

# Kinematic power corrections in TMD factorization theorem

---

**Alexey Vladimirov**

*Departamento de Física Teórica & IPARCOS, Universidad Complutense de Madrid,  
E-28040 Madrid, Spain*

*E-mail:* [alexeyvl@ucm.es](mailto:alexeyvl@ucm.es)

**ABSTRACT:** This work is dedicated to the study of power expansion in the transverse momentum dependent (TMD) factorization theorem. Each genuine term in this expansion gives rise to a series of kinematic power corrections (KPCs). All terms of this series exhibit the same properties as the leading term and share the same nonperturbative content. Among various power corrections, KPCs are especially important since they restore charge conservation and frame invariance, which are violated at a fixed power order. I derive and sum a series of KPCs associated with the leading-power term of the TMD factorization theorem. The resulting expression resembles a hadronic tensor computed with free massless quarks while still satisfying a proven factorization statement. Additionally, I provide an explicit check of this novel form of factorization theorem at the next-to-leading order (NLO) and demonstrate the restoration of the frame-invariant argument of the leading-power coefficient function. Numerical estimations show that incorporating the summed KPCs into the cross-section leads to an almost constant shift, which may help to explain the observed challenges in the TMD phenomenology.

**KEYWORDS:** Factorization, Renormalization Group, Resummation

**ARXIV EPRINT:** [2307.13054](https://arxiv.org/abs/2307.13054)

---

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Twist decomposition in TMD factorization</b>	<b>6</b>
2.1	TMD-twist	6
2.2	Example of twist-decomposition	8
2.3	Twist-decomposition for bi-quark TMD distribution	10
<b>3</b>	<b>KPCs for LP term</b>	<b>12</b>
3.1	TMD factorization at LP	13
3.2	KPC for LP term at all orders	16
3.3	KPCs beyond LO and the argument of the coefficient function	19
<b>4</b>	<b>Summation of KPC series</b>	<b>24</b>
4.1	Twist-two part of a general TMD correlator	24
4.2	Hadron tensor with summed KPCs	26
<b>5</b>	<b>KPC for the unpolarized Drell-Yan reaction</b>	<b>27</b>
5.1	Restoration of EM gauge invariance	29
5.2	Restoration of frame invariance	31
5.3	Estimation of numerical importance	32
<b>6</b>	<b>Conclusion</b>	<b>36</b>

---

## 1 Introduction

The transverse momentum dependent (TMD) factorization theorem is known to be an invaluable tool in understanding the transverse momentum distribution of inclusive and semi-inclusive processes across a wide range of energy regimes. The transverse momentum distribution of Z-boson or Higgs bosons, semi-inclusive deep-inelastic scattering (SIDIS), single- and double-spin asymmetries — these are only a few examples of successes of TMD factorization approach (for review see [1, 2]). As the list of accomplished applications continues to expand, it becomes evident that the current form of TMD factorization has practical limitations. It can only be confidently applied within a restricted phase-space, which leads to many problems and tensions already today (see f.i. discussions in [3–5]), and will be critical for the future generation of colliders [6, 7]. Therefore, a systematic extension and refinement of the TMD factorization approach are crucial priorities for modern studies.

A principal feature of the TMD factorization approach is its ability to systematically account for the transverse momenta of partons within the hadron. As a result, the main application domain is processes characterized by transverse momenta compatible with

typical hadronic scales. At the same time, the characteristic longitudinal momenta must be large enough to justify the application of the parton model. In this study, I focus on the Drell-Yan reaction ( $h_1 + h_2 \rightarrow \gamma^* + X$ ), which is the classical case of the TMD factorization [8]. In this particular reaction, the relevant momenta are the transverse ( $q_T$ ) and longitudinal ( $q^\pm$ ) components of the produced-photon momentum.

The leading term of the TMD factorization theorem is derived [9–11] in the limit of the longitudinal momenta being greater than any other scales. In the position-space representation, the leading term has the general form (omitting the scaling arguments for brevity)

$$\frac{d\sigma}{dq_T} \sim \sigma_0 \int_{-\infty}^{\infty} \frac{d^2b}{(2\pi)^2} e^{i(bq_T)} C_0 \tilde{F}_1(x_1, b) \tilde{F}_2(x_2, b), \quad (1.1)$$

where  $\tilde{F}_{1,2}$  are TMD parton distribution functions (TMDPDFs) in the Drell-Yan case,  $x_{1,2}$  are the longitudinal momentum fractions of partons.  $C_0$  is the hard coefficient function which depends on  $q^+q^-/\mu^2$  with  $\mu$  being the factorization scale. TMD distributions satisfy the pair of evolution equations [12–14], which have multiplicative kernels in the position space. Due to it, the position-space representation (1.1) is convenient in practice. The momentum-space representation of (1.1) is

$$\frac{d\sigma}{dq_T} \sim \sigma_0 \int_{-\infty}^{\infty} d^2k_{1T} d^2k_{2T} C_0 \delta^{(2)}(q_T - k_{1T} - k_{2T}) F_1(x_1, k_{1T}) F_2(x_2, k_{2T}), \quad (1.2)$$

where  $F_{1,2}$  are TMDPDFs in the momentum space obtained by the Fourier transform of  $\tilde{F}_{1,2}$ . Explicit examples of these structures for various processes can be found in refs. [15–17]. The perturbative elements of the factorization formula (1.1) (the hard coefficient function and the evolution kernels) are known up to next-to-next-to-next-to-next-to-leading order (N<sup>4</sup>LO), making it one of the most developed formula in the perturbation theory.

The corrections to the leading term are suppressed by powers of  $q^\pm$ . To simplify the exposition, I generally refer to the large scale as  $Q$  ( $\sim q^+ \sim q^-$ ). The power corrections can be categorized into four conceptual types:

**$k_T/Q$  power corrections:** in the position-space representation, these corrections manifest as transverse derivatives of TMD distributions, while in the momentum representation, as the powers of  $k_T$ 's. These corrections are commonly known as *kinematic power corrections* (KPCs).

**$q_T/Q$  power corrections:** in the position space, these corrections arise as terms accompanied by inverse powers of  $b$ . Once integrated, such factors transform into powers of  $q_T$ .

**$\Lambda/Q$  power corrections:** these corrections indicate terms that contain nonperturbative distributions of higher twist. Generally, a TMD distribution of twist- $n$  is considered as a  $(\Lambda/Q)^{n-2}$  correction. These corrections are often referred to as *higher twist* power corrections.

**Target-mass corrections:** these corrections account for the finite mass of the hadron. Currently, they are the least studied corrections.

Let me emphasize the difference between  $k_T/Q$  and  $q_T/Q$  power corrections since they were not distinguished until recently. Indeed, in the momentum-space representation (1.2), these corrections can be confused with each other due to the relation  $q_T = k_{1T} + k_{2T}$ . Nonetheless, they have conceptually different origins. The  $q_T/Q$ -corrections are related to the common ultra-violet (UV) singular behavior at  $b \rightarrow 0$  and are merely a part of the hard scattering sub-process. The  $k_T/Q$ -corrections are UV-finite and are a part of low-energy dynamics. In the position space, the  $k_T/Q$ -corrections appear as the transverse derivatives of  $\tilde{F}(x, b)$ 's that turn to parton's  $k_T$ 's in the momentum space representation, which suggests the name. Therefore, these corrections are entirely independent from the theory point of view. Practically, these power corrections are also distinct. KPCs are non-vanishing in the limit  $q_T \rightarrow 0$  (the traditional TMD limit) and remain of the same order for larger  $q_T$ . On the contrary, the  $q_T/Q$ -corrections vanish at  $q_T \rightarrow 0$  and grow at larger  $q_T$  (as suggested by name), and constitute the major part of the famous Y-term [2, 8, 9]. Below, I present more details on their definition.

Together, four types of power corrections form a complex system, which is discussed further in the remaining part of the introduction. Importantly, each type of power correction exhibits distinct theoretical and practical features, making some corrections more significant than others in different circumstances. In this study, I present the derivation, summation, and discussion of (pure) KPCs and examine their practical importance.

In recent years, there has been significant progress in the study of power corrections to the TMD factorization theorem [18–25]. This progress has been made possible by formulating the TMD factorization on the operator level, as opposed to the method-of-region analysis [9] used in the original derivation of the leading power (LP) term. As a matter of fact, the power corrections are simpler to analyze with operator methods since they allow for unambiguous identification of the different types of power corrections. This is crucial for a comprehensive understanding of the structure of the factorization theorem beyond the LP approximation. Still, the status of the TMD factorization theorem beyond the LP approximation remains an open problem. Currently, there exists a complete expression for TMD factorization at the next-to-leading power (NLP) level. It has been derived with three different methods [21–23] (and agrees between them), and checked at next-to-leading order (NLO) [22, 24, 26]. Furthermore, several general statements (such as the all-order structure of rapidity divergences for twist-three distributions [22, 23], the cancellation of special rapidity divergences [24, 26]) provide us hope that the NLP TMD factorization is valid at all perturbative orders. The research on the next-to-next-to-leading power (NNLP) TMD factorization is very limited [19–21].

The straight evaluation of power corrections is impractical. Already at NNLP, one faces many novel theoretical problems and a vast number of structures. Therefore, to proceed further, it is crucial to understand and account for the general hierarchy of power-suppressed terms in the TMD factorization framework. This hierarchy can be deduced from the power counting and the dimension analysis. Naturally, there is a degree of arbitrariness in such a generalization. Nonetheless, if TMD factorization holds at higher powers, one should

expect it to exhibit the following general structure (omitting target-mass corrections for simplicity)

$$\begin{aligned}
 W^{\mu\nu} = \frac{1}{N_c} \int \frac{d^2b}{(2\pi)^2} e^{-i(q_T b)} \left\{ \right. & \tag{1.3} \\
 & \Phi_2 \times \Phi_2 \\
 & + \frac{1}{Q} \left( D \Phi_2 \times \Phi_2 + \Phi_2 \times \Phi_3 \right) \\
 & + \frac{1}{Q^2} \left( D^2 \Phi_2 \times \Phi_2 + D \Phi_2 \times \Phi_3 + \Phi_3 \times \Phi_3 + \Phi_2 \times \Phi_4 + \frac{\Phi_2 \times \Phi_2}{b^2} \right) \\
 & + \frac{1}{Q^3} \left( D^3 \Phi_2 \times \Phi_2 + D^2 \Phi_2 \times \Phi_3 + D \Phi_3 \times \Phi_3 + D \Phi_2 \times \Phi_4 + \frac{D \Phi_2 \times \Phi_2}{b^2} + \Phi_3 \times \Phi_4 + \dots \right) \\
 & \left. + \dots \right\}^{\mu\nu},
 \end{aligned}$$

where  $\Phi_n$  is a TMD distribution of twist- $n$ , which depends on  $b$  and some number of collinear-momentum fractions. The power of  $D$  indicates the number of boost-invariant transverse derivatives (defined in (3.39)) in the term, which acts to TMD distributions in various combinations. The symbol  $\times$  indicates an integral convolution in the collinear-momentum fractions. Each term is equipped with a coefficient function which is not shown. The first line in the brackets is the LP term (1.1). The second line is the NLP term (see [19, 22]), and so on. Each term in expression (1.3) represents a complicated composition of various distributions.

In expression (1.3), one can easily recognize power corrections of different types. So, the first term in the third line ( $D\Phi_2 \times \Phi_2$ ) is the correction  $\sim k_T/Q$ , and the second term ( $\Phi_2 \times \Phi_3$ ) is the correction  $\Lambda/Q$ . In the fourth line, the first term ( $D^2\Phi_2 \times \Phi_2$ ) is  $k_T^2/Q^2$ , the second term ( $D\Phi_2 \times \Phi_3$ ) is  $k_T\Lambda/Q^2$ , the third and fourth terms ( $\Phi_3 \times \Phi_3$  and  $\Phi_2 \times \Phi_4$ ) are  $\Lambda^2/Q^2$ , and the last term ( $\Phi_2 \times \Phi_2/b^2$ ) is  $q_T^2/Q^2$ . And so on.

In eq. (1.3), the terms with different order of  $k_T/Q$  corrections but the same order of other corrections are aligned into columns. These columns represent a series of KPCs to the first term. Each column incorporates unique nonperturbative content, as it contains a combination of TMD distributions that do not appear in other columns. Hence, each column can be considered as an independent contribution to the TMD factorization theorem. All global properties expected from the hadronic tensor, such as charge conservation and frame invariance, must hold for each column individually.

It is important to emphasize that the decomposition (1.3) is based on the assumption that TMD distributions of different twists (i.e.,  $\Phi_n$  and  $\Phi_m$  for  $n \neq m$ ) are completely independent functions. This assumption holds if these distributions do not mix under TMD evolution. The requirement for non-mixture is a fundamental prerequisite because otherwise, power corrections of different types would entangle due to evolution or by higher perturbative orders of coefficient functions, and the presentation form (1.3) would be impossible. Moreover, it would violate the universality property of nonperturbative distributions.<sup>1</sup> Therefore, the

<sup>1</sup>In principle, one could allow for the triangular mixture, such that higher-twist distributions depend on lower-twist. The presence of a mixture is inconvenient and makes the analysis of the factorization theorem difficult. In this case, one can always eliminate the mixture by a proper change of the basis.

problem of systematization of power corrections in TMD factorization is equivalent to the problem of the definition of TMD-twist. This work adopts the definition of TMD twist introduced in ref. [22], which constitutes a TMD generalization of the concept of *geometrical twist* (aka “dimension minus spin”) used for collinear operators [27, 28]. This definition ensures, at least, the independence of non-mixture under ultraviolet renormalization.

Apart from the theoretical challenges posed by power-suppressed terms, there is also a practical concern regarding their relevance. It is evident that a plethora of TMD distributions appear in the power corrections, and the number of new functions grows faster than the number of observables. Already at NLP, one encounters 32 twist-three TMDPDFs [26] (in addition to the 8 TMDPDFs at the leading power). Identifying and constraining all these functions from experimental data alone is practically impossible. In this regard, higher twist power corrections can be characterized as “bad” as they introduce unknown functions (although offering insights into new facets of quantum physics). The other types of power corrections are “better” as they do not increase the number of unknowns but instead refine and extend the LP term.

Thus, it is sensible to incorporate power corrections by types rather than by orders. First, this approach preserves the validity of the factorization theorem, as different types of power corrections are additive and do not interfere with each other. Secondly, each type of correction contributes to a specific kinematic region and could be distinguished phenomenologically. Neglecting power corrections put the requirements to the applicability range of the factorization theorem. Each power correction leads to its own limitation, and therefore, the LP TMD factorization is valid in the regime

$$k_T \ll Q, \quad q_T \ll Q, \quad \Lambda \ll Q, \quad m \ll Q. \tag{1.4}$$

Including power corrections of a specific type softens the corresponding restriction, while summing the power corrections of a specific type eliminates the restriction (or replaces it with a strict inequality). At low  $Q$ , all types of power corrections are significant simultaneously. Notably, the  $q_T/Q$  corrections are special, as their magnitude is controlled by the experimental kinematics. By including only  $q_T/Q$  corrections, the cross-section can be described up to  $q_T \sim Q$ , while the other corrections remain negligible (for large  $Q$ ). The remaining corrections follow a hierarchy of  $k_T > m > \Lambda$  and are valid in the regime  $q_T \ll Q$ .

In this work, I study KPCs for the LP term. Within the expression (1.3), these power corrections are represented by the first column. Notably, these power corrections are of utmost importance within the conventional range of application for TMD factorization  $q_T \ll Q$ . Primarily, KPCs play a crucial role in restoring the electromagnetic (EM) gauge invariance (charge conservation) and frame invariance of the hadronic tensor, both of which are explicitly broken in the LP term. Consequently, KPCs can be viewed as an integral part of the LP factorization, as the absence of these corrections render the LP factorization inconsistent. Furthermore, unlike other scales such as  $q_T$ ,  $\Lambda$ , or  $m$ , the scale  $k_T$  is not fixed but rather an integration variable (1.2). This introduces a self-contradiction, as the factorization formula includes values of  $k_T$  that exceed  $Q$ , contradicting the initial assumption that  $Q$  is the largest scale. Moreover, the average value of  $k_T$  is of the order of the ultraviolet cut-off due to the large- $k_T$  asymptotic behavior of TMD distributions

( $\sim k_T^{-2}$ ). An extra point in favor of KPCs is that the factorization theorem for them does not require a dedicated proof. KPCs inherit the factorization property from the LP term, for which the theorem is already proven. Considering these reasons, it is sufficient to include KPCs independently of other power corrections.

The paper is structured as follows. In section 2 I review the concept of the TMD-twist and present in detail the simplest (but the most useful) example of twist-decomposition for the quark operator (section 2.2) and for a general bi-quark TMD distribution (section 2.3). section 3 is dedicated to the derivation of KPCs order-by-order in power expansion. Here, I start from the sketch of the operator derivation of the LP term (section 3.1) and continue to this derivation to all-powers (section 3.2). In section 3, I inspect the influence of loop corrections for KPCs and derive the argument of the coefficient function. The summation of the KPC series is done in section 4. To perform the summation, I first sum the series of twist-two terms for a general TMD correlator (section 4.1) and then insert it into the factorized expression (section 4.2). In section 5, I inspect the properties of the summed expression, using the example of a completely unpolarized contribution to hadron tensor of Drell-Yan. The subsections 5.1 and 5.2 are devoted to the EM gauge invariance and the frame invariance, correspondingly. Finally, the estimation of numerical impact is presented in section 5.3.

## 2 Twist decomposition in TMD factorization

The definition of TMD-twist is vital for systematisation of power corrections. Given this definition, it becomes possible to decompose operators of any dimension into irreducible components and express their matrix elements in terms of universal and independent nonperturbative TMD distributions. A suitable definition of TMD-twist was proposed in reference [22]. TMD-twist is determined by a pair of geometrical twists associated with the semi-compact operators that constitute any TMD operator. This approach strictly ensures non-mixture of operators with distinct TMD-twists under evolution with the ultraviolet (UV) renormalization scale  $\mu$ . Although the non-mixture under the rapidity evolution remains questionable, this property has been demonstrated in [22] for a wide range of TMD distributions (given by quasi-partonic operators), which encompasses all TMD distributions of twist-two and twist-three. Furthermore, the equivalence of rapidity evolution between TMD distributions of TMD-twist two and three was established in [23]. Therefore, this definition is satisfactory for the current objective, namely the determination of KPCs for the LP term, as they solely include TMD distributions of twist-two.

In this section, I provide a review of the general definition of TMD-twist and present the simplest yet non-trivial example of twist-decomposition. This example serves as the foundation for the subsequent computation of KPCs.

### 2.1 TMD-twist

At any power of TMD factorization theorem TMD distributions have the same general structure,

$$\tilde{\Phi}_{AB}(\{z\}_A, \{z\}_B; b) \sim \langle p, s | U_A(\{z\}_A; b) U_B(\{z\}_B; 0) | p, s \rangle, \quad (2.1)$$

where  $U$  is a light-cone operator, labels  $A$  and  $B$  indicate the quantum numbers of  $U$ , and  $\{z\}_A$  is a set of coordinates  $z$ . The operator  $U(\{z\}_A; b)$  is a T-ordered product of QCD fields positioned at  $(z_i n + b)$  for  $z_i \in \{z\}_A$ . The fields in the operator are connected by light-like Wilson lines, which continue to the light-cone infinity. So, the operator  $U$  spans an infinite range, and for that reason it is called *semi-compact*.

The renormalization of TMD distribution consists of two renormalization factors — one for each operator  $U$ . These factors are independent since the distance between operators is transverse and thus UV finite. The evolution equation reads

$$\mu^2 \frac{d}{d\mu^2} \tilde{\Phi}_{AB}(\{z\}_A, \{z\}_B; b) = (\tilde{\gamma}_A(\{z\}_A) + \tilde{\gamma}_B(\{z\}_B)) \otimes \tilde{\Phi}_{AB}(\{z\}_A, \{z\}_B; b), \quad (2.2)$$

where  $\tilde{\gamma}_A$  is the anomalous dimension of  $U_A$  and  $\otimes$  is the integral convolution in corresponding positions  $\{z_i\}$ . Note, that anomalous dimensions  $\tilde{\gamma}$  contain the double logarithmic part, which is resulted from gluon propagated along Wilson line to the infinity (collinear singularity). The same singularity is presented in the rapidity renormalization factor (or soft factor). Their cancellation leads to a logarithm contribution  $\sim \ln(\mu^2/\zeta)$  in the anomalous dimensions.

The equation (2.2) describes only the UV evolution, while TMD distributions also obey rapidity evolution equation (2.26). Currently, there is no general consensus regarding the rapidity divergences of operators that arise at higher powers. The only established result is that all quasi-partonic TMD operators conform to the leading power (LP) rapidity evolution [22]. Thus, it is plausible that the present definition of TMD-twist is incomplete, and some operators could potentially mix due to the rapidity evolution. However, this hypothetical situation would only apply to non-quasi-partonic TMD distributions, which have a minimum twist-four and emerge at NNLP for the first time, and thus, are irrelevant for the present discussion.

The light-cone operators  $U_A$  and  $U_B$  have independent anomalous dimensions if they have different geometrical twist (defined as “dimension-minus-spin” of the operator). Consequently, TMD distributions  $\tilde{\Phi}_{AB}$  and  $\tilde{\Phi}_{A'B'}$  do not mix with each other if the geometrical twist of  $U_A$  or  $U_B$  is different from the geometrical twist of  $U_{A'}$  or  $U_{B'}$ . It gives raise to a natural definition of TMD-twist for  $\tilde{\Phi}_{AB}$ , as a pair of integer numbers  $(N, M)$ , where  $N(M)$  is a geometrical twist of semi-compact operator  $U_{A(B)}$ .

