Self-Contained Definition of the Collins-Soper Kernel

Alexey A. Vladimirov[®]

Institut für Theoretische Physik, Universität Regensburg, D-93040 Regensburg, Germany

(Received 26 March 2020; revised 5 June 2020; accepted 9 October 2020; published 5 November 2020)

The rapidity anomalous dimension (RAD), or Collins-Soper kernel, defines the scaling properties of transverse momentum dependent distributions and can be extracted from the experimental data. I derive a self-contained nonperturbative definition that represents RAD without reference to a particular process. This definition makes possible exploration of the properties of RAD by theoretical methods on one side, and the properties of QCD vacuum with collider measurements on another side. To demonstrate these possibilities, I compute the power correction to RAD, its large-b asymptotic, and compare these estimations with recent phenomenological extractions.

DOI: 10.1103/PhysRevLett.125.192002

Introduction.—The nontrivial structure of the QCD vacuum raises a lot of fundamental and yet unsolved problems, such as mechanisms of quark confinement and hadronization. As a matter of fact, there is a little number of experimental observables that test properties of the QCD vacuum. In this Letter, I demonstrate that the evolution kernel of transverse momentum dependent (TMD) distributions is exclusively sensitive to the structure of QCD vacuum and thus is a valuable tool to study it.

The rapidity anomalous dimension (RAD), or Collins-Soper kernel, was introduced in Refs. [1,2], as a part of the factorized formula, which accumulates the doublelogarithm contributions. In later works, it has been shown that RAD is universal for different processes and receives nonperturbative (NP) corrections. The rigorous formulation of the TMD factorization theorem [3-6] has identified RAD as an independent NP function that contains the information about soft-gluon exchanges between partons and dictates the evolution properties of TMD distributions. Until recently, the NP terms of TMD evolutions were not examined individually, but as a constituent of the resummation exponent or TMD distributions. Although it does not necessarily contradict the theory, it makes it difficult to split effects related to different sides of strong dynamics. One of the main messages of this Letter is that RAD is an important function with a rich physical background, and thus must be seen as an independent distribution.

Despite the long history of RAD, very little is known about its NP nature from the theory side. Apart from a general identification that the NP part exists, I know only a

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP³.

few works that are dedicated to this problem at least partially [7–12]. In this Letter I would like to draw attention to this gap in the theory. As an initial step, I provide a fieldtheoretical and model-independent definition of RAD detached from the cross-section formula. Given the definition, RAD can be used as a self-contained phenomenological function of QCD, which measures properties of QCD vacuum. To demonstrate the power of the derived definition, I compute the leading terms of operator product expansion (OPE) and compute RAD within a simplistic model.

The appearance of RAD.—The cross section for the Drell-Yan pair production at small transverse momentum is described by the TMD factorization formula

$$\begin{split} \frac{d\sigma}{dq_T} &= \sigma_0(Q/\mu) \int d^2b e^{-i\mathbf{b}\cdot\mathbf{q}_T} \\ &\times \left(\frac{Q^2}{\zeta}\right)^{-2\mathcal{D}(b,\mu)} F_1(x_1,b;\mu,\zeta) F_2(x_2,b;\mu,\zeta), \end{split} \tag{1}$$

where Q is the virtuality of a photon, and q_T its transverse momentum [Scalar products of traverse vectors in bold font are defined as $(\mathbf{bq}_T) = -(bq_T)$. Consequently, $\mathbf{b}^2 = -b^2 > 0$.]. The functions F are TMD distributions, and \mathcal{D} is RAD [There is no common notation for RAD. The other popular notations are -K/2 [1,3,13], $-F_{q\bar{q}}$ [4].). Similar formulas describe the small-transverse momentum regime of semi-inclusive deep inelastic scattering (SIDIS), and $e^+e^- \rightarrow h_1h_2 + X$ process. In Eq. (1) and in the following, I omit the flavor indices for brevity, keeping in mind that RAD depends on the color representation of the parton.

