaboration, G. Aad et al., *Observation of Long-Range Elliptic Azimuthal Anisotropies in* $\sqrt{s}$ = 13 *and 2.76 TeV pp Collisions with the ATLAS Detector*, *Phys. Rev. Lett.* **116** (2016) 172301, [1509.04776].

[271] ATLAS collaboration, M. Aaboud et al., *Measurements of long-range azimuthal anisotropies and associated Fourier coefficients for p-p collisions at* $\sqrt{s}$ = 5.02 *and* 13 *TeV and p-Pb collisions at* $\sqrt{s_{NN}}$ = 5.02 *TeV with the ATLAS detector*, *Phys. Rev.* **C96** (2017) 024908, [1609.06213].

[272] ATLAS collaboration, M. Aaboud et al., *Measurement of multi-particle azimuthal correlations in p-p, p-Pb and low-multiplicity Pb-Pb collisions with the ATLAS detector*, *Eur. Phys. J.* **C77** (2017) 428, [1705.04176].

[273] ATLAS collaboration, M. Aaboud et al., *Measurement of long-range multiparticle azimuthal correlations with the subevent cumulant method in p-p and p-Pb collisions with the ATLAS detector at the CERN Large Hadron Collider*, *Phys. Rev.* **C97** (2018) 024904, [1708.03559].

[274] CMS collaboration, V. Khachatryan et al., *Observation of Long-Range Near-Side Angular Correlations in Proton-Proton Collisions at the LHC*, *JHEP* **09** (2010) 091, [1009.4122].

[275] CMS collaboration, S. Chatrchyan et al., *Observation of long-range near-side angular correlations in proton-lead collisions at the LHC*, *Phys. Lett.* **B718** (2013) 795–814, [1210.5482].

[276] CMS collaboration, S. Chatrchyan et al., *Multiplicity and transverse momentum dependence of two- and four-particle correlations in p-Pb and Pb-Pb collisions*, *Phys. Lett.* **B724** (2013) 213–240, [1305.0609].

[277] CMS collaboration, V. Khachatryan et al., *Measurement of long-range near-side two-particle angular correlations in pp collisions at $\sqrt{s} = 13$ TeV*, *Phys. Rev. Lett.* **116** (2016) 172302, [1510.03068].

[278] CMS collaboration, V. Khachatryan et al., *Evidence for transverse momentum and pseudorapidity dependent event plane fluctuations in Pb-Pb and p-Pb collisions*, *Phys. Rev.* **C92** (2015) 034911, [1503.01692].

[279] CMS collaboration, V. Khachatryan et al., *Evidence for Collective Multiparticle Correlations in p-Pb Collisions*, *Phys. Rev. Lett.* **115** (2015) 012301, [1502.05382].

[280] CMS collaboration, V. Khachatryan et al., *Evidence for collectivity in pp collisions at the LHC*, *Phys. Lett.* **B765** (2017) 193–220, [1606.06198].

[281] CMS collaboration, A. M. Sirunyan et al., *Pseudorapidity and transverse momentum dependence of flow harmonics in p-Pb and Pb-Pb collisions*, *Phys. Rev.* **C98** (2018) 044902, [1710.07864].

[282] CMS collaboration, A. M. Sirunyan et al., *Observation of Correlated Azimuthal Anisotropy Fourier Harmonics in p-p and p-Pb Collisions at the LHC*, *Phys. Rev. Lett.* **120** (2018) 092301, [1709.09189].

[283] LHCb collaboration, R. Aaij et al., *Measurements of long-range near-side angular correlations in $\sqrt{s_{NN}} = 5$ TeV proton-lead collisions in the forward region*, *Phys. Lett.* **B762** (2016) 473–483, [1512.00439].

[284] PHENIX collaboration, A. Adare et al., *Quadrupole Anisotropy in Dihadron Azimuthal Correlations in Central d-Au Collisions at $\sqrt{s_{NN}} = 200$ GeV*, *Phys. Rev. Lett.* **111** (2013) 212301, [1303.1794].

[285] PHENIX collaboration, A. Adare et al., *Measurement of long-range angular correlation and quadrupole anisotropy of pions and (anti)protons in central d-Au collisions at $\sqrt{s_{NN}} = 200$ GeV*, *Phys. Rev. Lett.* **114** (2015) 192301, [1404.7461].

[286] PHENIX collaboration, R. Belmont, *PHENIX results on multiparticle correlations in small systems*, *Nucl. Phys.* **A967** (2017) 341–344, [1704.04570].