Formally, the definition of geometrical twist is defined only for local operators. Its generalization for semi-compact operator can be done by the following procedure. First, one selects  $L$  such that  $\{|z|\} < |L|$ , and drop the part of Wilson line from  $Ln$  to infinity. Next, the operator is expanded as series at  $L$ . It takes a form  $U_A(z, 0) \sim \sum_n (z - L)^n U_{A,n}$ , where  $U_{A,n}$  are local operators. The geometric twist of local operators  $U_{A,n}$  is the geometric twist of semi-compact operator  $U_A$ . After the decomposition of operator into components with the same twist, each component can be summed over  $n$ . Finally, the limit  $L \rightarrow \infty$  is taken. This method properly reconstructs the properties of semi-compact operators, as it is demonstrated in ref. [29].

An indication of twist for TMD operator by a single number (used so far), refers to  $(N + M)$ . For example, it was used in the expression (1.3), where  $\Phi_{\mathbf{n}} = \Phi_{MN}$  with

$N + M = n$ . Meanwhile, all TMD distributions of twist-two have TMD-twist  $(1, 1)$ . However, the single-number terminology is ambiguous beyond the twist-two case. For example, a TMD distribution of twist-three can have TMD-twist  $(1, 2)$  or  $(2, 1)$ , which are two entirely independent distributions with separate evolution equations (see ref. [26]). TMD distribution of twist-four can have TMD-twist  $(1, 3)$ ,  $(3, 1)$  or  $(2, 2)$ , and so on. Still, the single-number indication is shorter and it is convenient to use if it does not create a confusion.

## 2.2 Example of twist-decomposition

As an example of twist decomposition let me consider the simplest semi-compact operator that appears in the TMD factorization. It is a quark field with an attached semi-infinite Wilson line

$$U_q(0, 0_T) = [-\infty n, 0]q(0), \tag{2.3}$$

where  $[an, bn]$  is a Wilson line along  $n$

$$[an, bn] = P \exp \left( -ig \int_a^b ds A_+(sn) \right). \tag{2.4}$$

As usually, the vectors  $n^\mu$  and  $\bar{n}^\mu$  are two independent light-cone vectors with normalization

$$n^2 = \bar{n}^2 = 0, \quad (n\bar{n}) = 1. \tag{2.5}$$

The decomposition of any vector  $v^\mu$  reads

$$v^\mu = \bar{n}^\mu v^+ + n^\mu v^- + v_T^\mu, \tag{2.6}$$

where  $v^+ = (nv)$ ,  $v^- = (\bar{n}v)$  and  $v_T$  is the transverse component orthogonal to  $(n, \bar{n})$ -plane  $v_T^\mu = g_T^{\mu\nu} v_\nu$  with

$$g_T^{\mu\nu} = g^{\mu\nu} - n^\mu \bar{n}^\nu - \bar{n}^\mu n^\nu. \tag{2.7}$$

Following the procedure described above the operator (2.3) can be presented as

$$U_q(0, 0_T) = \lim_{L \rightarrow -\infty} \sum_{n=0}^{\infty} \frac{i^n L^n}{n!} D_+^n q(Ln), \tag{2.8}$$

where  $D_\mu = \partial_\mu - igA_\mu$  is the QCD covariant derivative. The local operator  $D_{\mu_1} \dots D_{\mu_n} q_i$  has the mass-dimension  $n + 3/2$ . It is a Lorenz tensor of mixed nature with  $n$  vector indices and one spinor index. The maximum possible spin of this tensor is  $n + 1/2$ . To achieve it, one should symmetrize indices  $\mu$ , subtract traces (in the present case, it automatically achieved by contraction with  $n^{\mu_1} \dots n^{\mu_n}$ ), and make the spinor index  $\gamma$ -traceless [30], i.e. such that  $\gamma^{\mu_j} D_{\mu_1} \dots D_{\mu_j} \dots D_{\mu_n} q = 0$  (for a more formal discussion see refs. [31, 32]). For the present tensor (that is contracted with  $n^{\mu_1} \dots n^{\mu_n}$ ) it implies that the maximum-spin operator should vanish after multiplication by  $\gamma^-$ . It leads to a natural decomposition of the spinor into “good” and “bad” components [33], as  $q = \xi_{\bar{n}} + \eta_{\bar{n}}$ , where

$$\xi_{\bar{n}} = \frac{\gamma^- \gamma^+}{2} q, \quad \eta_{\bar{n}} = \frac{\gamma^+ \gamma^-}{2} q. \tag{2.9}$$

Here, the subscript  $\bar{n}$  indicates that these components are defined with respect to  $\gamma^-$ . Since  $\gamma^- \xi = 0$ , the operator  $D_+^n \xi$  has the maximum spin, and its geometrical twist equals one. Summing the series over  $n$  and limiting  $L \rightarrow -\infty$ , the twist-one part of (2.3) is

$$U_q(0, 0_T) \Big|_{\text{tw-1}} = U_1(0, 0_T) = [-\infty n, 0] \xi_{\bar{n}}(0). \quad (2.10)$$

The remaining part of  $U_q$  is  $[-\infty n, 0] \eta_{\bar{n}}(0)$ . It is twist-two, but the field  $\eta$  is not dynamically independent. The components  $\eta$  and  $\xi$  are related to each other via the QCD equation of motion (EOM). The EOM is  $\not{D}q = 0$  for massless quark. After projection (2.9) EOM splits into two equations, which are convenient to write as

$$D_+ \eta_{\bar{n}} = -\frac{1}{2} \gamma^+ \not{D}_T \xi_{\bar{n}}, \quad D_- \xi_{\bar{n}} = -\frac{1}{2} \gamma^- \not{D}_T \eta_{\bar{n}}. \quad (2.11)$$

Inserting the left equation into eq. (2.8) and commuting the transverse derivative to the outer position one gets

$$\begin{aligned} D_+^n \eta_{\bar{n}}(Ln) &= \frac{-1}{2} \gamma^+ \gamma_T^\mu D_+^{n-1} D_\mu \xi_{\bar{n}}(Ln) \\ &= \frac{-1}{2} \gamma^+ \gamma_T^\mu \left( D_\mu D_+^{n-1} \xi_{\bar{n}}(Ln) + ig \sum_{m=0}^{n-1} D_+^{n-m} F_{\mu+} D_+^m \xi_{\bar{n}}(Ln) \right), \end{aligned} \quad (2.12)$$

where  $F_{\mu\nu} = ig^{-1} [D_\mu, D_\nu]$  is the gluon field-strength tensor. The subscript on a gamma-matrix indicates that its index is transverse. The summing over  $n$  gives

$$[Ln, 0] \eta_{\bar{n}}(0) = \frac{-1}{2} \gamma^+ \gamma_T^\mu \int_{-\infty}^0 ds \left( D_\mu [Ln, sn] \xi_{\bar{n}}(sn) + ig \int_L^s ds_1 [Ln, s_1 n] F_{\mu+} [s_1 n, sn] \xi_{\bar{n}}(sn) \right). \quad (2.13)$$

The spin of the last term is  $n - 1/2$ , since it is traceless and  $\gamma$ -traceless, and has all, except one, indices symmetrized. Thus, the last term is twist-two operator.

The final expression for twist-decomposition of  $U_q$  can be assembled. Let the transverse component of the gluon field vanish at light-cone infinity,<sup>2</sup> so  $\lim_{L \rightarrow \infty} D_\mu(L) = \partial_\mu$ . Then combining (2.10) and (2.13), one receives

$$U_q(0, 0_T) = U_1(0, 0_T) - \frac{\gamma^+ \gamma_T^\mu}{2} \int_{-\infty}^0 ds \partial_\mu U_1(s, 0_T) - i \frac{\gamma^+ \gamma_T^\mu}{2} \int_{-\infty}^0 ds \int_{-\infty}^s ds_1 U_{2,\mu}(s_1, s, 0_T), \quad (2.14)$$

where  $U_1$  is the twist-one operator (2.10) and  $U_2$  is the twist-two operator

$$U_{2,\mu}(z_1, z_2, 0_T) = g[-\infty n, z_1 n] F_{\mu+} [z_1 n, z_2 n] \xi_{\bar{n}}(z_2 n). \quad (2.15)$$

Therefore, the operator  $U_q$  is decomposed into operators of twist-one (first term) and twist-two (last term), and the *total derivative* of twist-one operator (second term). The procedure presented above is standard for analysis of higher-twist collinear operators, see

<sup>2</sup>This assumption is valid in any regular gauge, and could be used as the boundary condition of singular gauges (such as the light-cone gauge). It is known that in the absence of such condition one should attach the link along transverse direction  $[-\infty n + \infty_T, -\infty n]$  [34]. However, such link could not be incorporated into the definition of twist in any form. Therefore, all presented consideration is done assuming that component of the gluon field is vanishing at appropriate light-cone infinity. Since the final result is gauge invariant the necessary transverse links can be inserted into the formulas after all derivatives are resolved.

e.g. examples in refs. [35–37]. The only difference in the present case is the open spinor index, which should be extra treated in the tensor decomposition.

Note that the “good” and “bad” components has also different counting with respect to  $\lambda \sim k_T/k_+$ . Assuming that a quark field has large  $k^+$  and almost massless  $k^2 \sim 0$ , one finds that  $\eta_{\bar{n}} \sim \lambda \xi_{\bar{n}}$  from EOMs (2.11). This counting is convenient to use for sorting operators in the power series, and it is widely used.

### 2.3 Twist-decomposition for bi-quark TMD distribution

The procedure of twist-decomposition for semi-compact operators leads to the decomposition of TMD matrix elements to definite twist parts. Let me demonstrate it for the bi-quark TMD matrix element, which is also important for derivation of KPCs. The matrix element reads

$$\begin{aligned}\tilde{\Phi}_{\bar{q}q}^{[\gamma^\mu]}(x, b) &= \int_{-\infty}^{\infty} \frac{dz}{2\pi} e^{-ixz p^+} \langle p | \bar{U}_q(zn, b) \frac{\gamma^\mu}{2} U_q(0, 0_T) | p \rangle, \\ &= \int_{-\infty}^{\infty} \frac{dz}{2\pi} e^{-ixz p^+} \langle p | \bar{q}(zn + b) [zn + b, -\infty n + b] \frac{\gamma^\mu}{2} [-\infty n, 0] q(0) | p \rangle,\end{aligned}\quad (2.16)$$

where  $|p\rangle$  is a hadron state with momentum  $p$ ,  $b$  is a transverse vector and  $\bar{U} = U^\dagger \gamma^0$ . Being contracted with  $n^\mu$  this matrix element results into the unpolarized and Sivers TMD distributions that are twist-two TMD distribution. For a general  $\gamma^\mu$ , the matrix element  $\tilde{\Phi}_{\bar{q}q}^{[\gamma^\mu]}$  does not represent a conventional distribution, but is a sum of distributions with different properties. It can be checked by computing the renormalization of matrix element (2.16) and observing that it could not be expressed via  $\tilde{\Phi}^{[\gamma^\mu]}$  (see f.i. computations in refs. [26, 38]). The reason is that the operator (2.16) has an undefinite twist.

Inserting the decomposition (2.14) into (2.16), the correlator  $\tilde{\Phi}_{\bar{q}q}^{[\gamma^\mu]}$  turns into

$$\begin{aligned}\tilde{\Phi}_{\bar{q}q}^{[\gamma^\mu]}(x, b) &= \\ &\int_{-\infty}^{\infty} \frac{dz}{2\pi} e^{-ixz p^+} \left[ \langle p | \bar{U}_1(z, b) \frac{\gamma^\mu}{2} U_1(0, 0_T) | p \rangle \right. \\ &\quad - \frac{1}{2} \int_{-\infty}^0 ds \langle p | \bar{U}_1(z, b) \frac{\gamma^\mu \gamma^+ \gamma_T^\nu}{2} \overrightarrow{\partial}_\nu U_1(s, 0_T) + \bar{U}_1(z+s, b) \overleftarrow{\partial}_\nu \frac{\gamma_T^\nu \gamma^+ \gamma^\mu}{2} U_1(0, 0_T) | p \rangle \\ &\quad + \frac{1}{4} \int_{-\infty}^0 ds dt \langle p | \bar{U}_1(z+s, b) \overleftarrow{\partial}_\nu \frac{\gamma_T^\nu \gamma^+ \gamma^\mu \gamma^+ \gamma_T^\rho}{2} \overrightarrow{\partial}_\rho U_1(t, 0_T) | p \rangle \\ &\quad + \frac{i}{2} \int_{-\infty}^0 ds \int_{-\infty}^s ds_1 \\ &\quad \left. \langle p | \bar{U}_{2,\nu}(z+s, z+s_1, b) \frac{\gamma_T^\nu \gamma^+ \gamma^\mu}{2} U_1(0, 0_T) - \bar{U}_1(z, b) \frac{\gamma^\mu \gamma^+ \gamma_T^\nu}{2} U_{2,\nu}(s_1, s, 0_T) | p \rangle \right. \\ &\quad - \frac{i}{4} \int_{-\infty}^0 dt \int_{-\infty}^0 ds \int_{-\infty}^s ds_1 \langle p | \bar{U}_{2,\nu}(z+s, z+s_1, b) \frac{\gamma_T^\nu \gamma^+ \gamma^\mu \gamma^+ \gamma_T^\rho}{2} \partial_\rho U_1(t, 0_T) \\ &\quad \quad \quad \left. - \bar{U}_1(z+t, b) \overleftarrow{\partial}_\rho \frac{\gamma_T^\rho \gamma^+ \gamma^\mu \gamma^+ \gamma_T^\nu}{2} U_{2,\nu}(s_1, s, 0_T) | p \rangle \right. \\ &\quad \left. + \frac{1}{4} \int_{-\infty}^0 dt \int_{-\infty}^t dt_1 \int_{-\infty}^s ds \int_{-\infty}^s ds_1 \langle p | \bar{U}_{2,\nu}(z+s, z+s_1, b) \frac{\gamma_T^\nu \gamma^+ \gamma^\mu \gamma^+ \gamma_T^\rho}{2} U_{2,\rho}(t, t_1, 0_T) | p \rangle \right],\end{aligned}\quad (2.17)$$

where  $\bar{U} = U^\dagger \gamma^0$ . The matrix element in the first three lines are the TMD distributions of TMD-twist-(1,1). The lines fifth to seventh are TMD distributions with TMD-twists-(2,1) and (1,2). The last line is the TMD distribution of TMD-twist-(2,2).

The gamma-structure of (2.17) can be simplified using that the spinor indices of operators  $U$  projected by matrices (2.9). For example, one finds that  $\bar{U}_1 \gamma^\mu U_1 = \bar{n}^\mu \bar{U}_1 \gamma^+ U_1$ . The total derivatives of operators  $U$  can be pushed outside of matrix elements using that

$$\langle p | \partial_\mu (\dots) | p \rangle = 0. \tag{2.18}$$

After these simplifications the result reads

$$\tilde{\Phi}_{\bar{q}q}^{[\gamma^\mu]}(x, b) = \left( \bar{n}^\mu + \frac{i}{xp^+} \frac{\partial}{\partial b_\mu} + \frac{n^\mu}{2(xp^+)^2} \frac{\partial^2}{\partial b^\nu \partial b_\nu} \right) \tilde{\Phi}_{11}^{[\gamma^+]}(x, b) + \dots, \tag{2.19}$$

where  $\Phi_{11}^{[\gamma^+]}$  is a TMD distribution of TMD-twist-(1,1)

$$\tilde{\Phi}_{11}^{[\Gamma]}(x, b) = \int_{-\infty}^{\infty} \frac{dz}{2\pi} e^{-ixzp^+} \langle p | \bar{U}_1(z, b) \frac{\Gamma}{2} U_1(0, 0_T) | p \rangle, \tag{2.20}$$

and dots denote contributions with TMD-twist-(1,2), (2,1) and (2,2) resulted from the last five lines of (2.17). The expression (2.19) is the twist-two part of the correlator (2.16).

The momentum space representation of TMD distribution is obtained by

$$\tilde{\Phi}^{[\Gamma]}(x, b) = \int d^2 k_T e^{i(bk_T)} \Phi^{[\Gamma]}(x, k_T). \tag{2.21}$$

Here and in the following, I use the (un)tilded notation to indicate the distributions in the (momentum) position space. The momentum space representation of (2.19) is

$$\Phi_{\bar{q}q}^{[\gamma^\mu]}(x, k_T) = \left( \bar{n}^\mu + \frac{k_T^\mu}{xp^+} - \frac{n^\mu k_T^2}{2(xp^+)^2} \right) \Phi_{11}^{[\gamma^+]}(x, k_T) + \dots. \tag{2.22}$$

The derivatives over  $b$  turn into powers of  $k_T$ .

It is also critical to observe that total derivatives of semi-compact operators does not always result to the definite-twist TMD distributions. It works out for total derivatives  $\partial_+$  (which produces factors of  $x$  in the momentum space), and  $\partial_{\mu T}$  (which produces factors of  $k_T^\mu$ ). However, total derivatives  $\partial_-$  cannot be moved outside of the operator. Nonetheless, they could be always expressed via  $\partial_+$ ,  $\partial_{\mu T}$  and extra gluon fields, thanks to EOMs (2.11). The example is given in section 3.2.

The expression (2.22) can be written as

$$\Phi_{\bar{q}q}^{[\gamma^\mu]}(x, k_T) = 2 \int_{-\infty}^{\infty} dk^- k^\mu \delta(k^2) \Phi_{11}^{[\gamma^+]}(x, k_T) + \dots, \tag{2.23}$$

with  $k^\mu = xp_+ \bar{n}^\mu + k_T^\mu + n^\mu k^-$  being the four-momentum of the parton. This representation gives rise to an interpretation of the leading twist distributions, as the part of the hadron's structure carried by the free-quark approximation. In the section 4.1 I provide a further generalization of eq. (2.23).

The twist-decomposition is a standard procedure for consideration of power corrections for factorization theorems, and the expressions (2.14) and (2.19) are rather typical for such kind of calculus. Many analogies can be found in the literature related to the collinear factorization. For example, the expressions (2.19) and (2.22) are ideologically analogous to the Wandzura-Wilczek approximation for collinear distributions [39]. In the present work, the main interest is the terms with total derivatives of the leading twist operator. These terms are responsible for generation of KPCs. It is analogous to the case of Deeply-Virtual Compton Scattering (DVCS) and Generalized parton distributions (GPDs) [40, 41].

Note that for the operator with a differently oriented light-like Wilson line, the vectors  $n$  and  $\bar{n}$  should be swapped. For example, for decomposition for the operator  $[-\infty\bar{n}, 0]q(0)$  is the same as (2.14) but with  $n \leftrightarrow \bar{n}$ . The “good” and “bad” components are defined as

$$\xi_n = \frac{\gamma^+\gamma^-}{2}q, \quad \eta_n = \frac{\gamma^-\gamma^+}{2}q. \tag{2.24}$$

The direction of Wilson line (future vs. past) does not play a role in this example.

For the following discussion, it is important to specify the evolution equation for twist-two TMDPDFs. these equations are [12, 13]

$$\mu^2 \frac{d}{d\mu^2} \tilde{\Phi}_{11}^{[\Gamma]}(x, b; \mu, \zeta) = \left( \frac{\Gamma_{\text{cusp}}(\mu)}{2} \ln \left( \frac{\mu^2}{\zeta} \right) - \frac{\gamma_V(\mu)}{2} \right) \tilde{\Phi}_{11}^{[\Gamma]}(x, b; \mu, \zeta), \tag{2.25}$$

$$\zeta \frac{d}{d\zeta} \tilde{\Phi}_{11}^{[\Gamma]}(x, b; \mu, \zeta) = -\mathcal{D}(b, \mu) \tilde{\Phi}_{11}^{[\Gamma]}(x, b; \mu, \zeta), \tag{2.26}$$

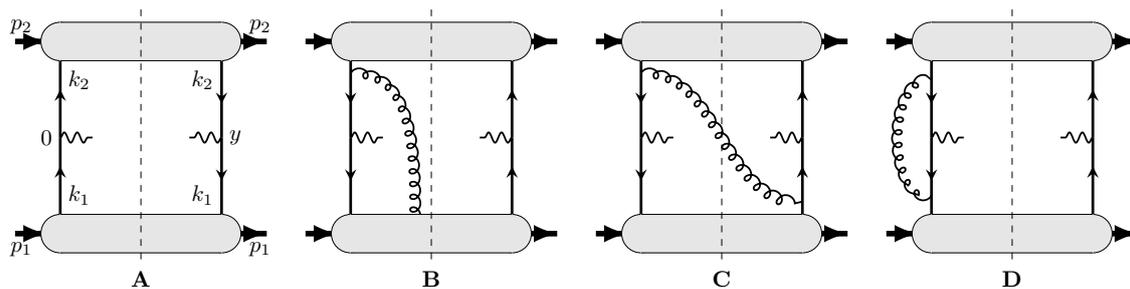
where  $\Gamma_{\text{cusp}}$ ,  $\gamma_V$  and  $\mathcal{D}$  are the cusp, quark-vector and rapidity anomalous dimensions, correspondingly. The rapidity anomalous dimension is also known as the Collins-Soper kernel [42]. It is a nonperturbative function that is associated with the properties of QCD vacuum [43]. In the momentum space the equation (2.26) turns to an integral equation [44].

### 3 KPCs for LP term

TMD operators of twist-two have a crucial property of incorporating only *two* “good” components of fields (quark and anti-quark, or two gluon fields). This simplifies the procedure of their extraction from higher-dimensional operators and allows for the unambiguous elimination of higher twist contributions at an early stage of computation. As a result, the determination of KPCs for the LP term can be achieved at all orders of the power series, as presented in this section. For concreteness, I focus on the Drell-Yan process

$$h_1(p_1) + h_2(p_2) \rightarrow \gamma^*(q) + X, \tag{3.1}$$

where the arguments denote the momenta of the respective particles. The generalization for other processes, such Semi-Inclusive Deep-Inelastic Scattering (SIDIS) or  $e^+e^- \rightarrow h_1h_2 + X$ , is straightforward. I start with the schematic review of the derivation of LP term, and then generalize this example including KPCs.