A distinctive feature of TMD distributions is their dependence on two scales [13,14]: the factorization scale μ and the scale of rapidities separation ζ . The dependence on these scales is given by a pair of equations, where the first one is an ordinary renomalization group equation for the scale μ and the second one is

$$\frac{dF(x,b;\mu,\zeta)}{d\ln\zeta} = -\mathcal{D}(b,\mu)F(x,b;\mu,\zeta). \tag{2}$$

The integrability condition for the pair of evolution equations gives the dependence of RAD on μ ,

$$\frac{d\mathcal{D}(b,\mu)}{d\ln\mu} = \Gamma_{\text{cusp}}(\mu),\tag{3}$$

where Γ_{cusp} is the anomalous dimension for the cusp of lightlike Wilson lines. In a conformal field theory, RAD is equivalent to the soft anomalous dimension and entirely perturbative [6,15]. In QCD RAD is a general NP function, although it still inherits some properties of an anomalous dimension, such as additive structure of renormalization group equation, see Eq. (3).

Equation (2) essentially mixes the definitions of two NP functions: a TMD distribution and RAD. For that reason, the separation of these functions with the data is a nontrivial phenomenological task. Nonetheless, it could be done observing that RAD governs the Q behavior of the cross section, whereas F's govern the x behavior. Therefore, analyzing a global set of data with a large span in x and Q, it is possible to decorrelate these functions. Such global studies were made recently [16-20]. The values of RAD obtained in these works are shown in Fig. 1. Clearly, there is no agreement between these extractions for $b > 2 \text{ GeV}^{-1}$. Another observation is that extraction based on the joined data of Drell-Yan and SIDIS cross sections [16,19] provide a higher value of RAD at $b \sim 1 \text{ GeV}^{-1}$ in comparison to extraction based only on the Drell-Yan data [17,20]. These contradictions could be resolved by adding more low- q_T data in the analysis, or by some alternative approaches to access RAD. One of promising approaches is

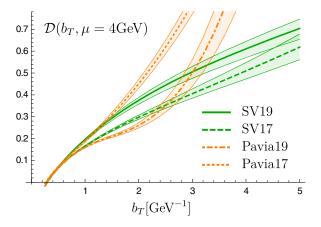


FIG. 1. Comparison of extracted values of RAD. The lines labeled as SV19, SV17, Pavia19, and Pavia17 correspond to Refs. [19,17,20,16].

the recently proposed method to compute RAD with lattice QCD [21–23].

Definition of RAD.—To derive the self-contained expression for RAD, I take a step backward in the derivation of Eq. (1) and recall the origin of scale ζ . At an intermediate stage, the expression for the cross section has the form $d\sigma \sim \tilde{F}_1 \times S \times \tilde{F}_2$ [3,5], where \tilde{F} are unsubtracted TMD distributions, and S is the TMD soft factor. Each of these terms contains the rapidity divergence(s) that cancel in the product. To obtain (1), the soft factor is factorized into parts with only rapidity divergences related to a particular lightlike direction. Afterwards, they are combined with \tilde{F} into physical TMD distributions [6,24,25]. The scale ζ in the definition of a physical TMD distribution (2) is the scale of rapidity divergence factorization. Thus, the soft factor is the primary object to define RAD.

The TMD soft factor is defined as

$$S_C(b,\mu) = \frac{\text{Tr}}{N_c} \langle 0|W_C|0\rangle Z_S^2(\mu), \tag{4}$$

where $W_C = P \exp(ig \int_C dx^\mu A_\mu(x))$ is a gauge link along the contour C (see Fig. 2), Z_S is the renormalization factor for lightlike cusps. In Ref. [6] it has been proven that the TMD soft factor with a properly designed regularization has the general form

$$S_C(b,\mu) = \exp\left[2\mathcal{D}(b,\mu)\ln(\varrho) + B(b,\mu) + \cdots\right], \quad (5)$$

where ϱ is the Lorenz-invariant combination of parameters of rapidity divergence regularization($\varrho \to 0$). The function B is the finite part of the soft factor, and the dots denote terms vanishing at $\varrho \to 0$. Consequently, RAD can be obtained from the TMD soft factor as

$$\mathcal{D}(b,\mu) = \frac{1}{2} \lim_{\rho \to 0} \frac{d \ln S_C(b,\mu)}{d \ln \rho}.$$
 (6)

The expression (5) is a general one, but it is difficult to use outside of the perturbation theory. The main complication is the definition of an appropriate rapidity divergence regulator. To guarantee Eq. (5) and make use of Eq. (6), the regulator must be given on the level of the operator, preserve the gauge invariance, and fully regularize rapidity divergences without generation of extra infrared divergences. None of the commonly used regulators in perturbative calculations (see, e.g., Refs. [3–5,26–28]) fulfill these requirements entirely. The discussion of the drawbacks in common regularizations can be found in Refs. [6,26,29]. All these requirements can be fulfilled by a deformation of the contour C such that it does not touch lightlike infinities [6]. The most straightforward deformation is the contour C_{Λ} shown in Fig. 2. In this case, the parameters Λ_+ regularize rapidity divergences at both infinities and $\varrho = (\Lambda_+ \Lambda_-)^{-1}$.