[287] STAR collaboration, L. Adamczyk et al., *Effect of event selection on jetlike correlation measurement in d-Au collisions at $\sqrt{s_{NN}} = 200$ GeV*, *Phys. Lett.* **B743** (2015) 333–339, [1412.8437].

[288] STAR collaboration, L. Adamczyk et al., *Long-range pseudorapidity dihadron correlations in d-Au collisions at $\sqrt{s_{NN}} = 200$ GeV*, *Phys. Lett.* **B747** (2015) 265–271, [1502.07652].

[289] D. d'Enterria, G. K. Eyyubova, V. L. Korotkikh, I. P. Lokhtin, S. V. Petrushanko, L. I. Sarycheva et al., *Estimates of hadron azimuthal anisotropy from multiparton interactions in proton-proton collisions at $\sqrt{s} = 14$ TeV*, *Eur. Phys. J.* **C66** (2010) 173–185, [0910.3029].

[290] P. Bożek, *Elliptic flow in proton-proton collisions at $\sqrt{s} = 7$ TeV*, *Eur. Phys. J.* **C71** (2011) 1530, [1010.0405].

[291] P. Bożek and W. Broniowski, *Correlations from hydrodynamic flow in p-Pb collisions*, *Phys. Lett.* **B718** (2013) 1557–1561, [1211.0845].

[292] G.-Y. Qin and B. Müller, *Elliptic and triangular flow anisotropy in deuteron-gold collisions at $\sqrt{s_{NN}} = 200$ GeV at RHIC and in proton-lead collisions at $\sqrt{s_{NN}} = 5.02$ TeV at the LHC*, *Phys. Rev.* **C89** (2014) 044902, [1306.3439].

[293] K. Werner, M. Bleicher, B. Guiot, I. Karpenko and T. Pierog, *Evidence for Flow from Hydrodynamic Simulations of p-Pb Collisions at 5.02 TeV from v_2 Mass Splitting*, *Phys. Rev. Lett.* **112** (2014) 232301, [1307.4379].

[294] P. Bożek, W. Broniowski and G. Torrieri, *Mass hierarchy in identified particle distributions in proton-lead collisions*, *Phys. Rev. Lett.* **111** (2013) 172303, [1307.5060].

[295] P. Bożek and W. Broniowski, *Collective dynamics in high-energy proton-nucleus collisions*, *Phys. Rev.* **C88** (2013) 014903, [1304.3044].

[296] A. Bzdak and G.-L. Ma, *Elliptic and triangular flow in p-Pb and peripheral Pb-Pb collisions from parton scatterings*, *Phys. Rev. Lett.* **113** (2014) 252301, [1406.2804].

[297] R. D. Weller and P. Romatschke, *One fluid to rule them all: viscous hydrodynamic description of event-by-event central p-p, p-Pb and Pb-Pb collisions at $\sqrt{s} = 5.02$ TeV*, *Phys. Lett.* **B774** (2017) 351–356, [1701.07145].

[298] K. Dusling, W. Li and B. Schenke, *Novel collective phenomena in high-energy proton-proton and proton-nucleus collisions*, *Int. J. Mod. Phys.* **E25** (2016) 1630002, [1509.07939].

[299] E. Avsar, C. Flensburg, Y. Hatta, J.-Y. Ollitrault and T. Ueda, *Eccentricity and elliptic flow in proton-proton collisions from parton evolution*, *Phys. Lett.* **B702** (2011) 394–397, [1009.5643].

[300] A. Dumitru, K. Dusling, F. Gelis, J. Jalilian-Marian, T. Lappi and R. Venugopalan, *The Ridge in proton-proton collisions at the LHC*, *Phys. Lett.* **B697** (2011) 21–25, [1009.5295].

[301] A. Kovner and M. Lublinsky, *Angular Correlations in Gluon Production at High Energy*, *Phys. Rev.* **D83** (2011) 034017, [1012.3398].

[302] E. Levin and A. H. Rezaeian, *The Ridge from the BFKL evolution and beyond*, *Phys. Rev.* **D84** (2011) 034031, [1105.3275].

[303] E. Iancu and D. N. Triantafyllopoulos, *Higher-point correlations from the JIMWLK evolution*, *JHEP* **11** (2011) 105, [1109.0302].

[304] A. Kovner and M. Lublinsky, *On Angular Correlations and High Energy Evolution*, *Phys. Rev.* **D84** (2011) 094011, [1109.0347].

[305] K. Dusling and R. Venugopalan, *Azimuthal collimation of long range rapidity correlations by strong color fields in high multiplicity hadron-hadron collisions*, *Phys. Rev. Lett.* **108** (2012) 262001, [1201.2658].