**Figure 1.** Example of diagrams contributing to the DY process. Gray blobs show the hadron nonperturbative state. The vertical dashed line denote the insertion of complete set of states.

### 3.1 TMD factorization at LP

The derivation of the LP term of the TMD factorization theorem has been extensively discussed in numerous references employing various formulations. Here are only some of them [9–11, 15–17, 22, 45]. In this section, I provide an overview of the derivation from the perspective of operator manipulations, similar to the TMD operator expansion, Soft-Collinear effective theories or High-Energy expansion. Emphasis is done on the technical aspects that are relevant for further explanations. I omit detailed proofs that can be found in the literature.

The QCD part of the cross-section for a hard reactions is given by the hadronic tensor. For the Drell-Yan process (3.1) the hadronic tensor reads

$$W^{\mu\nu} = \int \frac{d^4y}{(2\pi)^4} e^{-i(yq)} \sum_X \langle p_1, p_2 | J^\mu(y) | X \rangle \langle X | J^\nu(0) | p_1, p_2 \rangle, \quad (3.2)$$

where  $J^\mu$  is the electromagnetic (EM) current

$$J^\mu(y) = \bar{q}(y) \gamma^\mu q(y), \quad (3.3)$$

with  $q$  being the quark field. The quark charges and flavor part of the current are omitted for brevity and restored in the following sections.

The kinematics of the process is defined by the photon momentum  $q^\mu$  and the hadrons momenta  $p_1^\mu$  and  $p_2^\mu$ . To avoid complications related to the target mass corrections, the hadrons are considered as massless  $p_1^2 = p_2^2 = 0$ . These momenta introduce the natural system of vectors  $n$  and  $\bar{n}$  (2.5)

$$p_1^\mu = \bar{n}^\mu p_1^+, \quad p_2^\mu = n^\mu p_2^-. \quad (3.4)$$

The leading perturbative order contribution to the hadronic tensor of Drell-Yan process (3.2) is shown in figure 1(A). This diagram and its charge-conjugated are convenient to write as

$$W^{\mu\nu} = \int \frac{d^4y}{(2\pi)^4} e^{-i(yq)} \langle p_1 | \langle p_2 | J_{\bar{n}\bar{n}}^\mu(y) J_{n\bar{n}}^\nu(0) + J_{\bar{n}\bar{n}}^\mu(y) J_{\bar{n}n}^\nu(0) | p_1 \rangle | p_2 \rangle + \dots, \quad (3.5)$$

where dots represent the sub-leading terms, and

$$J_{\bar{n}\bar{n}}^\mu = \bar{q}_{\bar{n}} \gamma^\mu q_{\bar{n}}, \quad J_{\bar{n}n}^\mu = \bar{q}_n \gamma^\mu q_{\bar{n}}. \quad (3.6)$$

The fields  $q_{\bar{n}}$  ( $q_n$ ) are created by  $|p_1\rangle$  ( $|p_2\rangle$ ) states, and, according to the parton model, have momentum almost along  $p_1$  ( $p_2$ ). Usually, fields  $q_{\bar{n}}$  and  $q_n$  are referred as collinear and anti-collinear fields. At LP approximation the states  $|p_1\rangle$  and  $|p_2\rangle$  do not interact with each other directly. The complete set of states in-between currents is omitted for simplicity.

To proceed further, one should specify the counting rules for the elements of equation (3.5). For the TMD factorization the counting rules are [10, 11, 22]

$$\partial_\mu q_{\bar{n}} \sim A_{\mu,\bar{n}} \sim \{1, \lambda^2, \lambda\}, \quad \partial_\mu q_n \sim A_{\mu,n} \sim \{\lambda^2, 1, \lambda\}, \quad (3.7)$$

where  $\lambda \sim Q^{-1}$ . The components are presented in the order  $\{+, -, T\}$ . In the regime  $q^\pm \sim Q$  and  $q_T \sim Q^0$  the vector  $y$  satisfies the counting

$$y^\mu \sim Q^{-1}\{1, 1, \lambda^{-1}\} \quad (3.8)$$

The inhomogeneity of counting for components of  $y^\mu$  is the only conceptual difference between collinear and TMD factorizations.

The counting rules imply that the interactions between fields along certain directions are power-suppressed. The field must be expanded around these directions<sup>3</sup>

$$q_{\bar{n}}(y) = q_{\bar{n}}(y^- n + y_T) + y^+ \partial_- q_{\bar{n}}(y^- n + y_T) + \dots, \quad (3.9)$$

where the first term is  $\sim \lambda^0$ , the second term is  $\sim \lambda^2$ , and the dots indicate the further suppressed terms. Importantly, there is no expansion in the transverse direction, because  $y_T^\mu \partial_\mu q_{\bar{n}} \sim \lambda^0$  according to (3.7) and (3.8). Meanwhile, the transverse derivatives that appear in other parts of computation (e.g. in the loops) and that are not accompanied by  $y_T$ , produce power suppressed terms.

In addition to the diagram A, one should take into account the diagrams that are of the same order of power counting. These are diagrams with radiation of collinear components of gluon, such as diagram B in figure 1. Accounting of all such diagrams restores the QCD gauge invariance and equips the collinear and anti-collinear quark fields by half-infinite Wilson lines [34, 45]. Practically, it is convenient to impose the light-cone gauge for the background field. The gauge can be fixed for each collinear sector such that gluon fields with a  $\sim \lambda^0$  power counting vanish, i.e.  $A_{\bar{n},+} = 0$  and  $A_{n,-} = 0$ . The detailed justification of such choice can be found e.g. in ref. [22]. For the present work, it is enough to know that this set of gauges provides the correct factorization.

Dropping the power-suppressed terms in eq. (3.9), one gets the hadronic tensor in the form

$$W^{\mu\nu} = \int \frac{d^4 y}{(2\pi)^4} e^{-i(yq)} \quad (3.10)$$

$$\times \langle p_1 | \langle p_2 | \bar{q}_{\bar{n}}(y^- n + y_T) \gamma^\mu q_n(y^+ \bar{n} + y_T) \bar{q}_n(0) \gamma^\nu q_{\bar{n}}(0) + \{\bar{n} \leftrightarrow n\} | p_1 \rangle | p_2 \rangle + \dots$$

---

<sup>3</sup>For definiteness, the expansion is made at the origin. However, any other point can be used. The result is independent on this choice, thanks to eq. (2.18). However, in the factorization theorems for off-forward processes, which involve generalized TMD (GTMD) distributions, such as [46, 47], the result depends on the reference point. The independence on reference point should be restored in the sum of all KPCs similarly to the DVCS case [48].

The central assumption of the parton model is that a high-energy hadron consists of fields with corresponding collinear counting rules. Therefore, we can sort the fields with different collinearity into separate matrix elements. To do so, one needs to recouple the spinor and color indices, which could be done using the completeness relation of Dirac matrices

$$A = \frac{1}{2} \sum_a \bar{\Gamma}_a A^{[\Gamma_a]}, \quad A^{[\Gamma]} = \frac{1}{2} \text{Tr}(A\Gamma), \quad (3.11)$$

for any matrix  $A$ . Here,  $\Gamma$  is a complete basis of Dirac matrices. The result reads

$$W^{\mu\nu} = -\frac{1}{4N_c} \int \frac{d^4y}{(2\pi)^4} e^{-i(yq)} \sum_{a,b} \left( \text{Tr} \left( \gamma^\mu \bar{\Gamma}_b \gamma^\nu \bar{\Gamma}_a \right) \tilde{\Psi}_n^{[\Gamma_a]}(y^- n + y_T) \tilde{\Psi}_n^{[\Gamma_b]}(-y^+ \bar{n} - y_T) \right. \\ \left. + \text{Tr} \left( \gamma^\mu \bar{\Gamma}_a \gamma^\nu \bar{\Gamma}_b \right) \tilde{\Psi}_n^{[\Gamma_a]}(-y^- n - y_T) \tilde{\Psi}_n^{[\Gamma_b]}(y^+ \bar{n} + y_T) \right) + \dots, \quad (3.12)$$

where the common minus sign appears from the anti-commutation of the quark fields, and the factor  $1/N_c$  is due to the recoupling of color indices. The matrix elements are

$$\tilde{\Psi}_n^{[\Gamma]}(y^- n + y_T) = \langle p_1 | \bar{q}_n(y^- n + y_T) [y^- n + y_T, -\infty n + y_T] \frac{\Gamma}{2} [-\infty n, 0] q_n(0) | p_1 \rangle, \quad (3.13)$$

$$\tilde{\Psi}_n^{[\Gamma]}(y^+ \bar{n} + y_T) = \langle p_2 | \bar{q}_n(y^+ \bar{n} + y_T) [y^+ \bar{n} + y_T, -\infty \bar{n} + y_T] \frac{\Gamma}{2} [-\infty \bar{n}, 0] q_n(0) | p_2 \rangle. \quad (3.14)$$

These distributions are Fourier images of distributions  $\tilde{\Phi}$  (2.16) over the variables  $x$ . In expressions (3.13), (3.14) the complete set of states in-between quark fields can be omitted because all distances are space-like.

The expression (3.12) is of indefinite power, because the TMD distribution  $\Psi^{[\gamma^\mu]}$  is a mixture of terms with different power counting. It follows from eq. (2.17), where the first line is  $\sim \lambda^0$ , the second and fourth lines are  $\sim \lambda^1$ , etc. Therefore, the pure LP term is obtained by eliminating all, except the first, terms in eq. (2.17). It can be done by selecting only those  $\Gamma$ -matrices that project only “good” components. For the collinear part these matrices are

$$\Gamma_a^+ = \{\gamma^+, \gamma^+ \gamma^5, i\sigma^{\alpha+} \gamma^5\}, \quad (3.15)$$

with  $\alpha$  being transverse index. The corresponding elements of the basis are

$$\bar{\Gamma}_a^- = \{\gamma^-, -\gamma^- \gamma^5, -\frac{i}{2} \sigma^{\alpha-} \gamma^5\}. \quad (3.16)$$

For the decomposition in the anti-collinear sector one should use  $\Gamma_a^-$  and  $\bar{\Gamma}_a^+$  that are obtained by the replacement  $n \leftrightarrow \bar{n}$  in (3.15) and (3.16), correspondingly. After restriction of  $\Gamma$ -matrices to the set (3.15) the hadronic tensor contains only the LP contribution.

Finally, one passes to the momentum-fraction representation for TMD distributions and integrates over  $y^\pm$ . The result is

$$W_{\text{LP}}^{\mu\nu} = -\frac{p_1^+ p_2^-}{4N_c} \int \frac{d^2b}{(2\pi)^2} e^{-i(bq_T)} \int dx d\tilde{x} \delta(xp_1^+ - q^+) \delta(\tilde{x}p_2^- - q^-) \sum_{a,b} \left( \right. \\ \left. \text{Tr} \left( \gamma^\mu \bar{\Gamma}_b^+ \gamma^\nu \bar{\Gamma}_a^- \right) \tilde{\Phi}_{n11}^{[\Gamma_a^+]}(x, b) \tilde{\Phi}_{n11}^{[\Gamma_b^-]}(-\tilde{x}, -b) + \text{Tr} \left( \gamma^\mu \bar{\Gamma}_a^- \gamma^\nu \bar{\Gamma}_b^+ \right) \tilde{\Phi}_{n11}^{[\Gamma_a^-]}(-x, -b) \tilde{\Phi}_{n11}^{[\Gamma_b^+]}(\tilde{x}, b) \right), \quad (3.17)$$

where  $y_T$  is renamed as  $b$ , to match the standard notation. The expression (3.17) is the celebrated expression for the LP TMD factorization. The TMD distributions with negative  $x$  are related to the anti-quark distributions.

This is a complete derivation of LP TMD factorization at the tree order. To go beyond LO, one should include interaction diagrams. Some examples are presented in figure 1. The diagrams of type B (with a non-collinear gluon) or C do not contribute to the LP factorization. Indeed, any additional interaction with the hadrons (such as figure 1B) increases the number of fields in the operator, and hence increases its dimension. The only exception is the radiation of collinear gluons, which is already accounted in the LP term. Therefore, the interactions that do not violate the power counting are possible only in-between the hard fields. Even so, some hard interactions are power suppressed, due to inhomogeneous counting rules for  $y$ . Indeed, the diagram 1C is  $\sim 1/b^2$  in the position space, which leads to  $\sim q_T^2/Q^2$  in the momentum space.<sup>4</sup> Therefore, the perturbative correction at LP are given by the interactions in the vicinity of currents, such as the diagram 1D. As the result, the coefficient function is the product of coefficient functions for each current. It is in the one-to-one correspondence with the TMD-twist idea. Each semi-compact operator produces independent UV renormalization factor, that cancels the infrared poles of a corresponding part of coefficient function. This structure of pole cancellations is preserved at all powers of TMD factorization.

Finally, one should also take into account that separation of collinear and anti-collinear modes is ambiguous. For the small values of momenta the collinear and anti-collinear sectors overlap. There are several methods to avoid it. The double-counting of modes can be subtracted by means of the soft factor [9, 49] or rapidity renormalization factor [50], or modes can be defined with explicit cut parameter that prevents the overlap [22, 51]. The results coincide.<sup>5</sup> This procedure leads to an extra scales  $\zeta$  and  $\bar{\zeta}$ , and corresponding evolution equations (2.26).

### 3.2 KPC for LP term at all orders

The extension of the TMD factorization beyond LP is conceptually straightforward (although very complicated technically). One should systematically expand the interaction vertices in the background fields, and sort operators over the twists and power counting. Detailed discussions are presented in refs. [19, 22]. Performing this procedure, one receives the structure (1.3). In the following, I concentrate on the KPCs to the LP term. They are much simpler than other types of power corrections, because they incorporate only the TMD distributions of twist-two. Let me consider the series of power corrections and extract only the twist-two component from it. This component produces the complete series of KPCs.

The key observation is that the TMD distributions of twist-two are given by only two-point operators. Indeed, an insertion of any extra field increases the dimension and, thus,

---

<sup>4</sup>Since these power corrections appears due to the expansion of the propagator, they produces only even powers of  $q_T/Q$ -corrections. That describes the structure of eq. (1.3).

<sup>5</sup>To reach a complete agreement, different schemes of the rapidity renormalization must satisfy the same renormalization condition. Traditionally the TMDPDFs are defined such that the Drell-Yan cross-section does not incorporate any extra nonperturbative factor [50].

the TMD-twist by at least one. The geometrical-twist-three operators are all three-point operators. Such a simple rule (twist  $\sim$  number of fields) does not hold for higher-twist operators starting from twist-four, which could be three-point or four-point operators [52]. Still, one can formulate a rule: any two-point operator can contain geometrical twist-two part, but any three-point and higher-number-of-points operators necessarily have geometrical twist higher than two. It implies that any diagram with three or more particles in a single collinear sector (such as diagram B) does not contribute to KPCs for LP term. Consequently, the diagram A contains full information about twist-two part at all powers at LO in perturbative expansion.

Inspecting the LP derivation in section 3.1, one finds that the power corrections were dropped in two places: the multipole decomposition (3.9) and selection of “good” components of fields. If one does not perform these approximation the result will contain full information on KPCs. Lets release these approximations.

Starting with eq. (3.5) and performing the multipole expansion one obtains

$$\begin{aligned}
 W^{\mu\nu} = & \int \frac{d^4y}{(2\pi)^4} e^{-i(yq)} \sum_{n,m=0}^{\infty} \frac{(y^+)^n (y^-)^m}{n!m!} \\
 & \times \langle p_1 | \langle p_2 | \bar{q}_{\bar{n}}(y^- n + y_T) \overleftarrow{\partial}_-^n \gamma^\mu \overrightarrow{\partial}_+^m q_n(y^+ \bar{n} + y_T) J_{n\bar{n}}^\nu(0) + \{\bar{n} \leftrightarrow n\} | p_1 \rangle | p_2 \rangle + \dots .
 \end{aligned} \tag{3.18}$$

Here, the dots represent contributions with larger number of fields. The LP term (3.10) is at  $n = m = 0$ . Expressing the  $\Gamma$ -structure in the standard basis gives

$$\begin{aligned}
 W^{\mu\nu} = & \frac{-1}{4N_c} \int \frac{d^4y}{(2\pi)^4} e^{-i(yq)} \sum_{a,b} \sum_{n,m=0}^{\infty} \frac{(y^+)^n (y^-)^m}{n!m!} \left[ \right. \\
 & \text{Tr}(\gamma^\mu \bar{\Gamma}_b \gamma^\nu \bar{\Gamma}_a) \langle p_1 | \bar{U}_q(y^-, y_T) \overleftarrow{\partial}_-^n \frac{\Gamma_a}{2} U_q(0, 0_T) | p_1 \rangle \langle p_2 | \bar{U}_q(-y^+, -y_T) \frac{\Gamma_b}{2} \overrightarrow{\partial}_+^m U_q(0, 0_T) | p_2 \rangle \\
 & \left. + \{\bar{n} \leftrightarrow n\} \right] + \dots ,
 \end{aligned} \tag{3.19}$$

where the I use the notation (2.3) to highlight the semi-compact operators. Note that operators in the matrix elements with  $p_1(p_2)$  are (anti-)collinear. In contrast to the LP term, there are no constraints on  $\Gamma$ -matrices, since any Dirac component could produce a twist-two contribution, see example in eq. (2.19).

The semi-compact operators in eq. (3.19) contain the mixture of TMD distribution, and should be decomposed over the definite twist contributions. For the pure  $U_q$  operators the decomposition is given in eq. (2.14). However, the “improper” derivatives of  $U_q$  (such as  $\partial_- q_{\bar{n}}$ ) demolish this decomposition because they could not be presented as total derivatives of distributions. This issue is resolved by application of EOMs (2.11). The analysis of these operators is complicated in the general case, but essentially simpler if one seeks for only twist-one contribution. In this case, all gluon fields in EOMs could be dropped (since gluon-fields necessarily turn one-point operator to a higher-number-of-points operators), resulting to a simple rule

$$\partial_- \xi_{\bar{n}} = -\frac{\partial_T^2}{2\partial_+} \xi_{\bar{n}} + \dots, \quad \partial_- \eta_{\bar{n}} = \frac{\gamma^+ \not{\partial}_T \partial_T^2}{4\partial_+^2} \xi_{\bar{n}} + \dots, \tag{3.20}$$

where dots represent the terms with two or larger number of fields. Here inverse derivatives are understood as the integral, alike (2.13). Basically, the equation (3.20) tells that at the leading-twist approximation partons are free massless fields. The higher powers of  $\partial_-$  are

$$\partial_-^n \xi_{\bar{n}} = \left( -\frac{\partial_T^2}{2\partial_+} \right)^n \xi_{\bar{n}} + \dots, \quad \partial_-^n \eta_{\bar{n}} = -\frac{\gamma^+ \not{\partial}_T}{2\partial_+} \left( -\frac{\partial_T^2}{2\partial_+} \right)^n \xi_{\bar{n}} + \dots \quad (3.21)$$

After these transformations, the fields  $\xi$  could be promoted to the semi-compact operators  $U_1$ , by equipping them by semi-infinite Wilson lines, as it is demanded by the QCD gauge invariance.