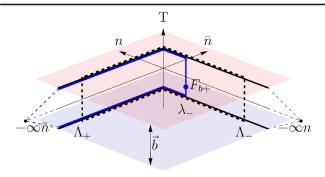


FIG. 2. Contours defining the TMD soft factor (in the Drell-Yan kinematics) and its derivatives. Axes n and \bar{n} are lightlike $(n^2 = \bar{n}^2 = 0)$, and the axis T is transverse. The black (blue) solid line shows contour C(C'). The black dashed lines show the contour C_{Λ} . The blue dot shows the insertion of gluon strength tensor

The regularized soft factor $S_{C_{\Lambda}}$ is a function of ϱ and b^2 (and μ^2), because these are the only nonzero scalar products in the task. The regularization is removed by limits $\Lambda_+ \to \infty$ and $\Lambda_- \to \infty$, but since the dependence on Λ 's is given by a single variable ϱ , one of these limits is obsolete. For definiteness, I fix $\Lambda_- = \lambda_-$. The derivative with respect to $\varrho = (\Lambda_+ \lambda_-)^{-1}$ can be replaced by derivative over λ_- , and Eq. (6) turns into

$$\mathcal{D}(b,\mu) = \frac{1}{2} \lim_{\Lambda_{+} \to \infty} \frac{d \ln S_{C_{\Lambda}}(b,\mu)}{d \ln \lambda_{-}}.$$
 (7)

The action of the derivative is

$$\mathcal{D}(b,\mu) = Z_{\mathcal{D}}(\mu) + \lim_{\Lambda_{+} \to \infty} \lambda_{-} \frac{ig \operatorname{Tr} \int_{0}^{1} d\beta \langle 0| F_{b+}(-\lambda_{-}n + b\beta) W_{C_{\Lambda}} | 0 \rangle}{\operatorname{Tr} \langle 0| W_{C_{\Lambda}} | 0 \rangle},$$

$$(8)$$

where $F_{b+}(x)=b^{\mu}n^{\nu}F_{\mu\nu}(x)$, with $F_{\mu\nu}$ being a gluon-field strength tensor, and $Z_{\mathcal{D}}(\mu)=d\ln Z_S/d\ln\lambda_-$. The contour in the numerator starts and ends at the point $(-\lambda_-n+b\beta)$, so the numerator is the Wilson loop with insertion of the gluon strength tensor. The limit $\Lambda_+\to\infty$ turns the contour C_Λ (with finite λ_-) to the contour C' shown in Fig. 2 in blue, and

$$\mathcal{D}(b,\mu) = \lambda_{-} \frac{ig}{2} \frac{\text{Tr} \int_{0}^{1} d\beta \langle 0| F_{b+}(-\lambda_{-}n + b\beta) W_{C'}|0\rangle}{\text{Tr}\langle 0| W_{C'}|0\rangle} + Z_{\mathcal{D}}(\mu).$$

$$(9)$$

Here, the numerator and the denominator have rapidity divergences, which cancel each other. So, to use Eq. (9) beyond tree order, a convenient regularization for these divergences should be introduced. The term $Z_{\mathcal{D}}(\mu) = d \ln Z_S/d \ln \lambda_-$ removes the ultraviolet divergences.

Peculiarly, it is additive rather than multiplicative, which produces the renormalization group equation of the form (3), with

$$\frac{dZ_{\mathcal{D}}(\mu)}{d\ln\mu} = \Gamma_{\rm cusp}(\mu). \tag{10}$$

The additional cusps present in C' do not introduces divergences since (nb) = 0 [30].

Despite that the left-hand side of Eq. (9) has an explicit entry of λ_- , the expression is independent on it. It is an outcome of the dependence of $S_{C_{\Lambda}}$ on ϱ . Alternatively, the λ_- independence can be seen as a consequence of the boost invariance. Different values of λ_- can be related by a boost in the n direction. Therefore, λ_- occurs in the numerator and the denominator of Eq. (9) only due to the rapidity divergences and cancel in the ratio. The independence on λ_- also demonstrates the universality of RAD for Drell-Yan and SIDIS processes, which is dictated by the sign of λ_- in the current context.