[306] A. Kovner and M. Lublinsky, *Angular and long range rapidity correlations in particle production at high energy*, *Int. J. Mod. Phys.* **E22** (2013) 1330001, [1211.1928].

[307] K. Dusling and R. Venugopalan, *Comparison of the color glass condensate to dihadron correlations in proton-proton and proton-nucleus collisions*, *Phys. Rev.* **D87** (2013) 094034, [1302.7018].

[308] A. Kovner and A. H. Rezaeian, *Diphoton production in high-energy pA collisions*, *Phys. Rev.* **D90** (2014) 014031, [1404.5632].

[309] A. Dumitru, L. D. McLerran and V. Skokov, *Azimuthal asymmetries and the emergence of 'collectivity' from multi-particle correlations in high-energy pA collisions*, *Phys. Lett.* **B743** (2015) 134–137, [1410.4844].

[310] B. Schenke, S. Schlichting and R. Venugopalan, *Azimuthal anisotropies in p-Pb collisions from classical Yang-Mills dynamics*, *Phys. Lett.* **B747** (2015) 76–82, [1502.01331].

[311] T. Altinoluk, N. Armesto, G. Beuf and A. H. Rezaeian, *Diffractive Dijet Production in Deep Inelastic Scattering and Photon-Hadron Collisions in the Color Glass Condensate*, *Phys. Lett.* **B758** (2016) 373–383, [1511.07452].

[312] K. Dusling, P. Tribedy and R. Venugopalan, *Energy dependence of the ridge in high multiplicity proton-proton collisions*, *Phys. Rev.* **D93** (2016) 014034, [1509.04410].

[313] A. Kovner and A. H. Rezaeian, *Diphoton ridge in pp and pA collisions at RHIC and the LHC*, *Phys. Rev.* **D92** (2015) 074045, [1508.02412].

[314] B. Schenke, S. Schlichting, P. Tribedy and R. Venugopalan, *Mass ordering of spectra from fragmentation of saturated gluon states in high multiplicity proton-proton collisions*, *Phys. Rev. Lett.* **117** (2016) 162301, [1607.02496].

[315] A. H. Rezaeian, *Photon-jet ridge at RHIC and the LHC*, *Phys. Rev.* **D93** (2016) 094030, [1603.07354].

[316] K. Dusling, M. Mace and R. Venugopalan, *Parton model description of multiparticle azimuthal correlations in pA collisions*, *Phys. Rev.* **D97** (2018) 016014, [1706.06260].

[317] K. Dusling, M. Mace and R. Venugopalan, *Multiparticle collectivity from initial state correlations in high energy proton-nucleus collisions*, *Phys. Rev. Lett.* **120** (2018) 042002, [1705.00745].

[318] K. Fukushima and Y. Hidaka, *General formulae for dipole Wilson line correlators with the Color Glass Condensate*, *JHEP* **11** (2017) 114, [1708.03051].

[319] T. Altinoluk, N. Armesto and D. E. Wertepny, *Correlations and the ridge in the Color Glass Condensate beyond the glasma graph approximation*, *JHEP* **05** (2018) 207, [1804.02910].

[320] M. K. Davy, C. Marquet, Y. Shi, B.-W. Xiao and C. Zhang, *Two particle azimuthal harmonics in pA collisions*, 1808.09851.

[321] A. Bzdak, B. Schenke, P. Tribedy and R. Venugopalan, *Initial state geometry and the role of hydrodynamics in proton-proton, proton-nucleus and deuteron-nucleus collisions*, *Phys. Rev.* **C87** (2013) 064906, [1304.3403].

[322] A. Dumitru and A. V. Giannini, *Initial state angular asymmetries in high energy pA collisions: spontaneous breaking of rotational symmetry by a color electric field and C-odd fluctuations*, *Nucl. Phys.* **A933** (2015) 212–228, [1406.5781].

[323] A. Dumitru and V. Skokov, *Anisotropy of the semiclassical gluon field of a large nucleus at high energy*, *Phys. Rev.* **D91** (2015) 074006, [1411.6630].

[324] T. Lappi, *Azimuthal harmonics of color fields in a high energy nucleus*, *Phys. Lett.* **B744** (2015) 315–319, [1501.05505].

[325] T. Lappi, B. Schenke, S. Schlichting and R. Venugopalan, *Tracing the origin of azimuthal gluon correlations in the color glass condensate*, *JHEP* **01** (2016) 061, [1509.03499].

[326] A. Kovner, M. Lublinsky and V. Skokov, *Exploring correlations in the CGC wave function: odd azimuthal anisotropy*, *Phys. Rev.* **D96** (2017) 016010, [1612.07790].