As a result of these manipulations, one obtains a long expression

$$\begin{aligned} W_{\text{KPC}}^{\mu\nu} = & \frac{-1}{4N_c} \int \frac{d^4 y}{(2\pi)^4} e^{-i(yq)} \sum_{a,b} \sum_{n,m=0}^{\infty} \frac{(y^+)^n (y^-)^m (-1)^{n+m}}{n!m!} \frac{1}{2^{n+m}} \left\{ \right. \\ & \text{Tr}[\gamma_T^\mu \bar{\Gamma}_b^+ \gamma_T^\nu \bar{\Gamma}_a^-] \frac{\partial_T^{2n}}{\partial_+^n} \tilde{\Psi}_{\bar{n}11}^{[\Gamma_a^+]} \frac{\partial_T^{2m}}{\partial_-^m} \tilde{\Psi}_{n11}^{[\Gamma_b^-]} \\ & - \left( n^\nu \text{Tr}[\gamma_T^\mu \bar{\Gamma}_b^+ \gamma_T^\alpha \bar{\Gamma}_a^-] + n^\mu \text{Tr}[\gamma_T^\alpha \bar{\Gamma}_b^+ \gamma_T^\nu \bar{\Gamma}_a^-] \right) \frac{\partial_\alpha}{\partial_+} \frac{\partial_T^{2n}}{\partial_+^n} \tilde{\Psi}_{\bar{n}11}^{[\Gamma_a^+]} \frac{\partial_T^{2m}}{\partial_-^m} \tilde{\Psi}_{n11}^{[\Gamma_b^-]} \\ & - \left( \bar{n}^\nu \text{Tr}[\gamma_T^\mu \bar{\Gamma}_b^+ \gamma_T^\alpha \bar{\Gamma}_a^-] + \bar{n}^\mu \text{Tr}[\gamma_T^\alpha \bar{\Gamma}_b^+ \gamma_T^\nu \bar{\Gamma}_a^-] \right) \frac{\partial_T^{2n}}{\partial_+^n} \tilde{\Psi}_{\bar{n}11}^{[\Gamma_a^+]} \frac{\partial_\alpha}{\partial_-} \frac{\partial_T^{2m}}{\partial_-^m} \tilde{\Psi}_{n11}^{[\Gamma_b^-]} \\ & - \frac{1}{2} \left( \text{Tr}[\gamma_T^\mu \bar{\Gamma}_b^+ \gamma_T^\beta \gamma_T^\nu \gamma_T^\alpha \bar{\Gamma}_a^-] + \text{Tr}[\gamma_T^\alpha \gamma_T^\mu \gamma_T^\beta \bar{\Gamma}_b^+ \gamma_T^\nu \bar{\Gamma}_a^-] \right) \frac{\partial_\alpha}{\partial_+} \frac{\partial_T^{2n}}{\partial_+^n} \tilde{\Psi}_{\bar{n}11}^{[\Gamma_a^+]} \frac{\partial_\beta}{\partial_-} \frac{\partial_T^{2m}}{\partial_-^m} \tilde{\Psi}_{n11}^{[\Gamma_b^-]} \\ & + \left( \bar{n}^\mu n^\nu \text{Tr}[\gamma_T^\beta \bar{\Gamma}_b^+ \gamma_T^\alpha \bar{\Gamma}_a^-] + n^\mu \bar{n}^\nu \text{Tr}[\gamma_T^\alpha \bar{\Gamma}_b^+ \gamma_T^\beta \bar{\Gamma}_a^-] \right) \frac{\partial_\alpha}{\partial_+} \frac{\partial_T^{2n}}{\partial_+^n} \tilde{\Psi}_{\bar{n}11}^{[\Gamma_a^+]} \frac{\partial_\beta}{\partial_-} \frac{\partial_T^{2m}}{\partial_-^m} \tilde{\Psi}_{n11}^{[\Gamma_b^-]} \\ & + \text{Tr}[\gamma_T^\alpha \bar{\Gamma}_b^+ \gamma_T^\beta \bar{\Gamma}_a^-] \left( \bar{n}^\mu \bar{n}^\nu \frac{\partial_T^{2n}}{\partial_+^n} \tilde{\Psi}_{\bar{n}11}^{[\Gamma_a^+]} \frac{\partial_\alpha}{\partial_-} \frac{\partial_\beta}{\partial_-} \frac{\partial_T^{2m}}{\partial_-^m} \tilde{\Psi}_{n11}^{[\Gamma_b^-]} + n^\mu n^\nu \frac{\partial_\alpha}{\partial_+} \frac{\partial_\beta}{\partial_+} \frac{\partial_T^{2n}}{\partial_+^n} \tilde{\Psi}_{\bar{n}11}^{[\Gamma_a^+]} \frac{\partial_T^{2m}}{\partial_-^m} \tilde{\Psi}_{n11}^{[\Gamma_b^-]} \right) \\ & + \frac{1}{2} \left( \bar{n}^\mu \text{Tr}[\gamma_T^\sigma \bar{\Gamma}_b^+ \gamma_T^\beta \gamma_T^\nu \gamma^\alpha \bar{\Gamma}_a^-] + \bar{n}^\nu \text{Tr}[\gamma_T^\alpha \gamma_T^\mu \gamma_T^\beta \bar{\Gamma}_b^+ \gamma_T^\sigma \bar{\Gamma}_a^-] \right) \frac{\partial_\alpha}{\partial_+} \frac{\partial_T^{2n}}{\partial_+^n} \tilde{\Psi}_{\bar{n}11}^{[\Gamma_a^+]} \frac{\partial_\beta}{\partial_-} \frac{\partial_\sigma}{\partial_-} \frac{\partial_T^{2m}}{\partial_-^m} \tilde{\Psi}_{n11}^{[\Gamma_b^-]} \\ & + \frac{1}{2} \left( n^\mu \text{Tr}[\gamma_T^\alpha \bar{\Gamma}_b^+ \gamma_T^\sigma \gamma_T^\nu \gamma^\beta \bar{\Gamma}_a^-] + n^\nu \text{Tr}[\gamma_T^\beta \gamma_T^\mu \gamma_T^\sigma \bar{\Gamma}_b^+ \gamma_T^\alpha \bar{\Gamma}_a^-] \right) \frac{\partial_\alpha}{\partial_+} \frac{\partial_\beta}{\partial_+} \frac{\partial_T^{2n}}{\partial_+^n} \tilde{\Psi}_{\bar{n}11}^{[\Gamma_a^+]} \frac{\partial_\sigma}{\partial_-} \frac{\partial_T^{2m}}{\partial_-^m} \tilde{\Psi}_{n11}^{[\Gamma_b^-]} \\ & \left. + \frac{1}{4} \text{Tr}[\gamma_T^\alpha \gamma_T^\mu \gamma_T^\sigma \bar{\Gamma}_b^+ \gamma_T^\rho \gamma_T^\nu \gamma_T^\beta \bar{\Gamma}_a^-] \frac{\partial_\alpha}{\partial_+} \frac{\partial_\beta}{\partial_+} \frac{\partial_T^{2n}}{\partial_+^n} \tilde{\Psi}_{\bar{n}11}^{[\Gamma_a^+]} \frac{\partial_\sigma}{\partial_-} \frac{\partial_\rho}{\partial_-} \frac{\partial_T^{2m}}{\partial_-^m} \tilde{\Psi}_{n11}^{[\Gamma_b^-]} \right\} + (n \leftrightarrow \bar{n}), \end{aligned} \quad (3.22)$$

where the argument of  $\tilde{\Psi}_{\bar{n}11}$  is  $(y^- n + y_T)$ , and the argument of  $\tilde{\Psi}_{n11}$  is  $(-y^+ \bar{n} - y_T)$ . This expression contains all twist-two terms of TMD factorization that appear at LO in perturbation theory. The eliminated terms have higher twist.

Finally, one integrates over  $y^\pm$  and obtains

$$\begin{aligned} W_{\text{KPC}}^{\mu\nu} & \quad (3.23) \\ = & \frac{-p_1^+ p_2^-}{4N_c} \int \frac{d^2 b}{(2\pi)^2} e^{-i(bq_T)} \int dx d\tilde{x} \delta(xp_1^+ - q^+) \delta(\tilde{x}p_2^- - q^-) \\ & \sum_{q,\bar{q}} \sum_{a,b} \sum_{n,m=0}^{\infty} \frac{\partial_{\tilde{x}}^n \partial_x^m}{n!m!} \frac{(-1)^{n+m}}{(2p_1^+ p_2^-)^{n+m}} \frac{1}{x^n \tilde{x}^m} \left\{ \right. \end{aligned}$$

$$\begin{aligned}
 & \text{Tr}[\gamma_T^\mu \bar{\Gamma}_b^+ \gamma_T^\nu \bar{\Gamma}_a^-] \partial_T^{2n} \tilde{\Phi}_{\bar{n}11}^{[\Gamma_a^+]} \partial_T^{2m} \tilde{\Phi}_{n11}^{[\Gamma_b^-]} \\
 & + \frac{i}{xp_1^+} \left( n^\nu \text{Tr}[\gamma_T^\mu \bar{\Gamma}_b^+ \gamma_T^\alpha \bar{\Gamma}_a^-] + n^\mu \text{Tr}[\gamma_T^\alpha \bar{\Gamma}_b^+ \gamma_T^\nu \bar{\Gamma}_a^-] \right) \partial_\alpha \partial_T^{2n} \tilde{\Phi}_{\bar{n}11}^{[\Gamma_a^+]} \partial_T^{2m} \tilde{\Phi}_{n11}^{[\Gamma_b^-]} \\
 & + \frac{i}{\tilde{x}p_2^-} \left( \bar{n}^\nu \text{Tr}[\gamma_T^\mu \bar{\Gamma}_b^+ \gamma_T^\alpha \bar{\Gamma}_a^-] + \bar{n}^\mu \text{Tr}[\gamma_T^\alpha \bar{\Gamma}_b^+ \gamma_T^\nu \bar{\Gamma}_a^-] \right) \partial_T^{2n} \tilde{\Phi}_{\bar{n}11}^{[\Gamma_a^+]} \partial_\alpha \partial_T^{2m} \tilde{\Phi}_{n11}^{[\Gamma_b^-]} \\
 & + \frac{1}{2x\tilde{x}p_1^+ p_2^-} \left( \text{Tr}[\gamma_T^\mu \bar{\Gamma}_b^+ \gamma_T^\beta \gamma_T^\nu \gamma_T^\alpha \bar{\Gamma}_a^-] + \text{Tr}[\gamma_T^\alpha \gamma_T^\mu \gamma_T^\beta \bar{\Gamma}_b^+ \gamma_T^\nu \bar{\Gamma}_a^-] \right. \\
 & \quad \left. + 2\bar{n}^\mu n^\nu \text{Tr}[\gamma_T^\beta \bar{\Gamma}_b^+ \gamma_T^\alpha \bar{\Gamma}_a^-] + 2n^\mu \bar{n}^\nu \text{Tr}[\gamma_T^\alpha \bar{\Gamma}_b^+ \gamma_T^\beta \bar{\Gamma}_a^-] \right) \partial_\alpha \partial_T^{2n} \tilde{\Phi}_{\bar{n}11}^{[\Gamma_a^+]} \partial_\beta \partial_T^{2m} \tilde{\Phi}_{n11}^{[\Gamma_b^-]} \\
 & - \text{Tr}[\gamma_T^\alpha \bar{\Gamma}_b^+ \gamma_T^\beta \bar{\Gamma}_a^-] \left( \frac{\bar{n}^\mu \bar{n}^\nu}{\tilde{x}^2 (p_2^-)^2} \partial_T^{2n} \tilde{\Phi}_{\bar{n}11}^{[\Gamma_a^+]} \partial_\alpha \partial_\beta \partial_T^{2m} \tilde{\Phi}_{n11}^{[\Gamma_b^-]} + \frac{n^\mu n^\nu}{x^2 (p_1^+)^2} \partial_\alpha \partial_\beta \partial_T^{2n} \tilde{\Phi}_{\bar{n}11}^{[\Gamma_a^+]} \partial_T^{2m} \tilde{\Phi}_{n11}^{[\Gamma_b^-]} \right) \\
 & + \frac{i}{2x\tilde{x}p_1^+ (p_2^-)^2} \left( \bar{n}^\mu \text{Tr}[\gamma_T^\sigma \bar{\Gamma}_b^+ \gamma_T^\beta \gamma_T^\nu \gamma_T^\alpha \bar{\Gamma}_a^-] + \bar{n}^\nu \text{Tr}[\gamma_T^\alpha \gamma_T^\mu \gamma_T^\beta \bar{\Gamma}_b^+ \gamma_T^\sigma \bar{\Gamma}_a^-] \right) \partial_\alpha \partial_T^{2n} \tilde{\Phi}_{\bar{n}11}^{[\Gamma_a^+]} \partial_\beta \partial_\sigma \partial_T^{2m} \tilde{\Phi}_{n11}^{[\Gamma_b^-]} \\
 & + \frac{i}{2x^2 \tilde{x} (p_1^+)^2 p_2^-} \left( n^\mu \text{Tr}[\gamma_T^\alpha \bar{\Gamma}_b^+ \gamma_T^\sigma \gamma_T^\nu \gamma_T^\beta \bar{\Gamma}_a^-] + n^\nu \text{Tr}[\gamma_T^\beta \gamma_T^\mu \gamma_T^\sigma \bar{\Gamma}_b^+ \gamma_T^\alpha \bar{\Gamma}_a^-] \right) \partial_\alpha \partial_\beta \partial_T^{2n} \tilde{\Phi}_{\bar{n}11}^{[\Gamma_a^+]} \partial_\sigma \partial_T^{2m} \tilde{\Phi}_{n11}^{[\Gamma_b^-]} \\
 & + \frac{1}{4(x\tilde{x}p_1^+ p_2^-)^2} \text{Tr}[\gamma_T^\alpha \gamma_T^\mu \gamma_T^\sigma \bar{\Gamma}_b^+ \gamma_T^\rho \gamma_T^\nu \gamma_T^\beta \bar{\Gamma}_a^-] \partial_\alpha \partial_\beta \partial_T^{2n} \tilde{\Phi}_{\bar{n}11}^{[\Gamma_a^+]} \partial_\sigma \partial_\rho \partial_T^{2m} \tilde{\Phi}_{n11}^{[\Gamma_b^-]} \Big\},
 \end{aligned}$$

where arguments of the TMD-distributions are  $(x, b)$  for  $\tilde{\Phi}_{\bar{n}11}$  and  $(\tilde{x}, b)$  for  $\tilde{\Phi}_{n11}$ . The derivatives  $\partial_x = \frac{\partial}{\partial x}$  ( $\partial_{\tilde{x}} = \frac{\partial}{\partial \tilde{x}}$ ) comes from the integration over  $y^- (y^+)$ . Herewith, distributions  $\tilde{\Phi}_{\bar{n}11}$  are quark distributions and  $\tilde{\Phi}_{n11}$  are anti-quark distributions. The sum  $\sum_{q, \bar{q}}$  indicates the addition of the term with quark and anti-quark distributions exchanged. This expression is valid for any polarizations. The example of unpolarized cases are given in section 5. In particular, the LP expression is given in eq. (5.8), NLP in eq. (5.9), and NNLP in eq. (5.10). The NLP part of eq. (3.23) exactly reproduces the KPC-part of the full NLP expression derived in refs. [19, 22].

It is complicated to investigate the properties (3.23) in the general form. However, one can confirm that the expression (3.23) is exactly EM-gauge invariant,

$$q_\mu W_{\text{KPC}}^{\mu\nu} = 0. \quad (3.24)$$

The truncated at  $N^n$ LP order part of series (3.23) is not transverse, but violates EM-gauge invariance up to order  $N^{n+1}$ LP. The exact restoration involves all terms. Moreover, the series (3.23) could not be split into independently gauge-invariant parts or at least such a split is not natural. The terms with derivatives  $\partial_x$  and  $\partial_{\tilde{x}}$  play important role in the restoration of gauge invariance, since they (after integration by parts) turn  $q^\pm$  to  $p^\pm$ , which is required for the cancellation between different orders. The expression (3.23) is also frame-invariant, although it is complicated to confirm. These invariances are discussed explicitly with the unpolarized example in section 5.1 and section 5.2, and they are obvious in the summed form (4.17).

### 3.3 KPCs beyond LO and the argument of the coefficient function

In the present approximation (solely twist-two TMD operators, neglecting  $q_T/Q$  corrections) the perturbative corrections come only from the corrections to the EM currents. Indeed,

an inclusion of a non-collinear external gluon increases the twist of operators, and the in-between currents interaction (alike figure 1C) induces  $q_T/Q$  corrections. Therefore, one can consider perturbative corrections for the currents  $J_{\bar{n}n}$  and  $J_{n\bar{n}}$  (3.6) independently. The NLO LP computation the coefficient function is performed in multiple works, see f.i. refs. [9–11, 22, 53]. Presently, it is known up to N<sup>4</sup>LO accuracy (4-loop) [54, 55]. The NLP NLO computation has been done explicitly in ref. [22], and the result coincides with LP. It is a non-trivial check since, intermediate expressions and the algebra of NLP computation are different from those at LP (see the detailed discussion in section 6 of ref. [22]). As it is explained above, the coefficient function of KPC must be the same as the LP coefficient function. In this section 1 would like to confirm it by explicit computation.

The perturbative computations beyond LP are much simpler in the position space. It is because Feynman propagators in the position space are free of external momenta, which appear only in the numerator of diagrams. Therefore, one should not worry about increasing singularity of power suppressed parts of integrals as it happen in the momentum space. The explanation of convenient technique for loop-computation with multiple examples can be found f.i. in refs. [22, 36, 56, 57].

The NLO correction for EM current is given by the diagram shown in figure 1 D. In the position space, this diagram (here for  $J_{\bar{n}n}(0)$ ) reads

$$\text{diag} = g^2 C_F \frac{\Gamma^2(2-\epsilon)\Gamma(1-\epsilon)}{16\pi^{d/2}} \int d^d x d^d z \frac{\bar{q}_{\bar{n}}(x)\gamma^\nu \not{x}\gamma^\mu \not{z}\gamma_\nu q_n(z)}{[-x^2+i0]^{2-\epsilon}[-z^2+i0]^{2-\epsilon}[-(x-z)^2+i0]^{1-\epsilon}}, \quad (3.25)$$

where  $C_F = (N_c^2 - 1)/(2N_c)$  and  $d = 4 - 2\epsilon$  is the parameter of the dimensional regularization. This expression incorporates operators of all powers and twists. The power expansion is obtained from the decomposition of external quark fields over “good” and “bad” components, and their expansion in the vicinity of the collinear direction. Importantly, the power expansion can be done at the level of the integrand (since integral is convergent due to the dimensional regularization). In this way, the only complications of higher-power computation is the increasing (from power to power) size of the numerator.

The LP and NLP computation has been carried out in ref. [22]. It is convenient to write the result in the mixed momentum-coordinate representation

$$\text{diag}_{\text{LP}} = \frac{a_s C(\epsilon)}{(-2k_1^+ k_2^-)^\epsilon} \bar{\xi}_{\bar{n}} \gamma_T^\mu \xi_n, \quad (3.26)$$

$$\text{diag}_{\text{NLP}} = \frac{a_s C(\epsilon)}{(-2k_1^+ k_2^-)^\epsilon} \left( -n^\mu \bar{\xi}_{\bar{n}} \overleftarrow{\not{\partial}}_T \xi_n - \bar{n}^\mu \bar{\xi}_{\bar{n}} \overrightarrow{\not{\partial}}_T \xi_n \right) \quad (3.27)$$

where

$$C(\epsilon) = 2C_F \frac{\Gamma(\epsilon)\Gamma(-\epsilon)\Gamma(2-\epsilon)}{\Gamma(3-2\epsilon)} (2-\epsilon+2\epsilon^2), \quad (3.28)$$

and  $k_1$  and  $k_2$  are the momenta of quark fields  $\bar{q}_{\bar{n}}$  and  $q_n$  (in the position space they appear as derivatives). The terms (3.26) and (3.27) provide the coefficient function for the LP and NLP parts EM-current. In this way, one confirms that the coefficient function of KPC at NLP is the same as for LP term (at least at NLO).

The coefficient  $C(\epsilon)$  contains infra-red divergences (since  $J^\mu$  is a conserved current,  $C(\epsilon)$  is UV finite). These divergences are compensated by the UV renormalization factors of operators  $U_1$ . These factors are  $Z_{U_1}$  for  $U_1$  (anti-collinear) and  $Z_{U_1}^\dagger$  for  $\bar{U}_1$  (collinear). These factors contain collinear divergence, which is removed by the analogous divergence in the rapidity renormalization. The LO expression for  $Z_{U_1}$  is

$$Z_{U_1}(\zeta) = 1 + \frac{a_s C_F}{\epsilon} \left( \frac{1}{\epsilon} + \frac{3}{2} + \ln \left( \frac{\mu^2}{\zeta} \right) \right) + \mathcal{O}(a_s^2), \quad (3.29)$$

where  $\zeta$  is the scale of the rapidity renormalization. Performing the renormalization procedure, one obtains

$$(1 + a_s C(\epsilon) + \mathcal{O}(a_s^2)) Z_{U_1}^\dagger(\zeta) Z_{U_1}(\bar{\zeta}) = C_V, \quad (3.30)$$

where  $a_s = g^2/(4\pi)^2$ . Importantly, for the cancellation of poles one must impose  $\zeta \bar{\zeta} = (2q^+q^-)^2$ , where it is used that  $k_1^+ = q^+$  and  $k_2^- = q^-$  due to the  $\delta$ -functions (3.17). The coefficient function  $C_V$  is

$$C_V = 1 + a_s C_F \left( -\mathbf{L}^2 + 3\mathbf{L} - 8 + \frac{\pi^2}{6} \right) + \mathcal{O}(a_s^2), \quad (3.31)$$

where  $\mathbf{L} = \ln(-2q^+q^-/\mu^2)$ . The coefficient function is multiplicative. The product of current has the coefficient function

$$C_0 = |C_V|^2 = 1 + 2a_s C_F \left( -|\mathbf{L}|^2 + 3|\mathbf{L}| - 8 + \frac{7\pi^2}{6} \right) + \mathcal{O}(a_s^2), \quad (3.32)$$

where it has been taken into account that  $q^+q^- > 0$ .

Let me emphasize that the argument of the coefficient function is  $2q^+q^-$ . This is the result of the formal computation in the factorization approach (see e.g. [22, 53]). Often, the argument of coefficient function at LP is replaced by  $q^2$  for simplicity. However, it is not entirely correct since  $q^2 = 2q^+q^- + q_T^2$  has an indefinite power counting.