The expression (9) is the main result of this Letter. In contrast to previous works, the definition (9) gives a direct access to RAD. In the next paragraphs, I demonstrate possible applications of it and make elementary checks.

OPE and perturbative computation.—RAD is very well studied in the perturbation theory, where it has been derived up to next-to-next-to-leading order (NNLO) [15,28]. All previous calculations have been done by evaluation of the TMD soft factor [26,28], or TMD distributions [13,29], with successive identification of rapidity divergent terms. Using Eq. (9) RAD can be computed directly.

The perturbative calculation is made in the regime $b \ll \Lambda_{\rm OCD}^{-1}$. In this regime, RAD can be written as

$$\mathcal{D}(b,\mu) = \mathcal{D}_0(b,\mu) + \mathbf{b}^2 \mathcal{D}_2(b) + (\mathbf{b}^2)^2 \mathcal{D}_4(b) + \cdots, (11)$$

where dots designate terms accompanied by a higher power of \mathbf{b}^2 . Each \mathcal{D}_n depends on b only logarithmically, via $\ln(b\mu)$. Importantly, the definition (9) is made for a finite b. The limit $b \to 0$ does not exist due to the presence of divergent renormalization constant $Z_{\mathcal{D}}$ that is independent on b. Indeed, already at LO $\mathcal{D}_0 \sim \alpha_s(\mu) \ln(b\mu)$. The terms with n > 0 do not depend on μ explicitly, as it follows from the independence of $Z_{\mathcal{D}}$ on b.

The computation of \mathcal{D}_n can be done, for example, by the background field method, similarly to calculations made in Refs. [31,32]. It is convenient to use the background field in the Schwinger gauge with a reference point at the origin. With this choice, Wilson lines of background gluons turn to unities at $b \to 0$, which crucially simplifies the calculation.

The LO contribution to term \mathcal{D}_0 is given by a one-loop diagram. The result of computation in the dimensional regularization ($d=4-2\epsilon$ with $\epsilon>0$, and $\overline{\text{MS}}$ -scheme) reads

$$\mathcal{D}_0(b,\mu) = -2C_F a_s \left[\Gamma(-\epsilon) \left(\frac{\mathbf{b}^2 \mu^2}{4e^{-\gamma_E}} \right)^{\epsilon} + \frac{1}{\epsilon} \right] + \mathcal{O}(a_s^2), \quad (12)$$

where $a_s = g^2/(4\pi)^2$, and the value of Z_S is taken from Ref. [29]. This expression coincides with the one derived in [26] at arbitrary ϵ , and in the limit $\epsilon \to 0$ reproduces the well-known result [1,3–5]

$$\mathcal{D}_0(b,\mu) = 2C_F a_s(\mu) \ln\left(\frac{\mathbf{b}^2 \mu^2}{4e^{-2\gamma_E}}\right) + \mathcal{O}(a_s^2). \quad (13)$$

Note, that there is no dependence on λ_{-} , as expected.

In contrast to \mathcal{D}_0 , power suppressed terms have treeorder contributions, which are the coefficients of the Taylor series at b=0. Each term introduces a new NP function which are matrix elements of gluon strength tensors connected by Wilson lines to the origin. The LO contribution to \mathcal{D}_2 is given by

$$-g^{2}\lambda_{-}\int_{0}^{1}d\beta\int_{-\infty}^{0}d\sigma\langle0|F_{b+}(-\lambda_{-}n)[..]F_{b-}(\sigma\bar{n})|0\rangle, \quad (14)$$

where $F_{b-}=b^{\mu}n^{\nu}F_{\mu\nu}$, and [..] stays for a gauge link (in the adjoint representation) between F's and the origin (see Fig. 3). It is a particular case of the following matrix element

$$\Phi_{\mu\nu}(x,y) = g^2 x^{\alpha} y^{\beta} \langle 0|F_{\mu\alpha}(x)[x,0][0,y]F_{\nu\beta}(y)|0\rangle.$$
 (15)

The matrix element $\Phi_{\mu\nu}$ satisfies $x^{\mu}\Phi_{\mu\nu}=y^{\nu}\Phi_{\mu\nu}=0$ and thus can be parametrized by two independent functions

$$\Phi_{\mu\nu}(x,y) = \left(g_{\mu\nu} - \frac{y_{\mu}x_{\nu}}{(xy)}\right) \varphi_1(r^2, x^2, y^2)
+ \frac{(x_{\mu}(xy) - y_{\mu}x^2)(y_{\nu}(xy) - x_{\nu}y^2)}{(xy)((xy)^2 - x^2y^2)} \varphi_2(r^2, x^2, y^2), \quad (16)$$

where $r^2 = (x - y)^2$. At $x^2 = y^2 = 0$, φ_2 vanishes and only φ_1 contributes to Eq. (14), but at higher orders of perturbative series both terms are present.