[327] Y. V. Kovchegov and V. V. Skokov, *How classical gluon fields generate odd azimuthal harmonics for the two-gluon correlation function in high-energy collisions*, *Phys. Rev.* **D97** (2018) 094021, [1802.08166].

[328] Y. Hagiwara, Y. Hatta, B.-W. Xiao and F. Yuan, *Elliptic Flow in Small Systems due to Elliptic Gluon Distributions?*, *Phys. Lett.* **B771** (2017) 374–378, [1701.04254].

[329] E. Iancu and A. H. Rezaeian, *Elliptic flow from color-dipole orientation in pp and pA collisions*, *Phys. Rev.* **D95** (2017) 094003, [1702.03943].

[330] E. Petreska, *TMD gluon distributions at small x in the CGC theory*, *Int. J. Mod. Phys.* **E27** (2018) 1830003, [1804.04981].

[331] M. Strikman and W. Vogelsang, *Multiple parton interactions and forward double pion production in pp and dA scattering*, *Phys. Rev.* **D83** (2011) 034029, [1009.6123].

[332] A. Stasto, B.-W. Xiao and F. Yuan, *Back-to-Back Correlations of Di-hadrons in d-Au Collisions at RHIC*, *Phys. Lett.* **B716** (2012) 430–434, [1109.1817].

[333] T. Lappi and H. Mäntysaari, *Forward dihadron correlations in deuteron-gold collisions with the Gaussian approximation of JIMWLK*, *Nucl. Phys.* **A908** (2013) 51–72, [1209.2853].

[334] M. Diehl and J. R. Gaunt, *Double parton scattering theory overview*, 1710.04408.

[335] A. Dumitru, A. Hayashigaki and J. Jalilian-Marian, *The Color glass condensate and hadron production in the forward region*, *Nucl. Phys.* **A765** (2006) 464–482, [hep-ph/0506308].

[336] Y. Hagiwara, Y. Hatta and T. Ueda, *Wigner, Husimi, and generalized transverse momentum dependent distributions in the color glass condensate*, *Phys. Rev.* **D94** (2016) 094036, [1609.05773].

[337] J. Zhou, *Elliptic gluon generalized transverse-momentum-dependent distribution inside a large nucleus*, *Phys. Rev.* **D94** (2016) 114017, [1611.02397].

[338] Y. Hatta, B.-W. Xiao and F. Yuan, *Gluon Tomography from Deeply Virtual Compton Scattering at Small-x*, *Phys. Rev.* **D95** (2017) 114026, [1703.02085].

[339] Y. Hagiwara, Y. Hatta, R. Pasechnik, M. Tasevsky and O. Teryaev, *Accessing the gluon Wigner distribution in ultraperipheral pA collisions*, *Phys. Rev.* **D96** (2017) 034009, [1706.01765].

[340] B. Blok, Yu. Dokshitzer, L. Frankfurt and M. Strikman, *The Four jet production at LHC and Tevatron in QCD*, *Phys. Rev.* **D83** (2011) 071501, [1009.2714].

[341] S. Voloshin and Y. Zhang, *Flow study in relativistic nuclear collisions by Fourier expansion of Azimuthal particle distributions*, *Z. Phys.* **C70** (1996) 665–672, [hep-ph/9407282].

[342] R. D. Woods and D. S. Saxon, *Diffuse Surface Optical Model for Nucleon-Nuclei Scattering*, *Phys. Rev.* **95** (1954) 577–578.

[343] M. Diehl, T. Feldmann, R. Jakob and P. Kroll, *Generalized parton distributions from nucleon form-factor data*, *Eur. Phys. J.* **C39** (2005) 1–39, [hep-ph/0408173].

[344] M. Diehl, T. Kasemets and S. Keane, *Correlations in double parton distributions: effects of evolution*, *JHEP* **05** (2014) 118, [1401.1233].

[345] A. V. Belitsky, *Leading order analysis of the twist-three space- and time-like cut vertices in QCD*, in *Annual 31st PNPI Winter School on Nuclear and Particle Physics St. Petersburg, Russia, February 24-March 2, 1997*. hep-ph/9703432.

[346] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*. Addison-Wesley, Reading, USA, 1995.

[347] J. C. Romão and J. P. Silva, *A resource for signs and Feynman diagrams of the Standard Model*, *Int. J. Mod. Phys.* **A27** (2012) 1230025, [1209.6213].

[348] D. Boer, L. Gamberg, B. Musch and A. Prokudin, *Bessel-Weighted Asymmetries in Semi Inclusive Deep Inelastic Scattering*, *JHEP* **10** (2011) 021, [1107.5294].