The NLO computation can be easily automatized (necessary integrals in general form are given in appendix B or ref. [22]). Using the package *FeynCalc* [58], I have generated the first few powers.<sup>6</sup> They are

$$\begin{aligned} \text{diag}_{\text{N}^2\text{LP}} &= \frac{a_s C(\epsilon)}{(-2k_1^+k_2^-)^\epsilon} \left[ -\frac{1}{2} \bar{\xi}_n \overleftarrow{\frac{\partial}{\partial_+}} \gamma_T^\mu \overrightarrow{\frac{\partial}{\partial_+}} \xi_n - \epsilon \frac{(k_1k_2)_T}{k_1^+k_2^-} \bar{\xi}_n \gamma_T^\mu \xi_n \right], \\ \text{diag}_{\text{N}^3\text{LP}} &= \frac{a_s C(\epsilon)}{(-2k_1^+k_2^-)^\epsilon} \left( -\epsilon \frac{(k_1k_2)_T}{k_1^+k_2^-} \right) \left( -n^\mu \bar{\xi}_n \overleftarrow{\frac{\partial}{\partial_+}} \xi_n - \bar{n}^\mu \bar{\xi}_n \overrightarrow{\frac{\partial}{\partial_+}} \xi_n \right), \\ \text{diag}_{\text{N}^4\text{LP}} &= \frac{a_s C(\epsilon)}{(-2k_1^+k_2^-)^\epsilon} \left[ \frac{\epsilon}{2} \frac{(k_1k_2)_T}{k_1^+k_2^-} \bar{\xi}_n \overleftarrow{\frac{\partial}{\partial_+}} \gamma_T^\mu \overrightarrow{\frac{\partial}{\partial_+}} \xi_n + \left( \frac{\epsilon(1+\epsilon)}{2} \frac{(k_1k_2)_T}{k_1^+k_2^-} - \frac{\epsilon}{4} \frac{k_{1T}^2 k_{2T}^2}{(k_1^+k_2^-)^2} \right) \bar{\xi}_n \gamma_T^\mu \xi_n \right], \\ \text{diag}_{\text{N}^5\text{LP}} &= \frac{a_s C(\epsilon)}{(-2k_1^+k_2^-)^\epsilon} \left( \frac{\epsilon(1+\epsilon)}{2} \frac{(k_1k_2)_T^2}{(k_1^+k_2^-)^2} - \frac{\epsilon}{4} \frac{k_{1T}^2 k_{2T}^2}{(k_1^+k_2^-)^2} \right) \left( -n^\mu \bar{\xi}_n \overleftarrow{\frac{\partial}{\partial_+}} \xi_n - \bar{n}^\mu \bar{\xi}_n \overrightarrow{\frac{\partial}{\partial_+}} \xi_n \right), \end{aligned} \quad (3.33)$$

<sup>6</sup>I thank Oscar del Rio for the cross-check of the N<sup>2</sup>LP computation.

$$\text{diag}_{\text{N}^6\text{LP}} = \frac{a_s C(\epsilon)}{(-2k_1^+ k_2^-)^\epsilon} \left[ \left( \frac{\epsilon(1+\epsilon)}{2} \frac{(k_1 k_2)_T^2}{(k_1^+ k_2^-)^2} - \frac{\epsilon}{4} \frac{k_{1T}^2 k_{2T}^2}{(k_1^+ k_2^-)^2} \right) \bar{\xi}_n \overleftarrow{\frac{\partial}{\partial_+}} \gamma_T^\mu \overrightarrow{\frac{\partial}{\partial_+}} \xi_n \right. \\ \left. + \left( -\frac{\epsilon(1+\epsilon)(2+\epsilon)}{6} \frac{(k_1 k_2)_T^3}{(k_1^+ k_2^-)^3} - \frac{\epsilon(1+\epsilon)}{4} \frac{k_{1T}^2 k_{2T}^2 (k_1 k_2)_T}{(k_1^+ k_2^-)^3} \right) \bar{\xi}_n \gamma_T^\mu \xi_n \right],$$

and so on. For the convincingness, the computation has been performed up to  $\text{N}^{10}\text{LP}$ , and it reproduces the presented pattern.

At the first glance, expressions (3.33) demonstrate the violation of the factorization theorem. Indeed, naively, one should expect to obtain the effective current (3.35) multiplied by  $C(\epsilon)$ . I.e. all terms except the first for  $\text{diag}_{\text{N}^2\text{LP}}$  should vanish. Instead, the expressions contain the new terms, equipped by different coefficients. It is especially confusing that these new terms have different order of  $\epsilon$  singularities. Due to the latter, these terms cannot be renormalized by  $Z_{U1}$ 's (3.30), and the factorization theorem seems to be violated at  $\text{NNLP}$  and higher.

Further inspection of corrections (3.33) reveals a certain pattern. The all-power series (3.33) can be rewritten as

$$\sum_{k=0}^{\infty} \text{diag}_{\text{N}^k\text{LP}} = \frac{a_s C(\epsilon)}{\left( -2k_1^+ k_2^- - 2(k_1 k_2)_T - \frac{k_{1T}^2 k_{2T}^2}{2k_1^+ k_2^-} \right)^\epsilon} J_{\bar{n}n}^{\text{tw}2}, \quad (3.34)$$

where

$$J_{\bar{n}n}^{\text{tw}2} = \bar{\xi}_n \gamma_T^\mu \xi_n - n^\mu \bar{\xi}_n \overleftarrow{\frac{\partial}{\partial_+}} \xi_n - \bar{n}^\mu \bar{\xi}_n \overrightarrow{\frac{\partial}{\partial_+}} \xi_n - \frac{1}{2} \bar{\xi}_n \overleftarrow{\frac{\partial}{\partial_+}} \gamma_T^\mu \overrightarrow{\frac{\partial}{\partial_+}} \xi_n, \quad (3.35)$$

is the twist-two part of the EM current. Therefore, the coefficient function for all KPCs is the same, but has a different argument. The argument of the coefficient function is

$$X = (2q^+ q^- - 2(k_1 k_2)_T - k_{1T}^2 k_{2T}^2 / (2q^+ q^-)), \quad (3.36)$$

instead of  $2q^+ q^-$ , that appears at LP and NLP cases. The cancellation of poles also imposes

$$\zeta \bar{\zeta} = X^2.$$

The coefficient function  $C_0(X)$  can be used only as the whole, because any strict fixed-power expansion would lead to the non-cancellation of poles between infra-red poles of coefficient function, and UV poles of twist-one operators. To my best knowledge, it is the first example of accumulation of power corrections into the argument of the coefficient function in factorization theorems. In section 4.2, it is shown that  $X^2$  turns to  $Q^2$  once KPCs are summed.

In the position space the variables  $k_{1,2}$  are represented by the differential operators. Therefore, in the position space, the coefficient function (3.34) is a differential operator in the transverse space. That is the main reasons to switch to the momentum space in the following discussions.

Another contribution that reveals only beyond LO is the restoration of the boost-invariance. The boost-invariance demands that the factorization theorem is independent of

the values of  $\zeta$ 's as far as  $\zeta\bar{\zeta} = X^2$  (this can be demonstrated using different approaches [50, 59, 60]). Therefore, the factorized expression should be invariant under

$$\zeta \rightarrow \alpha\zeta, \quad \bar{\zeta} \rightarrow \alpha^{-1}\bar{\zeta}, \quad (3.37)$$

at each given power. The expression (3.23) does not preserve this property. Indeed, the combinations with transverse derivatives, such as  $\Phi_{\bar{n}}(\zeta)\partial_{T\mu}\Phi_n(\bar{\zeta})$  are not invariant but produces the terms  $\sim (\alpha - 1)\partial_{\mu}\mathcal{D}$  (for  $\alpha \sim 1$ ), where  $\mathcal{D}$  is the Collins-Soper kernel (2.26).

The restoration of the boost-invariance happens with the accounting of the higher-twist terms. The higher-twist distributions appears in the factorized expressions with the divergent convolution. These divergences are called the special rapidity divergences [26]. They are somewhat similar to the end-point divergences in the heavy-quark factorizations [61, 62], but have a different nature, and are rapidity divergences according to the general criterion [50]. The leading contribution of these divergences is proportional to the derivative of the Collins-Soper kernel and twist-two TMD distributions. In the sum of all terms these divergences cancel, and leave the remnant proportional to  $\ln(\zeta/\bar{\zeta})$ , which restores the boost-invariance. Practically, it leads to the replacement

$$\partial_{T\mu}\Phi_n \rightarrow D_{\mu}\left(\frac{\bar{\zeta}}{\zeta}\right)\Phi_n = \left[\frac{\partial}{\partial b^{\mu}} - \frac{1}{2}\left(\frac{\partial\mathcal{D}(b,\mu)}{\partial b^{\mu}}\right)\ln\left(\frac{\bar{\zeta}}{\zeta}\right)\right]\Phi_n, \quad (3.38)$$

and analogous for  $\Phi_{\bar{n}}$ . Here,

$$D_{\rho}\left(\frac{\bar{\zeta}}{\zeta}\right) = \frac{\partial}{\partial b^{\rho}} - \frac{1}{2}\left(\frac{\partial\mathcal{D}(b,\mu)}{\partial b^{\rho}}\right)\ln\left(\frac{\bar{\zeta}}{\zeta}\right). \quad (3.39)$$

It is straightforward to check that the structure  $\Phi_{\bar{n}}(\zeta)D_{\mu}(\bar{\zeta}/\zeta)\Phi_n(\bar{\zeta})$  is invariant under (3.37)

The mechanism of special rapidity divergences is well-understood at NLP, [24, 25], but not beyond. The explicit check requires the computation with increased number of loops for each next power. The terms  $\sim (\partial_T\mathcal{D})^n$ , which required at  $n$ 'th order, are  $\mathcal{O}(a_s^n)$  in perturbative order and could be obtained only with  $n$ -loop computation. Also, one needs the complete higher-twist structure to extract the divergences. Such computation goes far beyond the present work.<sup>7</sup> Nonetheless, it seems safe to extend the NLP case to all powers and declare that the complete expression for KPCs (3.23) requires the replacement (3.38). This replacement restores the boost-invariance (3.37) at each order of power expansion independently.

In this section, I have demonstrated at NLO that the coefficient functions for all KPCs are the same. It is in the full agreement with the expectations, since KPCs must reproduce LP term, in order to preserve EM-gauge invariance (discussed in section 5.1). However, it appears that the coefficient function has the mixed-power argument (as shown

---

<sup>7</sup>I have checked explicitly that the cancellation of special-rapidity divergences between twist-three part of the NNLP terms, which are indicated as  $D\Phi_2 \times \Phi_3$  and  $\Phi_3 \times \Phi_3$  in the expression (1.3). These terms are relatively simple to obtain using NLP factorization. I have found that their one-loop special rapidity divergences produce the terms required for restoration of boost-invariance. The performed check is incomplete (and for that reason is not presented here explicitly), because it was done using only one-loop computation and caught only the terms  $\sim \partial_T\mathcal{D}$ , but does not catch terms  $\sim (\partial_T\mathcal{D})^2$ .

in eq. (3.34)). Also, I argue that the transverse derivatives must be replaced by the boost-invariant transverse derivatives (3.38). Equipped by these modifications the expression (3.23) represents the all-power part of the TMD factorization that contains solely twist-two distributions, without  $q_T/Q$  corrections.

## 4 Summation of KPC series

The series of KPCs can be summed to a simpler expression. The direct summation of expression (3.33) is possible but very tedious. The main complication is the derivatives over  $x$  and  $\tilde{x}$  which produce generating functions with complicated argument (see (5.29)). The simpler way is to sum the multipole expansion separately for each collinear sector before the tensor decomposition. It leads a simple and intuitive picture of KPCs as the scattering of free massless partons. The details of computation are presented in this section.

### 4.1 Twist-two part of a general TMD correlator

The procedure of the assembling fields into matrix elements and the multipole expansion are commuting operations. Due to it, one can collect the fields of different modes into matrix elements already in eq. (3.5). The result is (compare with (3.12))

$$\begin{aligned}
 W^{\mu\nu} = & -\frac{1}{4N_c} \int \frac{d^4y}{(2\pi)^4} e^{-i(yq)} \sum_{a,b} \left( \text{Tr} \left( \gamma^\mu \bar{\Gamma}_b \gamma^\nu \bar{\Gamma}_a \right) \tilde{\Psi}_n^{[\Gamma_a]}(y) \tilde{\Psi}_n^{[\Gamma_b]}(-y) \right. \\
 & \left. + \text{Tr} \left( \gamma^\mu \bar{\Gamma}_a \gamma^\nu \bar{\Gamma}_b \right) \tilde{\Psi}_n^{[\Gamma_a]}(-y) \tilde{\Psi}_n^{[\Gamma_b]}(y) \right), \tag{4.1}
 \end{aligned}$$

where  $\Psi(y)$  are the generalization of (3.13) for a four-dimensional argument,

$$\tilde{\Psi}_n^{[\Gamma]}(y) = \langle p_1 | \bar{q}_n(y) [y, -\infty n + y_T + \bar{n}y^+] \frac{\Gamma}{2} [-\infty n, 0] q_n(0) | p_1 \rangle, \tag{4.2}$$

$$\tilde{\Psi}_n^{[\Gamma]}(y) = \langle p_2 | \bar{q}_n(y) [y, -\infty \bar{n} + y_T + ny^-] \frac{\Gamma}{2} [-\infty \bar{n}, 0] q_n(0) | p_2 \rangle. \tag{4.3}$$

These matrix elements are alike ordinary TMD distributions, but with a generally-valued separation between semi-compact operators. The direction of Wilson lines is preserved.

The expression (4.2) is not a factorization theorem, but only an intermediate expression that is used to sum KPCs. The functions  $\Psi$  should be understood as the generating functions for the series of power corrections (3.19). I.e

$$\tilde{\Psi}_n^{[\Gamma]}(y) = \sum_{n=0}^{\infty} \frac{(y^+)^n}{n!} [\partial_-^n \tilde{\Psi}_n^{[\Gamma]}(y^- n + y_T)], \tag{4.4}$$

where

$$[\partial_-^n \tilde{\Psi}_n^{[\Gamma]}(y^- n + y_T)] = \langle p_1 | \bar{q}_n(y^- n + y_T) [y^- n + y_T, -\infty n + y_T] \overleftarrow{\partial}^n \frac{\Gamma}{2} [-\infty n, 0] q_n(0) | p_1 \rangle, \tag{4.5}$$

and similar for  $\Psi_n$ . Substituting this expression into eq. (4.1) and making the twist-decomposition one gets eq. (3.19).

The function  $\Psi$  is an infinite series of TMD distributions of different twists. The twist-two part of (4.4) can be extracted term-by-term using EOMs (2.11), (3.20) in the same way as it was done in section 3.2. It is convenient to present the result in the following form

$$\begin{aligned} [\tilde{\Psi}_{\bar{n}}^{[\Gamma]}(y)]_{\text{tw-2}} &= \sum_{n=0}^{\infty} \frac{(y^+)^n}{n!} \left( \frac{-\partial_T^2}{2\partial_+} \right)^n \left[ \tilde{\Psi}_{\bar{n}11}^{[\Gamma]}(y^-n+y_T) - \frac{\partial_{T\alpha}}{2\partial_+} \tilde{\Psi}_{\bar{n}11}^{[\gamma^\alpha\gamma^+\Gamma+\Gamma\gamma^+\gamma^\alpha]}(y^-n+y_T) \right. \\ &\quad \left. + \frac{\partial_{T\alpha}\partial_{T\beta}}{4\partial_+^2} \tilde{\Psi}_{\bar{n}11}^{[\gamma^\alpha\gamma^+\Gamma+\Gamma\gamma^+\gamma^\beta]}(y^-n+y_T) \right], \end{aligned} \quad (4.6)$$

where  $\partial_{T\alpha} = \frac{\partial}{\partial y_T^\alpha}$ ,  $\partial_+ = \frac{\partial}{\partial y_-}$  and

$$[\Gamma] = \left[ \frac{\gamma^+\gamma^-}{2} \Gamma \frac{\gamma^-\gamma^+}{2} \right]. \quad (4.7)$$

The dropped terms are the distributions of twist-three and higher. Note that on the level of the cross-section the derivatives turns to the boost-invariant derivatives (3.39), and thus the expression satisfy the leading-twist TMD evolution equation (2.25), (2.26).

The summation of series (4.6) is straightforward in the momentum space, where it reads

$$\begin{aligned} [\tilde{\Psi}_{\bar{n}}^{[\Gamma]}(y)]_{\text{tw-2}} &= \sum_{n=0}^{\infty} \frac{p^+}{n!} \int dx d^2 k_T e^{ixp^+y^- + i(ky)_T} \left( \frac{-iy^+k_T^2}{2xp^+} \right)^n \left[ \right. \\ &\quad \left. \Phi_{\bar{n}11}^{[\Gamma]}(x, k_T) - \frac{k_{T\alpha}}{2xp^+} \Phi_{\bar{n}11}^{[\gamma^\alpha\gamma^+\Gamma+\Gamma\gamma^+\gamma^\alpha]}(x, k_T) + \frac{k_{T\alpha}k_{T\beta}}{4(xp^+)^2} \Phi_{\bar{n}11}^{[\gamma^\alpha\gamma^+\Gamma+\Gamma\gamma^+\gamma^\beta]}(x, k_T) \right]. \end{aligned} \quad (4.8)$$

The sum over  $n$  produces the exponent  $\exp(-iy^+k_T^2/(2xp^+))$ . Introducing an auxiliary variable  $k^- = -k_T^2/(2xp^+)$ , one can rewrite (4.8) in a compact form

$$\begin{aligned} [\tilde{\Psi}_{\bar{n}}^{[\Gamma]}(y)]_{\text{tw-2}} &= p^+ \int dx d^4 k (2k^+) e^{i(ky)} \delta(k^+ - xp^+) \delta(k^2) \left[ \right. \\ &\quad \left. \Phi_{\bar{n}11}^{[\Gamma]}(x, k_T) - \frac{k_{T\alpha}}{2xp^+} \Phi_{\bar{n}11}^{[\gamma^\alpha\gamma^+\Gamma+\Gamma\gamma^+\gamma^\alpha]}(x, k_T) + \frac{k_{T\alpha}k_{T\beta}}{4(xp^+)^2} \Phi_{\bar{n}11}^{[\gamma^\alpha\gamma^+\Gamma+\Gamma\gamma^+\gamma^\beta]}(x, k_T) \right], \end{aligned} \quad (4.9)$$

where the factor  $2k^+$  is the Jacobian for  $\delta(k^2)$  and all scalar products are four-dimensional. The same expression is valid for  $\tilde{\Psi}_n$  after the replacement of  $n \leftrightarrow \bar{n}$ .

The twist-two part of a general TMD distribution is related to the matrix element of the free massless parton. In fact, this statement is correct for leading-twist (twist-two) distributions of any kind. For example, in the case of PDF the equivalent derivation can be found in ref. [36].

The hadron tensor with inclusion of all KPCs is given by eq. (4.1) with replacement  $\tilde{\Psi} \rightarrow [\tilde{\Psi}]_{\text{tw-2}}$ . For the particular elements of the Dirac basis one finds

$$[\tilde{\Psi}_{\bar{n}}^{[1]}(y)]_{\text{tw-2}} = 0, \quad (4.10)$$

$$[\tilde{\Psi}_{\bar{n}}^{[\gamma^5]}(y)]_{\text{tw-2}} = 0, \quad (4.11)$$

$$[\tilde{\Psi}_{\bar{n}}^{[\gamma^\mu]}(y)]_{\text{tw-2}} = 2p^+ \int dx d^4 k \delta(k^+ - xp^+) \delta(k^2) e^{i(ky)} k^\mu \Phi_{\bar{n}11}^{[\gamma^+]}, \quad (4.12)$$

$$\left[\tilde{\Psi}_{\bar{n}}^{[\gamma^{\mu}\gamma^5]}(y)\right]_{\text{tw-2}} = 2p^+ \int dx d^4k \delta(k^+ - xp^+) \delta(k^2) e^{i(ky)} k^{\mu} \Phi_{\bar{n}11}^{[\gamma^+\gamma^5]}(x, k_T), \quad (4.13)$$

$$\left[\tilde{\Psi}_{\bar{n}}^{[i\sigma^{\mu\nu}\gamma^5]}(y)\right]_{\text{tw-2}} = 2p^+ \int dx d^4k \delta(k^+ - xp^+) \delta(k^2) e^{i(ky)} \left( g_T^{\mu\alpha} k^{\nu} - g_T^{\nu\alpha} k^{\mu} + \frac{k^{\mu}n^{\nu} - n^{\mu}k^{\nu}}{k^+} k_T^{\alpha} \right) \Phi_{\bar{n}11}^{[i\sigma^{\alpha+}\gamma^5]}(x, k_T). \quad (4.14)$$

One can see that the TMD correlators are divergence-free and satisfy Laplace equation

$$\frac{\partial^2}{\partial y^{\mu} \partial y_{\mu}} \left[\tilde{\Psi}_{\bar{n}}^{[\Gamma]}(y)\right]_{\text{tw-2}} = 0, \quad (4.15)$$

$$\frac{\partial}{\partial y^{\mu}} \left[\tilde{\Psi}_{\bar{n}}^{[\gamma^{\mu}]}(y)\right]_{\text{tw-2}} = \frac{\partial}{\partial y^{\mu}} \left[\tilde{\Psi}_{\bar{n}}^{[\gamma^{\mu}\gamma^5]}(y)\right]_{\text{tw-2}} = \frac{\partial}{\partial y^{\mu}} \left[\tilde{\Psi}_{\bar{n}}^{[i\sigma^{\mu\nu}\gamma^5]}(y)\right]_{\text{tw-2}} = 0. \quad (4.16)$$

These equations are identical to the equations for twist-two part of collinear operators [32, 36, 63].

## 4.2 Hadron tensor with summed KPCs

Substituting expressions (4.10)–(4.14) into (4.1) one obtains the hadron tensor with re-summed KPCs. The general expression reads

$$W_{\text{KPC}}^{\mu\nu} = -\frac{4p_1^+ p_2^-}{N_c} C_0 \left(\frac{Q^2}{\mu^2}\right) \int d\xi_1 d\xi_2 \int d^4k_1 d^4k_2 \delta^4(q - k_1 - k_2) \delta(k_1^+ - \xi_1 p_1^+) \delta(k_2^- - \xi_2 p_2^-) \delta(k_1^2) \delta(k_2^2) \left\{ \begin{aligned} & ((k_1 k_2) g^{\mu\nu} - k_1^{\mu} k_2^{\nu} - k_2^{\mu} k_1^{\nu}) \left( \Phi_{\bar{n}11}^{[\gamma^+]} \Phi_{n11}^{[\gamma^-]} + \Phi_{\bar{n}11}^{[\gamma^+\gamma^5]} \Phi_{n11}^{[\gamma^-\gamma^5]} \right) \\ & + i\epsilon^{\mu\nu\alpha\beta} k_1^{\alpha} k_2^{\beta} \left( \Phi_{\bar{n}11}^{[\gamma^+]} \Phi_{n11}^{[\gamma^-\gamma^5]} - \Phi_{\bar{n}11}^{[\gamma^+\gamma^5]} \Phi_{n11}^{[\gamma^-]} \right) + t_{\alpha\beta}^{\mu\nu} \Phi_{\bar{n}11}^{[i\sigma^{\alpha+}\gamma^5]} \Phi_{n11}^{[i\sigma^{\beta-}\gamma^5]} \end{aligned} \right\}, \quad (4.17)$$

where

$$t_{\alpha\beta}^{\mu\nu} = \frac{\text{Tr}[\gamma^{\mu} \sigma^{\gamma\rho} \gamma^{\nu} \sigma^{\sigma\delta}]}{16} \left( g_T^{\sigma\alpha} k_1^{\delta} - g_T^{\delta\alpha} k_1^{\sigma} + \frac{k_1^{\sigma} n^{\delta} - n^{\delta} k_1^{\sigma}}{k_1^+} k_T^{\alpha} \right) \left( g_T^{\gamma\beta} k_2^{\rho} - g_T^{\rho\beta} k_2^{\gamma} + \frac{k_2^{\gamma} \bar{n}^{\rho} - \bar{n}^{\rho} k_2^{\gamma}}{k_2^-} k_T^{\beta} \right)$$

In this expression, the product of two distributions should be understood as following

$$\Phi_{\bar{n}}^{[\Gamma_1]} \Phi_n^{[\Gamma_2]} = \Phi_{\bar{n},q}^{[\Gamma_1]}(\xi_1, k_{1T}; \mu^2, Q^2) \Phi_{n,\bar{q}}^{[\Gamma_2]}(\xi_2, k_{2T}; \mu^2, Q^2) + \Phi_{\bar{n},\bar{q}}^{[\Gamma_1]}(\xi_1, k_{1T}; \mu^2, Q^2) \Phi_{n,q}^{[\Gamma_2]}(\xi_2, k_{2T}; \mu^2, Q^2), \quad (4.18)$$

where in the second term the quark and anti-quark TMDPDFs are exchanged. The collinear momentum fractions are denoted by  $\xi$  in order to distinguish them from the ordinary Drell-Yan momentum fractions  $x$ . Note, that in eq. (4.17) all indices and momenta are four-dimensional. The argument of the coefficient function is restored to  $Q^2$ . It is due to the  $\delta$ -functions in the integrand of eq. (4.17), which allows to rewrite  $X$  (3.36) as  $X = 2(k_1 k_2) = (k_1 + k_2)^2 = Q^2$ .