Using the parametrization (16) I receive

$$\mathcal{D}_2(b) = \frac{1}{2} \int_0^\infty d\mathbf{r}^2 \frac{\varphi_1(\mathbf{r}^2, 0, 0)}{\mathbf{r}^2} + \mathcal{O}(a_s). \tag{17}$$

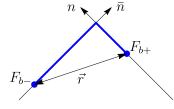


FIG. 3. Structure of the operator that describes the leading power correction to RAD. Blue lines are the gauge links, and dots are insertions of gluon strength tensors.

For the first time, the power correction to RAD is expressed in a model-independent way in terms of QCD vacuum correlations. It could be compared to expressions derived in Refs. [8,10], which involves the jet-algorithm-modified QCD vacuum.

The function φ_1 is unknown, nonetheless, its value could be estimated. In particular, at $\mathbf{r}^2 \to 0$ it is

$$\lim_{\mathbf{r}^2 \to 0} \frac{\varphi_1(\mathbf{r}^2, x^2, y^2)}{\mathbf{r}^2} = \frac{\pi^2}{36} G_2, \tag{18}$$

where $G_2 = (g^2/4\pi^2)\langle 0|: F_{\mu\nu}^a F_{\mu\nu}^a: |0\rangle$ is the gluon condensate [33]. At large ${\bf r}^2$, φ_1 decays at least as ${\bf r}^{-2}$ (more realistically, it decays exponentially). Assuming φ_1 has an effective radius $\sim \Lambda_{\rm OCD}^{-1}$, \mathcal{D}_2 can be estimated as

$$\mathcal{D}_2 \sim \frac{\pi^2}{72} \frac{G_2}{\Lambda_{\text{OCD}}^2} \simeq (1.-5.) \times 10^{-2} \text{ GeV}^2.$$
 (19)

The values of parameters are taken from the review in Ref. [34]. The estimation (19) is notably a small number. Nonetheless it is in agreement with recent extractions that are collected in the following table (see also Fig. 1).

These values are obtained with LO approximation at $\mu=2$ GeV. Let me note that earlier considerations, which are often used in high-energy phenomenology, such as BLNY-fit [35,36], use significantly higher values, $\mathcal{D}_2 \sim 0.2$ –0.35 GeV².

Example of NP modeling: Stochastic vacuum model.— One of the most promising applications of the expression (9) is the computation of RAD with various NP models. It would help to select appropriate phenomenological ansatz and give an intuitive interpretation for RAD. As an example, I evaluated \mathcal{D} in the stochastic vacuum model (SVM) [37]. Although this model cannot be considered realistic, it catches some global features of QCD, such as the area law.

In SVM one assumes that the QCD dynamics is dominated by two-point correlators, whereas multipoint correlators give a negligible contribution. Additionally, one ignores the gauge links connecting fields, assuming their unimportance at large distances. In this way, all gluonic observables are written in terms of two functions Δ and Δ_1 , defined as [37]

$$g^{2}\langle 0|F_{\mu\nu}(x)F_{\alpha\beta}(0)|0\rangle$$

$$= (g_{\mu\alpha}g_{\nu\beta} - g_{\mu\beta}g_{\nu\alpha})(\Delta(x^{2}) + \Delta_{1}(x^{2})) + (g_{\mu\alpha}x_{\nu}x_{\beta}$$

$$- g_{\nu\alpha}x_{\mu}x_{\beta} - g_{\mu\beta}x_{\nu}x_{\alpha} + g_{\nu\beta}x_{\mu}x_{\alpha})\frac{\partial\Delta_{1}(x^{2})}{\partial x^{2}}.$$
(20)

Applying the non-Abelian Stockes theorem to Eq. (9), dropping multipoint correlators, and using Eq. (20), after some simplification I arrive at the expression

$$\mathcal{D}(b) = \mathbf{b}^2 \int_{\mathbf{b}^2}^{\infty} d\mathbf{y}^2 \left(\Delta(\mathbf{y}^2) + \frac{\Delta_1(\mathbf{y}^2)}{2} \right) + \int_0^{\mathbf{b}^2} d\mathbf{y}^2 \left[\frac{\mathbf{y}^2}{2} \Delta_1(\mathbf{y}^2) + (2\sqrt{\mathbf{b}^2 \mathbf{y}^2} - \mathbf{y}^2) \Delta(\mathbf{y}^2) \right].$$
(21)