The expression (4.17) has the same hard-coefficient function as the LP TMD factorization, and the TMDPDFs are the usual TMDPDFs obeying the standard evolution equations. The main limitation of eq. (4.17) is that the rapidity scales are taken in the symmetric point,  $\zeta = \bar{\zeta} = Q^2$ . It is necessary in order to remove the terms proportional to  $\partial_\mu \mathcal{D} \ln(\bar{\zeta}/\zeta)$ . Otherwise the summation of KPCs is not possible. However, if required, these terms could be restored infinitesimally. Practically, the restriction  $\zeta = \bar{\zeta} = Q^2$  is not essential, since all phenomenological applications are done at the symmetric point.

The expression (4.17) is the TMD factorization with resummed KPCs and the main result of this work. The dropped terms either contain TMD distributions of higher-twist, either proportional to  $q_T/Q$ . Being expanded at large- $Q$   $W_{\text{KPC}}^{\mu\nu}$  reproduces the ordinary fixed-power TMD factorization power-by-power.

The hadron tensor  $W_{\text{KPC}}^{\mu\nu}$  is exactly transverse to the vector  $q^\mu$  (3.24). It is also frame-invariant, up to the corrections which come due to the selection direction of the Wilson lines. The later follows from the initial assumption on the momenta counting rules for partons. These counting rules also receive power corrections and modify the definition of TMD distributions. The resulting factorization theorem does not depend on the definition of vectors  $n$  and  $\bar{n}$ . The inclusion of these corrections goes beyond the present work, and will be performed in the future.

## 5 KPC for the unpolarized Drell-Yan reaction

To exemplify the general structure derived above, let me consider the case of the spherically-symmetric contribution to the unpolarized Drell-Yan. In this case, the only<sup>8</sup> contributing TMDPDFs are unpolarized TMDPDFs  $f_1$ . They are defined as

$$\Phi_{11}^{[\gamma^+]}(x, k_T) = f_1(x, k_T) + \dots, \quad (5.1)$$

where dots indicate the omitted Siverson function. In this case the series of KPCs (3.23) reads

$$\begin{aligned} W_{f_1 f_1}^{\mu\nu} = & \frac{p_1^+ p_2^-}{N_c} \int d^2 k_{1T} d^2 k_{2T} \delta^{(2)}(q_T - k_1 - k_2) \int dx d\tilde{x} \delta(q^+ - x p_1^+) \delta(q^- - \tilde{x} p_2^-) \quad (5.2) \\ & C_0 \left( \frac{X}{\mu^2} \right) \sum_{n,m=0}^{\infty} \frac{(\partial_{\tilde{x}})^n (\partial_x)^m}{n! m!} \frac{(k_{1T}^2)^n (k_{2T}^2)^m}{(2p_1^+ p_2^-)^{n+m} x^n \tilde{x}^m} \left\{ \right. \\ & - g_T^{\mu\nu} + \frac{n^\nu k_{1T}^\mu + n^\mu k_{1T}^\nu}{x p_1^+} + \frac{\bar{n}^\nu k_{2T}^\mu + \bar{n}^\mu k_{2T}^\nu}{\tilde{x} p_2^-} - \frac{g^{\mu\nu} (k_1 k_2)_T - k_{1T}^\mu k_{2T}^\nu - k_{1T}^\nu k_{2T}^\mu}{x \tilde{x} p_1^+ p_2^-} \\ & - \frac{\bar{n}^\mu \bar{n}^\nu k_{2T}^2}{(\tilde{x} p_2^-)^2} - \frac{n^\mu n^\nu k_{1T}^2}{(x p_1^+)^2} - \frac{\bar{n}^\mu k_{1T}^\nu + \bar{n}^\nu k_{1T}^\mu}{2x p_1^+ (\tilde{x} p_2^-)^2} k_{2T}^2 - \frac{n^\mu k_{2T}^\nu + n^\nu k_{2T}^\mu}{2(x p_1^+)^2 \tilde{x} p_2^-} k_{1T}^2 \\ & \left. - g_T^{\mu\nu} \frac{k_{1T}^2 k_{2T}^2}{(2x \tilde{x} p_1^+ p_2^-)^2} \right\} f_1(x, k_{1T}; \mu, X) f_1(\tilde{x}, k_{2T}; \mu, X), \end{aligned}$$

---

<sup>8</sup>Another pair of TMDPDFs that contributes to the unpolarized Drell-Yan hadronic tensor is the pair of Boer-Mulder functions  $h_1^\perp$ . Generally, all angular structure functions contain contributions proportional to  $f_1 f_1$  and  $h_1^\perp h_1^\perp$ , see f.i. refs. [17, 21]. I have checked explicitly that the spherically-symmetric structure function with KPCs does not contain contribution  $\sim h_1^\perp h_1^\perp$ .

where  $X$  is given in eq. (3.36). The summed expression (4.17) turns to

$$\begin{aligned}
 W_{\text{KPC}f_1f_1}^{\mu\nu} &= -\frac{4p_1^+p_2^-}{N_c}C_0\left(\frac{Q^2}{\mu^2}\right)\int d\xi_1d\xi_2\int d^4k_1d^4k_2\delta^4(q-k_1-k_2) \\
 &\delta(k_1^+-\xi_1p_1^+)\delta(k_2^--\xi_2p_2^-)\delta(k_1^2)\delta(k_2^2) \\
 &((k_1k_2)g^{\mu\nu}-k_1^\mu k_2^\nu-k_2^\mu k_1^\nu)f_1(\xi_1,k_{1T};\mu,Q^2)f_1(\xi_2,k_{2T};\mu,Q^2).
 \end{aligned}
 \tag{5.3}$$

The same expression could be obtained computing the Drell-Yan reaction with free massless quarks produced by TMD distributions. In other words, this result exactly reproduces the naive parton model. However, it is not naive, because it is known which terms of factorization series were dropped and thus, it could be systematically improved.

For the examples discussed in this section, it is instructive to have the first three terms of the standard power expansion in the explicit form. Let me denote the terms of expansion as

$$W_{f_1f_1}^{\mu\nu} = W_{0,f_1f_1}^{\mu\nu} + W_{1,f_1f_1}^{\mu\nu} + W_{2,f_1f_1}^{\mu\nu} + \dots, \tag{5.4}$$

where  $W_n \sim Q^{-n}$  in power counting. Explicitly, these terms are

$$W_{0,f_1f_1}^{\mu\nu} = \frac{-g_T^{\mu\nu}}{N_c}\int d^2k_{1T}d^2k_{2T}\delta^{(2)}(q_T-k_{1T}-k_{2T})C_0f_1(x_1,k_{1T})f_1(x_2,k_{2T}), \tag{5.5}$$

$$\begin{aligned}
 W_{1,f_1f_1}^{\mu\nu} &= \frac{1}{N_c}\int d^2k_{1T}d^2k_{2T}\delta^{(2)}(q_T-k_{1T}-k_{2T})C_0 \\
 &\times\left(\frac{\bar{n}^\mu k_{2T}^\nu+k_{2T}^\mu\bar{n}^\nu}{q^-}+\frac{n^\mu k_{1T}^\nu+k_{1T}^\mu n^\nu}{q^+}\right)f_1(x_1,k_{1T})f_1(x_2,k_{2T}),
 \end{aligned}
 \tag{5.6}$$

$$\begin{aligned}
 W_{2,f_1f_1}^{\mu\nu} &= \frac{1}{N_c}\int d^2k_{1T}d^2k_{2T}\delta^{(2)}(q_T-k_{1T}-k_{2T})C_0 \\
 &\times\left[-\frac{g^{\mu\nu}(k_1\cdot k_2)_T-k_{1T}^\mu k_{2T}^\nu-k_{2T}^\mu k_{1T}^\nu}{q^+q^-}-\frac{n^\mu n^\nu}{(q^+)^2}k_{1T}^2-\frac{\bar{n}^\mu\bar{n}^\nu}{(q^-)^2}k_{2T}^2\right. \\
 &\left.+\frac{g_T^{\mu\nu}}{2q^+q^-}\left(x_2k_{1T}^2\frac{\partial}{\partial x_2}+x_1k_{2T}^2\frac{\partial}{\partial x_1}\right)\right]f_1(x_1,k_{1T})f_1(x_2,k_{2T}),
 \end{aligned}
 \tag{5.7}$$

where  $x_1 = q^+/p_1^+$  and  $x_2 = q^-/p_2^-$ . The rapidity-scaling arguments for the LP expression are  $f_1(x_1, k_{1T}; \mu, \zeta)f_1(x_2, k_{2T}; \mu, \bar{\zeta})$ , and for the NLP they are  $f_1(x_1, k_{1T}; \mu, 2q^+q^-)f_1(x_2, k_{2T}; \mu, 2q^+q^-)$ . The argument for the coefficient function at LP and NLP is  $2q^+q^-/\mu^2$ . For the (pure) NNLP term one could not specify rapidity scales and the argument of the coefficient function without introducing higher-power terms as it is discussed above.

In the position space expressions for the first three power are

$$W_{0,f_1f_1}^{\mu\nu} = \frac{-g_T^{\mu\nu}}{N_c}\int\frac{d^2b}{(2\pi)^2}e^{-i(q_Tb)}C_0\tilde{f}_1(x_1,b)\tilde{f}_1(x_2,b), \tag{5.8}$$

$$\begin{aligned}
 W_{1,f_1f_1}^{\mu\nu} &= \frac{1}{N_c}\int\frac{d^2b}{(2\pi)^2}e^{-i(q_Tb)}C_0 \\
 &\times\left(-\frac{i\bar{n}^\mu}{q^-}\tilde{f}_1(x_1,b)D^\nu\tilde{f}_1(x_2,b)-\frac{in^\mu}{q^+}D^\nu\tilde{f}_1(x_1,b)\tilde{f}_1(x_2,b)\right),
 \end{aligned}
 \tag{5.9}$$

$$\begin{aligned}
 W_{2,f_1 f_1}^{\mu\nu} = & \frac{1}{N_c} \int \frac{d^2 b}{(2\pi)^2} e^{-i(q_T b)} C_0 \left[ \frac{g^{\mu\nu} D_\alpha \tilde{f}_1 D^\alpha \tilde{f}_1 - D^\mu \tilde{f}_1 D^\nu \tilde{f}_1 - D^\nu \tilde{f}_1 D^\mu \tilde{f}_1}{q^+ q^-} \right. \\
 & \left. + \frac{n^\mu n^\nu}{(q^+)^2} D^2 \tilde{f}_1 \tilde{f}_1 + \frac{\bar{n}^\mu \bar{n}^\nu}{(q^-)^2} \tilde{f}_1 D^2 \tilde{f}_1 + \frac{g_T^{\mu\nu}}{2q^+ q^-} \left( x_1 \frac{\partial \tilde{f}_1}{\partial x_1} D^2 \tilde{f}_1 + x_2 D^2 \tilde{f}_1 \frac{\partial \tilde{f}_1}{\partial x_2} \right) \right], \quad (5.10)
 \end{aligned}$$

where arguments for TMDPDFs are omitted completely for the NNLP case. The operator  $D$  is the ‘‘long’’ derivative defined in eq. (3.39). In contrast to the momentum-space expressions (5.5)–(5.7), the position space expressions are boost-invariant (3.37), and the rapidity-scaling arguments are  $\zeta$  and  $\bar{\zeta}$ .

The expressions (5.8)–(5.10) can be compared to the literature. Naturally, the NLP part agrees with the one computed in ref. [22]. Another computation was performed in refs. [19, 21] using the small- $x$  approximation. These results can be directly compared to eq. (5.2). I have found that the LP and NLP parts (5.8), (5.9) agree with ref. [19], but the NNLP part disagrees. The same holds for the double Boer-Mulders term, which is not presented here. The hadron tensor derived in refs. [19, 21] is transverse exactly at NNLP, while (5.2) is transverse in the sum of all powers. Possibly, it is due to the differences in the definition of twist-four terms.

### 5.1 Restoration of EM gauge invariance

The EM gauge-invariance (or electric-charge conservation) imposes

$$\partial_\mu J^\mu = 0, \quad (5.11)$$

which on the level hadronic tensor turns to

$$q_\mu W^{\mu\nu} = W^{\nu\mu} q_\mu = 0. \quad (5.12)$$

This statement is exact in QCD. Nonetheless, EM gauge-invariance is violated at each power of TMD factorization. It is a consequence of the fact that the relation (5.12) does not have a definite power counting, since it contains a sum of  $q^\pm$  and  $q_T$ . The EM gauge-invariance is restored by the power corrections. Moreover, it is restored for each terms with unique nonperturbative content by a the series of KPCs.

Let me demonstrate the restoration of EM gauge invariance using the unpolarized DY example in the momentum space.<sup>9</sup> Contracting the LP term (5.5) with  $q^\mu$ , one obtains

$$\begin{aligned}
 q_\mu W_{0,f_1 f_1}^{\mu\nu} = & \frac{-1}{N_c} \int d^2 k_{1T} d^2 k_{2T} \delta^{(2)}(q_T - k_{1T} - k_{2T}) C_0 \\
 & \times (k_{1T}^\nu + k_{2T}^\nu) f_1(x_1, k_{1T}) f_1(x_2, k_{2T}) \neq 0. \quad (5.13)
 \end{aligned}$$

So, the EM gauge-invariance is violated, but the violation goes beyond the LP approximation, because the left-hand-side(l.h.s.) of the formula is  $\sim Q^1$  while the r.h.s. is  $\sim Q^0$ . The

---

<sup>9</sup>The same exercise can be done without assumptions on  $\zeta$ 's in the position space. In this case, the integrand of  $\tilde{W}^{\mu\nu}$  should be acted by the operator  $\hat{q}^\mu = \bar{n}^\mu q^+ + n^\mu q^- - i \frac{\partial}{\partial b_\mu}$ . The results of this section are reproduced also for the terms that contain derivatives of the Collins-Soper kernel.

inclusion of the NLP term (5.6) leads to

$$q_\mu(W_{0,f_1f_1}^{\mu\nu} + W_{1,f_1f_1}^{\mu\nu}) = \frac{1}{N_c} \int d^2k_{1T} d^2k_{2T} \delta^{(2)}(q_T - k_{1T} - k_{2T}) C_0 \quad (5.14)$$

$$\times \left[ \left( \frac{\bar{n}^\nu(k_{2T})_T}{q^-} + \frac{n^\nu(k_{1T})_T}{q^+} \right) f_1(x_1, k_{1T}) f_1(x_2, k_{2T}) \right] \neq 0.$$

The r.h.s. expression (5.14) is of order  $\sim Q^{-1}$ , i.e. NNLP in comparison to the l.h.s. The LP violation term (5.13) is canceled by the NLP violation term, but leaves an NNLP remnant. The twist-three contributions are not presented here, but it could be checked using expression in refs. [21–23] that they also result into  $\sim Q^{-1}$  terms. The NNLP remnant in eq. (5.14) is then canceled by NNLP (5.7) term:

$$q_\mu(W_{0,f_1f_1}^{\mu\nu} + W_{1,f_1f_1}^{\mu\nu} + W_{2,f_1f_1}^{\mu\nu}) = \quad (5.15)$$

$$\frac{1}{N_c} \int d^2k_{1T} d^2k_{2T} \delta^{(2)}(q_T - k_{1T} - k_{2T}) C_0$$

$$\times \left[ \frac{k_{1T}^\nu k_{2T}^2 + k_{2T}^\nu k_{1T}^2}{q^+ q^-} - \frac{q_T^\nu}{2q^+ q^-} \left( x_2 k_{1T}^2 \frac{\partial}{\partial x_2} + x_1 k_{2T}^2 \frac{\partial}{\partial x_1} \right) \right] f_1(x_1, k_{1T}) f_1(x_2, k_{2T}) \neq 0.$$

Here, the r.h.s. is of order  $\sim Q^{-2}$ , i.e. N<sup>3</sup>LP in comparison to l.h.s. This N<sup>3</sup>LP violation term is to be canceled by the corresponding part of  $q_\mu W_3^{\mu\nu}$ , and so on.

In this way, EM gauge invariance is restored only in the sum of all powers of TMD factorization, and violated at any fixed power. A truncation of the series at power  $n$  results into the violation of EM gauge invariance at order  $1/Q^{n+1}$ . This mechanism is standard for the factorization theorems with inhomogeneous counting of  $q^\mu$  components. Another very well studied example is DVCS, see refs. [40, 41, 64, 65] for detailed discussion.

Importantly, the restoration of EM gauge invariance takes place independently for the terms proportional to twist-two TMDPDFs, and twist-three TMDPDFs. This is due to the proper definition of the TMD-twist. With the present definition TMD distributions of different twists do not mix with each other, and are entirely independent nonperturbative functions. Each independent combination of TMDPDFs forms its own series of KPCs that restores the EM gauge invariance.

The complete series (5.2) is exactly EM-gauge invariant. The check is straightforward, one should only take into account that  $q^\pm$  turns to  $xp$  by  $\delta$ -function and does not commute with the derivative. All terms of expression (5.2) participate in the restoration, which is in contrast to the small- $x$ -based derivation made in ref. [21], where NNLP term already completes the gauge-invariant sequence. The summed expression (4.17) is also EM-gauge invariant since  $q^\mu = k_1^\mu + k_2^\mu$  and

$$q^\mu ((k_1 k_2) g^{\mu\nu} - k_1^\mu k_2^\nu - k_2^\mu k_1^\nu) = -k_1^2 k_2^\nu - k_2^2 k_1^\nu = 0, \quad (5.16)$$

because  $k_{1,2}^2 = 0$  due to the  $\delta$ -functions. The part  $\sim t^{\mu\nu}$  in (4.17) is also transverse.

The cancellation between successive terms takes place only if their hard coefficient functions are identical. This has been checked up in section 3.3 explicitly at NLO, but the gauge-invariance guaranties the equivalence at all perturbative orders. The EM-gauge invariance does not determine the coefficient function of other genuine contributions, since they have independent nonperturbative content.

## 5.2 Restoration of frame invariance

Another important symmetry that is violated by the factorization approach is the Lorentz invariance. In the factorization theorems approach it is often called the frame-invariance or the reparameterization invariance. It has been intensively studied for the case of collinear factorization for DVCS, see refs. [40, 66], and also in the soft-collinear/heavy-quark effective fields theories, see refs. [67–69]. Meanwhile, I do not know any dedicated discussion for the TMD factorization case.

The frame- or reparameterization-invariance is based on the fact that the field-mode separation depends on the directions  $n$  and  $\bar{n}$ , which, in turn, are defined only approximately. In other words, the counting rules (3.7) can be modified by power corrections, if they preserve the leading counting. The light-cone vector  $n^\mu$ , which defines the collinear direction, could be turned  $n^\mu \rightarrow n'^\mu$  by a power-suppressed amount such that  $n'^2 = 0$ , and the factorization theorem remains the same. If there are several collinear directions the transformation should preserve the relation between them. This is an obvious statement in the deep-inelastic scattering. In other cases the frame-invariance is rather involved, and usually is violated by the LP term and restored by power corrections similarly to EM gauge invariance.

Let me inspect the implication of frame-invariance for the TMD factorization approach. There are two collinear directions given by  $n^\mu$  and  $\bar{n}^\mu$ , which satisfy  $(n\bar{n}) = 1$ . It is straightforward to see that there are two possible transformations  $n \rightarrow n'$ , that preserve  $n'^2 = \bar{n}'^2 = 0$  and  $(n'\bar{n}') = 1$ . They are

$$\text{I: } \begin{cases} n^\mu \rightarrow n'^\mu = n^\mu + \frac{\Delta^\mu}{q^-} - \frac{\Delta^2}{2(q^-)^2} \bar{n}^\mu, \\ \bar{n}^\mu \rightarrow \bar{n}'^\mu = \bar{n}^\mu, \end{cases} \quad \text{II: } \begin{cases} n^\mu \rightarrow n'^\mu = n^\mu, \\ \bar{n}^\mu \rightarrow \bar{n}'^\mu = \bar{n}^\mu + \frac{\bar{\Delta}^\mu}{q^+} - \frac{\bar{\Delta}^2}{2(q^+)^2} n^\mu, \end{cases} \quad (5.17)$$

where  $\Delta$  and  $\bar{\Delta}$  are transverse vectors in the original frame, i.e.  $(\Delta n) = (\bar{\Delta} \bar{n}) = 0$ . The vectors have the counting  $\Delta \sim Q^0$ , and could be treated as infinitesimal parameters. Note, that one cannot transform  $n$  and  $\bar{n}$  simultaneously, because in this case, the equation  $(n\bar{n}) = 1$  requires  $\Delta \sim Q^1$ , which is not a small transformation. Nonetheless, transformations I and II can be applied successively, simulating a simultaneous rotation of  $n$  and  $\bar{n}$ . It is important to mention that the vectors  $p_1^\mu$  and  $p_2^\mu$  are external vectors, and thus they are not transformed under (5.17). The definition (3.4) is not modified by  $\Delta$ .

The transformations I and II are symmetric to each other. Therefore, the result valid for one of them is automatically valid for another. Thus, in the rest of the section, I consider only the transformation I.

On the level of factorized expression transformations (5.17) change the definition of TMDPDFs, since  $(k_T n') \neq 0$ . Therefore, for the comparison of momentum space integrands (similar analysis can be done for position space integrands), one should compensate the redefinition of  $k$ 's, such that the definition of  $f_1(x, k_T)$  remains the same. Recalling that TMDPDF is obtained by Fourier transforms (2.21) one finds that the same definitions can be achieved by the contra-rotation

$$k_{1T}^\mu \rightarrow k'_{1T}^\mu = k_{1T}^\mu, \quad k_{2T}^\mu \rightarrow k'_{2T}^\mu = k_{2T}^\mu - \Delta^\mu - ((k_2 \Delta) - \Delta^2) \frac{\bar{n}^\mu}{q^-}. \quad (5.18)$$

Here,  $(k_{1T}\Delta) = 0$  is imposed, in order to preserve the definition of TMDPDF  $f_1(x_1, k_{1T})$ . The variables  $x$  are also impacted by the rotation because they are defined via  $q^\pm$ . One finds that

$$x_1 \rightarrow x'_1 = x_1 \left( 1 + \frac{2(k_2\Delta) - \Delta^2}{2q^+q^-} \right), \quad x_2 \rightarrow x'_2 = x_2, \quad (5.19)$$

where it is used that  $(q\Delta) = (k_2\Delta)$ , due to the delta-function. So, the simultaneous transformations (5.17) (part I), (5.18) and (5.19) should keep the integrand of the factorized hadronic tensor invariant.

Applying the transformation to the LP term (5.5) one finds

$$W_{0,f_1f_1}^{\mu\nu} \rightarrow W_{0,f_1f_1}^{\mu\nu} + \frac{1}{N_c} \int d^2k_{1T} d^2k_{2T} \delta^{(2)}(q_T - k_{1T} - k_{2T}) C_0 \left\{ \frac{\bar{n}^\mu \Delta^\nu + \Delta^\mu n^\nu}{q^-} + \dots \right\} f_1(x_1, k_{1T}) f_1(x_2, k_{2T}), \quad (5.20)$$

where the dots denote the higher power terms. The first term in brackets is  $\sim Q^{-1}$  i.e. NLP. So, the LP term is frame-invariant up to NLP corrections. Adding the NLP contribution (5.6) one obtains

$$W_{0,f_1f_1}^{\mu\nu} + W_{1,f_1f_1}^{\mu\nu} \rightarrow W_{0,f_1f_1}^{\mu\nu} + W_{1,f_1f_1}^{\mu\nu} + \frac{1}{N_c} \int d^2k_{1T} d^2k_{2T} \delta^{(2)}(q_T - k_{1T} - k_{2T}) C_0 \left\{ \left( \frac{\Delta^\mu k_{1T}^\nu + k_{1T}^\mu \Delta^\nu}{q^+q^-} + \frac{\bar{n}^\mu \bar{n}^\nu (\Delta^2 - 2(k_2\Delta))}{(q^-)^2} \right) f_1 f_1 - x_1 g_T^{\mu\nu} \frac{2(k_2\Delta) - \Delta^2}{2q^+q^-} \frac{\partial f_1}{\partial x_1} f_1 + \dots \right\}. \quad (5.21)$$

The arguments of TMDPDFs are omitted but the relative order of TMDPDFs is preserved. Again the violation term starts with  $\sim Q^{-2}$  and is NNLP. The derivative term is obtained from the expansion of  $x'_1$ . Finally, one can check that

$$W_0^{\mu\nu} + W_1^{\mu\nu} + W_2^{\mu\nu} \rightarrow W_0^{\mu\nu} + W_1^{\mu\nu} + W_2^{\mu\nu} + \mathcal{O}(N^3\text{LP}), \quad (5.22)$$

where  $\mathcal{O}(N^3\text{LP})$  is a long expression. Importantly, that the frame-invariance involves the derivative-of-TMDPDF terms, which do not participate in the check of gauge-invariance at NNLP order. Therefore, the series of KPCs cannot be split into sub-series that are gauge and frame invariant. Alike the gauge-invariance, the frame-invariance is not complete at any given power-order. Moreover, at each power order it generates the infinite series of derivative corrections, due to the Taylor expansion of  $f_1(x'_1, k_{1T})$ , which altogether restore the frame-invariance of the full series. The frame invariance holds for each nonperturbative sector, i.e. separately for the power series involving twist-two TMDPDFs, separately for the power series involving twist-three TMDPDFs, etc. The expression with summed KPCs (5.3) is invariant under the transformations (5.17) since it does not depend vectors  $n$  and  $\bar{n}$ . Here, I recall that the external momenta  $p_1^\mu, p_2^\mu$  are not transformed, and thus  $\delta$ -functions with arguments  $(\xi p^\pm - k^\pm)$  are not modified.

### 5.3 Estimation of numerical importance

The factorization theorem with summed KPCs is valid in the traditional regime of TMD phenomenology,  $Q \gg q_T$  and  $Q \gg \Lambda$  but does not have restriction  $Q \gg k_T$ . In fact, the

restriction  $Q \gg k_T$  was ignored in all phenomenological studies, assuming that  $k_T \sim \Lambda$ . The assumption  $k_T \sim \Lambda$  is not entirely correct because  $k_T$  is an integration variable. In this section, I test the impact of inclusion of KPCs. For it, I compare LP and KPC-summed cross-sections of the unpolarized Drell-Yan (only  $\gamma$ -channel for simplicity).

The cross-section for the Drell-Yan reaction is computed from the hadronic tensor by

$$\frac{d\sigma}{dQ^2 dy dq_T^2} = \frac{2\alpha_{\text{em}}}{sQ^4} \sum_q e_q^2 \int d\Omega L_{\mu\nu} W^{\mu\nu}, \quad (5.23)$$

where  $\mathbf{q}_T^2 = -q_T^2 > 0$ ,  $\alpha_{\text{em}}$  is the QED coupling, and  $e_q$  is the charge of the quark  $q$ . The leptonic tensor is

$$L^{\mu\nu} = 4(l^\mu l'^\nu + l'^\mu l^\nu - g^{\mu\nu}(ll')), \quad (5.24)$$

where  $l$  and  $l'$  are leptons' momenta and  $l + l' = q$ .

Substituting the hadronic tensor from eq. (5.5) one obtains the cross-section in the LP TMD factorization

$$\begin{aligned} \left. \frac{d\sigma}{dQ^2 dy dq_T^2} \right|_{\text{LP}} &= \frac{(2\pi)^2 \alpha_{\text{em}}}{3N_c s Q^2} \left( 1 + \frac{\mathbf{q}_T^2}{2Q^2} \right) \sum_q e_q^2 C_0 \left( \frac{2q^+ q^-}{\mu^2} \right) \\ &\times \int d^2 \mathbf{k}_{1T} d^2 \mathbf{k}_{2T} \delta^{(2)}(\mathbf{q}_T - \mathbf{k}_{1T} - \mathbf{k}_{2T}) f_{1q}(x_1, k_{1T}; \mu, \zeta) f_{1\bar{q}}(x_2, k_{2T}; \mu, \bar{\zeta}), \end{aligned} \quad (5.25)$$

where

$$x_1 = \frac{q^+}{p_1^+} = \frac{Qe^y}{\sqrt{s}} \sqrt{1 + \frac{\mathbf{q}_T^2}{Q^2}}, \quad x_2 = \frac{q^-}{p_2^-} = \frac{Qe^{-y}}{\sqrt{s}} \sqrt{1 + \frac{\mathbf{q}_T^2}{Q^2}}. \quad (5.26)$$

The correction  $\mathbf{q}_T^2/2Q^2$  in the common factor is the result of convolution of the leptonic tensor with  $g_T^{\mu\nu}$ . The inhomogeneity of this factor in power-counting is due fact that the leptonic tensor is exact (i.e. it contains all powers) while the hadronic tensor is pure LP. Exactly this expression is usually used for the phenomenology of unpolarized TMD distributions, see e.g. [70–72].

The integral over transverse momenta  $k_{1T}$  and  $k_{2T}$  is convenient to present as the integral over  $\mathbf{k}_{1T}^2$  and  $\mathbf{k}_{2T}^2$ . The delta-function restricts the values of these variables as

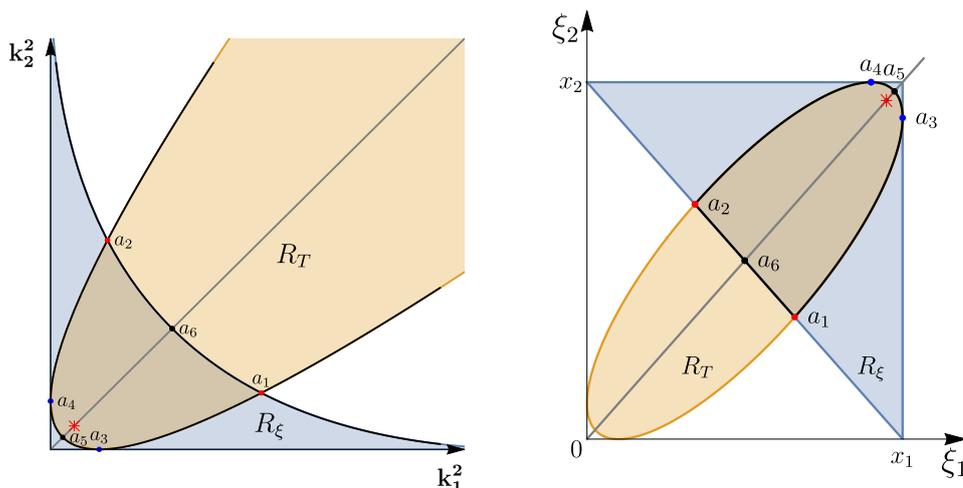
$$R_T : \quad \mathbf{k}_{1T}^2 > 0, \quad \mathbf{k}_{1T}^2 + \mathbf{q}_T^2 - \sqrt{\mathbf{k}_{qT}^2 \mathbf{q}_T^2} < \mathbf{k}_{2T}^2 < \mathbf{k}_{1T}^2 + \mathbf{q}_T^2 + \sqrt{\mathbf{k}_{qT}^2 \mathbf{q}_T^2}. \quad (5.27)$$

The region  $R_T$  is shown in the figure 2(left) by yellow color. Notably, the region spans to infinite values of  $\mathbf{k}_T^2$ 's, despite initial assumption of  $Q$  being the largest scale.

The cross-section with resummed KPC is obtained from eq. (4.17) and reads

$$\begin{aligned} \left. \frac{d\sigma}{dQ^2 dy dq_T^2} \right|_{\text{KPC}} &= \frac{(2\pi)^2 \alpha_{\text{em}}}{3N_c s} \sum_q e_q^2 C_0 \left( \frac{Q^2}{\mu^2} \right) \int d^4 k_1 d^4 k_2 \int d\xi_1 d\xi_2 \delta^{(4)}(q - k_1 - k_2) \\ &\delta(k_1^2) \delta(k_2^2) \delta(k_1^+ - \xi_1 p_1^+) \delta(k_2^- - \xi_2 p_2^-) f_{1q}(\xi_1, k_{1T}; \mu, Q^2) f_{1\bar{q}}(\xi_2, k_{2T}; \mu, Q^2). \end{aligned} \quad (5.28)$$

It is straightforward to check that in the limit  $q^\pm \rightarrow \infty$  the expression (5.28) reproduces (5.25) (up to the power-suppressed term in the common factor). Note that to make this comparison



**Figure 2.** (left) The region of integration for variables  $k_{1T}^2$  and  $k_{2T}^2$  in the TMD factorization formulas (5.25) and (5.28). (right) The region of values of parameters  $\xi_{1,2}$  (5.26) covered during the integration over  $k_{1T}^2$  and  $k_{2T}^2$ . The points  $a_i$  at both plots corresponds to the same values of  $k_{1T}^2$  and  $k_{2T}^2$ .

one should commute the limit  $Q \rightarrow \infty$  and the integration operation. In other words, one should assume that  $k_T^2 \ll Q^2$ , despite the integral range of  $k_T^2$ 's is infinite.

The values collinear momentum-fractions of TMDPDFs are not fixed in the summed formula (5.28), but integrated in a particular range. It could be anticipated apriori, because the reparametrization invariance modifies the value of  $x$ 's (5.19), and thus any fixed choice is not frame-invariant. The  $\delta$ -functions express  $\xi$ 's as

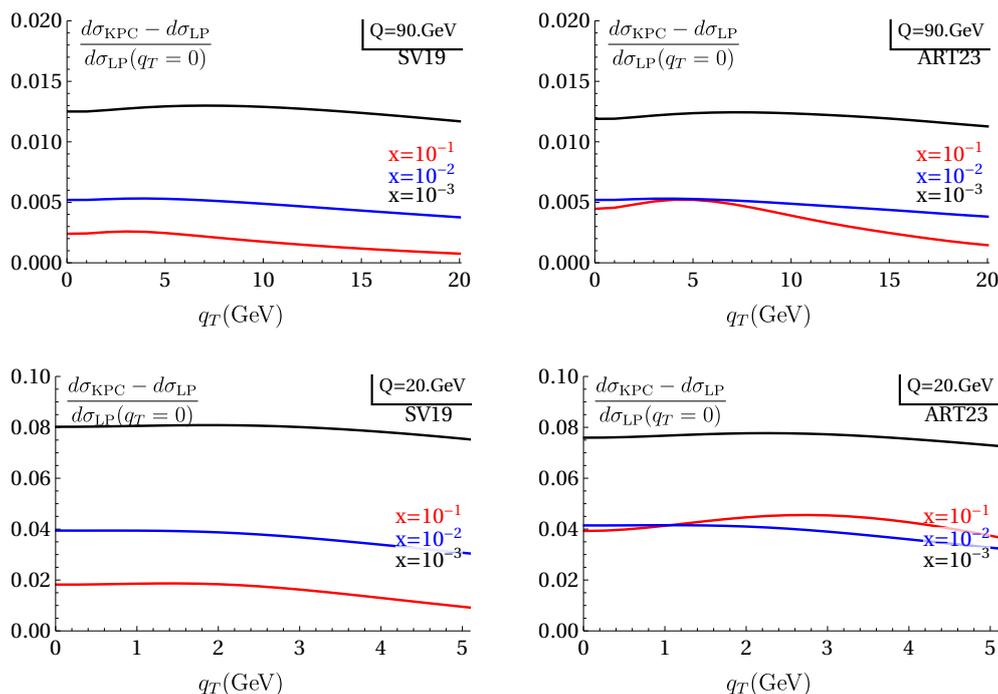
$$\begin{aligned} \xi_1 &= \frac{x_1}{2} \left( 1 + \frac{k_{1T}^2}{\tau^2} - \frac{k_{2T}^2}{\tau^2} + \frac{\sqrt{\lambda(k_{1T}^2, k_{2T}^2, \tau^2)}}{\tau^2} \right), \\ \xi_2 &= \frac{x_2}{2} \left( 1 - \frac{k_{1T}^2}{\tau^2} + \frac{k_{2T}^2}{\tau^2} + \frac{\sqrt{\lambda(k_{1T}^2, k_{2T}^2, \tau^2)}}{\tau^2} \right), \end{aligned} \quad (5.29)$$

where  $\tau^2 = 2q^+q^- = Q^2 + \mathbf{q}_T^2$ , and  $\lambda(a, b, c) = a^2 + b^2 + c^2 - 2ab - 2ac - 2bc$  is the kinematic function. The values of  $\xi_{1,2}$  are restricted  $0 < \xi_{1,2} < 1$  (due to the support properties of TMDPDF), which constraints the integration region of  $k_{1T}^2$  and  $k_{2T}^2$  to

$$R_\xi : \quad 0 < k_{1T}^2 < \tau^2, \quad 0 < k_{2T}^2 < \tau + k_{1T}^2 - 2\sqrt{k_{1T}^2\tau^2}. \quad (5.30)$$

The region  $R_\xi$  is shown in the figure 2(left) by blue color. The variables  $\xi$  (5.29) can be seen as a kind-of-Nachmann variables [73, 74] for the TMD factorization theorem, which correct for the convolution integral for non-zero transverse momenta of partons.

In full the integration region for (5.28) is  $R_\xi \cap R_T$ . It is compact. The maximum values of  $k_T^2$  are  $(\sqrt{\tau^2} + \sqrt{\mathbf{q}_T^2})^2/4$  ( $= Q^2/4$  at  $\mathbf{q}_T^2 = 0$ ) which are reached at the points  $a_{1,2}$  in figure 2. The arguments  $\xi_{1,2}$  belongs to the region shown in figure 2(right). This region



**Figure 3.** The difference of KPC-summed and LP cross-sections relative to the  $q_T = 0$  LP cross-section for the Drell-Yan reaction at different  $q_T$  and different  $x$ . The plot in left(right) panels are based on SV19 [71] (ART23 [72]) extraction of TMDPDF. Upper row is for  $Q = 90\text{GeV}$ , and lower row is for  $Q = 20\text{GeV}$ . Lines of different color corresponds to different values of  $x$  (at  $y = 0$ ).

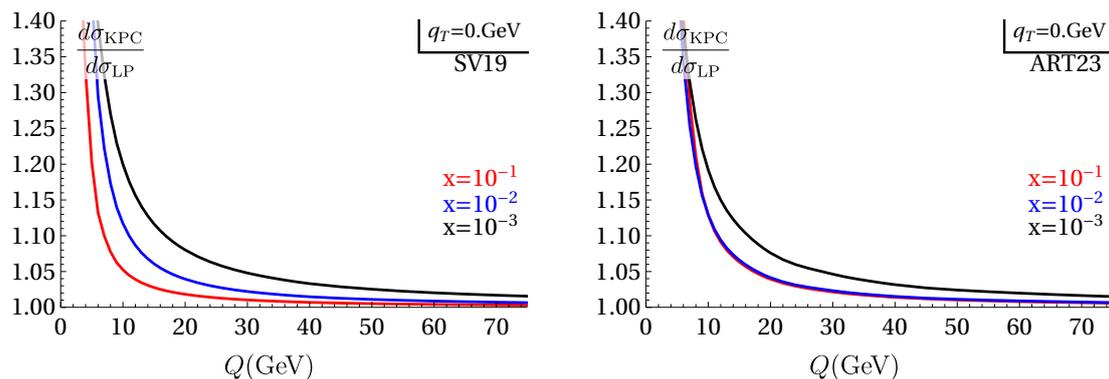
does not include the LP values of collinear momentum-fraction  $x_{1,2}$  (5.26), but it includes the values  $\sqrt{Q^2}/se^{\pm y}$ , which are often used in the phenomenology and is shown by the red star in figure 2.

In figures 3 and 4, the comparison of KPC-summed (5.28) and LP (5.25) cross-sections is shown. For the comparison the extractions SV19 [71] (presented in the left panels) and ART23 [72] (presented in the right panels) were used. The perturbative orders are the same as used in the extractions. They are  $N^3\text{LL}$  for SV19, and  $N^4\text{LL}$  for ART23. The evolution of distributions is performed in the position space, then the TMDPDFs are Fourier-transformed and substituted into (5.28), (5.25). The results for different extraction have generally the same size but somewhat different behavior at  $x > 10^{-2}$ .

The inclusion of KPCs results into an almost flat increase of the cross-section as a function of  $q_T$ , see figure 3. The size of the shift grows at smaller  $x$ 's and  $Q$ . In other words, the inclusion of KPCs does not significantly modify the shape of the LP prediction, but significantly changes the normalization. At the typical kinematic of the LHC (aka Z-boson production) the size of corrections is of the order of 1%. It grows to  $\sim 30 - 40\%$  at  $Q \sim 10\text{GeV}$ . Using the presented curves one can estimate the effective size of KPCs. Their order of magnitude is approximately described by

$$\frac{d\sigma|_{\text{KPC}}}{d\sigma|_{\text{LP}}} \sim 1 + \frac{(2\text{GeV})^2}{x^{0.4}Q^2}. \quad (5.31)$$

This approximation is obtained by a two-parameter fit of SV19 and ART23 curves.



**Figure 4.** The ratio of KPC-summed to LP cross-sections for the Drell-Yan reaction at  $q_T = 0$  at different  $Q$  and different  $x$ . The plot in left(right) panel is based on SV19 [71] (ART23 [72]) extraction of TMDPDF. Lines of different color corresponds to different values of  $x$  (at  $y = 0$ ).

It is interesting to mention that the modern TMD phenomenology faces problems with the description of normalization of cross-section, while perfectly describes the shape. For LHC energies the issue is of order of few percents, while for low-energy fix-target data ( $Q \sim 5 - 20\text{GeV}$ ) is tenths of percents [3, 75, 76]. The situation is worse for  $Q \sim 2 - 4\text{GeV}$  which are typical for Semi-Inclusive Deep-Inelastic Scattering (SIDIS). In this case the problem with normalization is of order of factor 2-3 [5]. Therefore, one could hope that inclusion of KPCs into the phenomenology will resolve this important problem.

## 6 Conclusion

The power corrections in the TMD factorization theorem provide a rich field for investigation. As described in the introduction, there are four conceptual types of power corrections:  $q_T/Q$ ,  $k_T/Q$ ,  $\Lambda/Q$ , and target-mass corrections. These corrections have distinct origins and characteristics. Importantly, different power corrections are significant in different kinematic regimes, which justifies considering them independently. In this work, I have derived the series of kinematic power corrections (KPCs) or  $k_T/Q$ -corrections that follow the leading power (LP) term. These corrections play a special role for the factorization theorem and are essential for its consistency because KPCs are responsible for restoring electromagnetic (EM) gauge invariance (charge conservation) and frame invariance, which are broken by the LP approximation. In that sense, the series of KPCs is the entailed part of the LP factorization theorem.

The KPCs that follow the LP term possess a distinctive mark: they only involve TMD distributions of twist-two. This feature allows for their straightforward extraction from the generic power expansion of operators. The resulting series is presented in equation (3.23). This series is summed in a simple expression (4.17), which is the main result of the work. The computation is done for the Drell-Yan process but can be generalized for other processes without conceptual problems.

One of the most significant features of KPCs is that they must obey the factorization theorem, even if other types of power corrections may violate it. This statement follows

from the fact that KPCs are responsible for restoring the fundamental properties of the LP term, namely charge conservation and frame invariance. Consequently, the coefficient function for KPCs remains identical to the LP coefficient function. In section 3.3, I have explicitly verified this statement at the next-to-leading order (NLO) and demonstrated the restoration of the argument of the LP coefficient function to  $q^2$ . The latter is a non-trivial result as it indicates the impossibility of a strict power expansion beyond NLP.

The final formula for the cross-section is almost trivial — it tells that the cross-section can be computed with free massless quarks similar to a naive parton picture. However, with the present derivation, it attains a different status. It obeys the factorization theorem, and it is clear which part of the power expansion is included and which is neglected. The excluded terms consist of TMD distributions of twist-three and higher, as well as powers of  $q_T/Q$ .

It is important to note that the derivation of the series of KPCs is contingent upon the definition of TMD-twist and higher twist distributions. The current derivation is based on the approach outlined in ref. [22], which leads to a consistent result. However, if an alternative definition is proposed, the series of KPCs could differ. In such a case, some of the present higher twist terms would be absorbed into the KPC series. In particular, the difference in the treatment of higher twist terms may describe the discrepancy between the present computation and ref. [19].

Incorporating KPCs into TMD phenomenology is essential, as they play a crucial role in restoring the consistency of the formalism. Moreover, all perturbative ingredients remain unchanged since KPCs adhere to the LP factorization theorem. Therefore, including KPCs in phenomenological studies can be done without encountering conceptual difficulties while maintaining the achieved level of accuracy (currently at the  $N^4LL$  order [72]). The summed formula is valid in the same kinematic range of application as ordinary TMD factorization. Estimations made in section 5.3 demonstrate that including KPCs results in an almost constant increment of the cross-section. The magnitude of this correction depends on  $Q$  and  $x$ . For typical LHC kinematics, the correction is around 1%, while at  $Q \sim 4 - 5$  GeV, the correction can reach 100%. Interestingly, the deficiency in normalization for the TMD factorization at low energies has been reported by multiple groups. One could expect that these problems will be resolved with the inclusion of KPCs.

## Acknowledgments

I thank Vladimir Braun, Ian Balitky, Oscar del Rio, Simone Rodini, and Ignazio Scimemi for numerous discussions and helpful comments. I also thank Sara Alvarez for critical reading and corrections to the manuscript. This work is funded by the *Atracción de Talento Investigador* program of the Comunidad de Madrid (Spain) No. 2020-T1/TIC-20204, and is also supported by the Spanish Ministry grant PID2019-106080GB-C21.

**Open Access.** This article is distributed under the terms of the Creative Commons Attribution License ([CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/)), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

## References

- [1] R. Angeles-Martinez et al., *Transverse Momentum Dependent (TMD) parton distribution functions: status and prospects*, *Acta Phys. Polon. B* **46** (2015) 2501 [[arXiv:1507.05267](#)] [[INSPIRE](#)].
- [2] R. Boussarie et al., *TMD Handbook*, [arXiv:2304.03302](#) [[INSPIRE](#)].
- [3] A. Vladimirov, *Pion-induced Drell-Yan processes within TMD factorization*, *JHEP* **10** (2019) 090 [[arXiv:1907.10356](#)] [[INSPIRE](#)].
- [4] M. Bury, A. Prokudin and A. Vladimirov, *Extraction of the Sivers function from SIDIS, Drell-Yan, and  $W^\pm/Z$  boson production data with TMD evolution*, *JHEP* **05** (2021) 151 [[arXiv:2103.03270](#)] [[INSPIRE](#)].
- [5] MAP (MULTI-DIMENSIONAL ANALYSES OF PARTONIC DISTRIBUTIONS) collaboration, *Unpolarized transverse momentum distributions from a global fit of Drell-Yan and semi-inclusive deep-inelastic scattering data*, *JHEP* **10** (2022) 127 [[arXiv:2206.07598](#)] [[INSPIRE](#)].
- [6] R. Abdul Khalek et al., *Science Requirements and Detector Concepts for the Electron-Ion Collider: EIC Yellow Report*, *Nucl. Phys. A* **1026** (2022) 122447 [[arXiv:2103.05419](#)] [[INSPIRE](#)].
- [7] A. Accardi et al., *Strong Interaction Physics at the Luminosity Frontier with 22 GeV Electrons at Jefferson Lab*, [arXiv:2306.09360](#) [[INSPIRE](#)].
- [8] J.C. Collins, D.E. Soper and G.F. Sterman, *Transverse Momentum Distribution in Drell-Yan Pair and  $W$  and  $Z$  Boson Production*, *Nucl. Phys. B* **250** (1985) 199 [[INSPIRE](#)].
- [9] J. Collins, *Foundations of perturbative QCD*, Cambridge University Press (2013) [[DOI:10.1017/9781009401845](#)] [[INSPIRE](#)].
- [10] T. Becher and M. Neubert, *Drell-Yan Production at Small  $q_T$ , Transverse Parton Distributions and the Collinear Anomaly*, *Eur. Phys. J. C* **71** (2011) 1665 [[arXiv:1007.4005](#)] [[INSPIRE](#)].
- [11] M.G. Echevarria, 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 [[arXiv:1111.4996](#)] [[INSPIRE](#)].
- [12] S.M. Aybat and T.C. Rogers, *TMD Parton Distribution and Fragmentation Functions with QCD Evolution*, *Phys. Rev. D* **83** (2011) 114042 [[arXiv:1101.5057](#)] [[INSPIRE](#)].
- [13] 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 [[arXiv:1202.0814](#)] [[INSPIRE](#)].
- [14] I. Scimemi and A. Vladimirov, *Systematic analysis of double-scale evolution*, *JHEP* **08** (2018) 003 [[arXiv:1803.11089](#)] [[INSPIRE](#)].
- [15] P.J. Mulders and R.D. Tangerman, *The complete tree level result up to order  $1/Q$  for polarized deep inelastic leptonproduction*, *Nucl. Phys. B* **461** (1996) 197 [[hep-ph/9510301](#)] [[INSPIRE](#)].
- [16] A. Bacchetta et al., *Semi-inclusive deep inelastic scattering at small transverse momentum*, *JHEP* **02** (2007) 093 [[hep-ph/0611265](#)] [[INSPIRE](#)].
- [17] S. Arnold, A. Metz and M. Schlegel, *Dilepton production from polarized hadron hadron collisions*, *Phys. Rev. D* **79** (2009) 034005 [[arXiv:0809.2262](#)] [[INSPIRE](#)].

- [18] I. Balitsky and A. Tarasov, *Power corrections to TMD factorization for Z-boson production*, *JHEP* **05** (2018) 150 [[arXiv:1712.09389](#)] [[INSPIRE](#)].
- [19] I. Balitsky, *Gauge-invariant TMD factorization for Drell-Yan hadronic tensor at small x*, *JHEP* **05** (2021) 046 [[arXiv:2012.01588](#)] [[INSPIRE](#)].
- [20] M. Inglis-Whalen, M. Luke, J. Roy and A. Spourdalakis, *Factorization of power corrections in the Drell-Yan process in EFT*, *Phys. Rev. D* **104** (2021) 076018 [[arXiv:2105.09277](#)] [[INSPIRE](#)].
- [21] I. Balitsky, *Drell-Yan angular lepton distributions at small x from TMD factorization*, *JHEP* **09** (2021) 022 [[arXiv:2105.13391](#)] [[INSPIRE](#)].
- [22] A. Vladimirov, V. Moos and I. Scimemi, *Transverse momentum dependent operator expansion at next-to-leading power*, *JHEP* **01** (2022) 110 [[arXiv:2109.09771](#)] [[INSPIRE](#)].
- [23] M.A. Ebert, A. Gao and I.W. Stewart, *Factorization for azimuthal asymmetries in SIDIS at next-to-leading power*, *JHEP* **06** (2022) 007 [*Erratum ibid.* **07** (2023) 096] [[arXiv:2112.07680](#)] [[INSPIRE](#)].
- [24] S. Rodini and A. Vladimirov, *Factorization for quasi-TMD distributions of sub-leading power*, *JHEP* **09** (2023) 117 [[arXiv:2211.04494](#)] [[INSPIRE](#)].
- [25] S. Rodini and A. Vladimirov, *Transverse momentum dependent factorization for SIDIS at next-to-leading power*, [arXiv:2306.09495](#) [[INSPIRE](#)].
- [26] S. Rodini and A. Vladimirov, *Definition and evolution of transverse momentum dependent distribution of twist-three*, *JHEP* **08** (2022) 031 [*Erratum ibid.* **12** (2022) 048] [[arXiv:2204.03856](#)] [[INSPIRE](#)].
- [27] R.A. Brandt and G. Preparata, *Operator product expansions near the light cone*, *Nucl. Phys. B* **27** (1971) 541 [[INSPIRE](#)].
- [28] D.J. Gross and S.B. Treiman, *Light cone structure of current commutators in the gluon quark model*, *Phys. Rev. D* **4** (1971) 1059 [[INSPIRE](#)].
- [29] V. Moos and A. Vladimirov, *Calculation of transverse momentum dependent distributions beyond the leading power*, *JHEP* **12** (2020) 145 [[arXiv:2008.01744](#)] [[INSPIRE](#)].
- [30] W. Rarita and J. Schwinger, *On a theory of particles with half integral spin*, *Phys. Rev.* **60** (1941) 61 [[INSPIRE](#)].
- [31] E.G. Delgado Acosta, V.M. Banda Guzmán and M. Kirchbach, *Bosonic and fermionic Weinberg-Joos  $(j, 0) \oplus (0, j)$  states of arbitrary spins as Lorentz tensors or tensor-spinors and second-order theory*, *Eur. Phys. J. A* **51** (2015) 35 [[arXiv:1503.07230](#)] [[INSPIRE](#)].
- [32] B. Geyer, M. Lazar and D. Robaschik, *Decomposition of nonlocal light cone operators into harmonic operators of definite twist*, *Nucl. Phys. B* **559** (1999) 339 [[hep-th/9901090](#)] [[INSPIRE](#)].
- [33] R.L. Jaffe and X.-D. Ji, *Chiral odd parton distributions and polarized Drell-Yan*, *Phys. Rev. Lett.* **67** (1991) 552 [[INSPIRE](#)].
- [34] A.V. Belitsky, X. Ji and F. Yuan, *Final state interactions and gauge invariant parton distributions*, *Nucl. Phys. B* **656** (2003) 165 [[hep-ph/0208038](#)] [[INSPIRE](#)].
- [35] R.L. Jaffe, *Spin, twist and hadron structure in deep inelastic processes*, in the proceedings of the *Ettore Majorana International School of Nucleon Structure: 1st Course: The Spin Structure of the Nucleon*, Erice, Italy, August 03–10 (1995), p. 42–129 [[hep-ph/9602236](#)] [[INSPIRE](#)].

- [36] I.I. Balitsky and V.M. Braun, *Evolution Equations for QCD String Operators*, *Nucl. Phys. B* **311** (1989) 541 [[INSPIRE](#)].
- [37] A.V. Belitsky and D. Mueller, *Twist- three effects in two photon processes*, *Nucl. Phys. B* **589** (2000) 611 [[hep-ph/0007031](#)] [[INSPIRE](#)].
- [38] L. Gamberg et al., *Transverse-momentum-dependent factorization at next-to-leading power*, [arXiv:2211.13209](#) [[INSPIRE](#)].
- [39] S. Wandzura and F. Wilczek, *Sum Rules for Spin Dependent Electroproduction: Test of Relativistic Constituent Quarks*, *Phys. Lett. B* **72** (1977) 195 [[INSPIRE](#)].
- [40] V.M. Braun and A.N. Manashov, *Operator product expansion in QCD in off-forward kinematics: Separation of kinematic and dynamical contributions*, *JHEP* **01** (2012) 085 [[arXiv:1111.6765](#)] [[INSPIRE](#)].
- [41] V.M. Braun and A.N. Manashov, *Kinematic power corrections in off-forward hard reactions*, *Phys. Rev. Lett.* **107** (2011) 202001 [[arXiv:1108.2394](#)] [[INSPIRE](#)].
- [42] J.C. Collins and D.E. Soper, *Back-To-Back Jets in QCD*, *Nucl. Phys. B* **193** (1981) 381 [*Erratum ibid.* **213** (1983) 545] [[INSPIRE](#)].
- [43] A.A. Vladimirov, *Self-contained definition of the Collins-Soper kernel*, *Phys. Rev. Lett.* **125** (2020) 192002 [[arXiv:2003.02288](#)] [[INSPIRE](#)].
- [44] M.A. Ebert and F.J. Tackmann, *Resummation of Transverse Momentum Distributions in Distribution Space*, *JHEP* **02** (2017) 110 [[arXiv:1611.08610](#)] [[INSPIRE](#)].
- [45] D. Boer, P.J. Mulders and F. Pijlman, *Universality of T odd effects in single spin and azimuthal asymmetries*, *Nucl. Phys. B* **667** (2003) 201 [[hep-ph/0303034](#)] [[INSPIRE](#)].
- [46] S. Bhattacharya et al., *Exclusive double quarkonium production and generalized TMDs of gluons*, *Phys. Lett. B* **833** (2022) 137383 [[arXiv:1802.10550](#)] [[INSPIRE](#)].
- [47] M.G. Echevarria, P.A. Gutierrez Garcia and I. Scimemi, *GTMDs and the factorization of exclusive double Drell-Yan*, *Phys. Lett. B* **840** (2023) 137881 [[arXiv:2208.00021](#)] [[INSPIRE](#)].
- [48] V.M. Braun, A.N. Manashov, D. Mueller and B. Pirnay, *Resolving kinematic ambiguities in QCD predictions for Deeply Virtual Compton Scattering*, *PoS DIS2014* (2014) 225 [[arXiv:1407.0815](#)] [[INSPIRE](#)].
- [49] A.V. Manohar and I.W. Stewart, *The Zero-Bin and Mode Factorization in Quantum Field Theory*, *Phys. Rev. D* **76** (2007) 074002 [[hep-ph/0605001](#)] [[INSPIRE](#)].
- [50] A. Vladimirov, *Structure of rapidity divergences in multi-parton scattering soft factors*, *JHEP* **04** (2018) 045 [[arXiv:1707.07606](#)] [[INSPIRE](#)].
- [51] I. Balitsky, *Rapidity-only TMD factorization at one loop*, *JHEP* **03** (2023) 029 [[arXiv:2301.01717](#)] [[INSPIRE](#)].
- [52] V.M. Braun, A.N. Manashov and J. Rohrwild, *Renormalization of Twist-Four Operators in QCD*, *Nucl. Phys. B* **826** (2010) 235 [[arXiv:0908.1684](#)] [[INSPIRE](#)].
- [53] A.V. Manohar, *Deep inelastic scattering as  $x \rightarrow 1$  using soft collinear effective theory*, *Phys. Rev. D* **68** (2003) 114019 [[hep-ph/0309176](#)] [[INSPIRE](#)].
- [54] R.N. Lee et al., *Quark and Gluon Form Factors in Four-Loop QCD*, *Phys. Rev. Lett.* **128** (2022) 212002 [[arXiv:2202.04660](#)] [[INSPIRE](#)].
- [55] I. Moulton, H.X. Zhu and Y.J. Zhu, *The four loop QCD rapidity anomalous dimension*, *JHEP* **08** (2022) 280 [[arXiv:2205.02249](#)] [[INSPIRE](#)].

- [56] I. Scimemi, A. Tarasov and A. Vladimirov, *Collinear matching for Sivers function at next-to-leading order*, *JHEP* **05** (2019) 125 [[arXiv:1901.04519](#)] [[INSPIRE](#)].
- [57] V.M. Braun, Y. Ji and A. Vladimirov, *QCD factorization for twist-three axial-vector parton quasidistributions*, *JHEP* **05** (2021) 086 [[arXiv:2103.12105](#)] [[INSPIRE](#)].
- [58] V. Shtabovenko, R. Mertig and F. Orellana, *FeynCalc 9.3: New features and improvements*, *Comput. Phys. Commun.* **256** (2020) 107478 [[arXiv:2001.04407](#)] [[INSPIRE](#)].
- [59] J. Collins, *CSS Equation, etc, Follow from Structure of TMD Factorization*, [arXiv:1212.5974](#) [[INSPIRE](#)].
- [60] M.G. Echevarría, A. Idilbi and I. Scimemi, *Soft and Collinear Factorization and Transverse Momentum Dependent Parton Distribution Functions*, *Phys. Lett. B* **726** (2013) 795 [[arXiv:1211.1947](#)] [[INSPIRE](#)].
- [61] M. Beneke, T. Feldmann and D. Seidel, *Systematic approach to exclusive  $B \rightarrow Vl^+l^-$ ,  $V\gamma$  decays*, *Nucl. Phys. B* **612** (2001) 25 [[hep-ph/0106067](#)] [[INSPIRE](#)].
- [62] M. Beneke and M. Neubert, *QCD factorization for  $B \rightarrow PP$  and  $B \rightarrow PV$  decays*, *Nucl. Phys. B* **675** (2003) 333 [[hep-ph/0308039](#)] [[INSPIRE](#)].
- [63] I.I. Balitsky, V.M. Braun and A.V. Kolesnichenko, *Power corrections  $1/Q^2$  to parton sum rules for deep inelastic scattering from polarized targets*, *Phys. Lett. B* **242** (1990) 245 [*Erratum ibid.* **318** (1993) 648] [[hep-ph/9310316](#)] [[INSPIRE](#)].
- [64] A.V. Belitsky and D. Mueller, *Resummation of target mass corrections in two photon processes: Twist two sector*, *Phys. Lett. B* **507** (2001) 173 [[hep-ph/0102224](#)] [[INSPIRE](#)].
- [65] A.V. Belitsky and D. Mueller, *Exclusive electroproduction revisited: treating kinematical effects*, *Phys. Rev. D* **82** (2010) 074010 [[arXiv:1005.5209](#)] [[INSPIRE](#)].
- [66] V.M. Braun, A.N. Manashov, D. Müller and B.M. Pirnay, *Deeply Virtual Compton Scattering to the twist-four accuracy: Impact of finite- $t$  and target mass corrections*, *Phys. Rev. D* **89** (2014) 074022 [[arXiv:1401.7621](#)] [[INSPIRE](#)].
- [67] M.E. Luke and A.V. Manohar, *Reparameterization invariance constraints on heavy particle effective field theories*, *Phys. Lett. B* **286** (1992) 348 [[hep-ph/9205228](#)] [[INSPIRE](#)].
- [68] A.V. Manohar, T. Mehen, D. Pirjol and I.W. Stewart, *Reparameterization invariance for collinear operators*, *Phys. Lett. B* **539** (2002) 59 [[hep-ph/0204229](#)] [[INSPIRE](#)].
- [69] C. Marcantonini and I.W. Stewart, *Reparameterization Invariant Collinear Operators*, *Phys. Rev. D* **79** (2009) 065028 [[arXiv:0809.1093](#)] [[INSPIRE](#)].
- [70] I. Scimemi and A. Vladimirov, *Analysis of vector boson production within TMD factorization*, *Eur. Phys. J. C* **78** (2018) 89 [[arXiv:1706.01473](#)] [[INSPIRE](#)].
- [71] I. Scimemi and A. Vladimirov, *Non-perturbative structure of semi-inclusive deep-inelastic and Drell-Yan scattering at small transverse momentum*, *JHEP* **06** (2020) 137 [[arXiv:1912.06532](#)] [[INSPIRE](#)].
- [72] V. Moos, I. Scimemi, A. Vladimirov and P. Zurita, *Extraction of unpolarized transverse momentum distributions from fit of Drell-Yan data at  $N^4LL$* , [arXiv:2305.07473](#) [[INSPIRE](#)].
- [73] O. Nachtmann, *Positivity constraints for anomalous dimensions*, *Nucl. Phys. B* **63** (1973) 237 [[INSPIRE](#)].
- [74] H. Georgi and H.D. Politzer, *Freedom at Moderate Energies: Masses in Color Dynamics*, *Phys. Rev. D* **14** (1976) 1829 [[INSPIRE](#)].

- [75] V. Bertone, I. Scimemi and A. Vladimirov, *Extraction of unpolarized quark transverse momentum dependent parton distributions from Drell-Yan/Z-boson production*, *JHEP* **06** (2019) 028 [[arXiv:1902.08474](#)] [[INSPIRE](#)].
- [76] A. Bacchetta et al., *Transverse-momentum-dependent parton distributions up to  $N^3LL$  from Drell-Yan data*, *JHEP* **07** (2020) 117 [[arXiv:1912.07550](#)] [[INSPIRE](#)].