This expression predicts the linear behavior of \mathcal{D} at large b,

$$\lim_{\mathbf{b}^2 \to \infty} \mathcal{D}(b) = \sqrt{\mathbf{b}^2} \int_0^\infty d\mathbf{y}^2 2\sqrt{\mathbf{y}^2} \Delta(\mathbf{y}^2). \tag{22}$$

A similar calculation but in the momentum space has been done in Ref. [7]. Although the final results could not be compared, some intermediate steps and Eq. (22) are in agreement. The integral can be roughly estimated using lattice computations [38–40] as $c_{\infty} \simeq 0.01$ –0.4 GeV. The value can be compared to $c_{\infty} = 0.06 \pm 0.01$ GeV extracted in Ref. [19], which uses the model for $\mathcal D$ with the linear asymptotic.

Considering various relations derived in SVM (in particular, the static interquark potential [41]), I found that for the internal consistency of the model one has to demand

$$\lim_{\mathbf{b}^2 \to \infty} \mathcal{D}(b) \sim (\mathbf{b}^2)^{1/2 - \delta}, \qquad \delta \ge 0.$$
 (23)

It significantly restricts the shape of \mathcal{D} at large b. In fact, the expression (23) disregards almost all models for RAD used in phenomenology, since the dominant part of studies (e.g., Refs. [16,17,35,36]) use quadratic asymptotic $\mathcal{D} \sim \mathbf{b}^2$ or even stronger [20]. The same or equivalent conclusion as Eq. (23) has been also made in Refs. [7,11].

Conclusion.—The expression (9) is the main result of this Letter. This expression is unique in several aspects. It is a definition of a scaling kernel through the matrix elements. It grants the opportunity to study RAD without referring to TMD distributions. It gives a connection between the vacuum structure and the particle scattering. Each of these aspects is a promising direction for further studies.

To provide an elementary check and the demonstration of definition (9) I present the LO perturbative computation at $b \to 0$ and recover the well-known expression. It is clear that for perturbative computations, which are already performed at the NNLO level [6,28], the new definition is not more advantageous than an ordinary one. The main power of the new definition is that it gives a direct operator definition of RAD. It allows us to compute power corrections and apply the nonperturbative modeling for RAD, which is not possible in a standard approach.

The derived LO power correction (17) is model independent. To my best knowledge, it is the first derivation of

this object. The expression could be systematically improved by computing higher-order terms. The LO computation predicts a small size of the power correction, which is in agreement with the most recent extractions. The model calculation, performed in SVM, put a serious restriction on the shape of RAD at large values of b (23). Altogether, these findings severely constrain the evolution properties of TMD distributions and should be accounted for in the analysis. The calculations are done for RAD of quark TMD distributions. It could be easily repeated for the gluon case. The only modification is the color representation for gauge links. Consequently, all expressions derived in the Letter are also valid for gluon RAD after the Casimir rescaling (that is valid up to N³LO), which consists of the multiplication by $C_A/C_F (= 9/4)$.

The possibility to investigate the QCD vacuum in high energy collisions sounds contradictory to the intuitive picture that the structure of accelerated particles is cleared from low-energy effects. Indeed, the partons do not interact with each other within a highly energetic hadron. Nonetheless, their temperate transverse motion is sensitive to the structure of the underlying vacuum. Therefore, measuring the low- q_T behavior of high-energy scattering at different energies, one examines the QCD vacuum. In fact, the measurements by LHC restricts RAD significantly, as it is shown in Refs. [18,20,42].

A. A. V. is thankful to V. Braun, O.Teryaev, A.Schäfer, and I.Scimemi for stimulating discussions. This work was supported by DFG (FOR 2926 "Next Generation pQCD for Hadron Structure: Preparing for the EIC," Project No. 430824754).

^{*}Alexey.Vladimirov@physik.uni-regensburg.de

^[1] J. C. Collins and D. E. Soper, Nucl. Phys. **B193**, 381 (1981);**B213**, 545 (1983).

^[2] J. C. Collins and D. E. Soper, Nucl. Phys. **B197**, 446 (1982).

^[3] J. Collins, Foundations of Perturbative QCD (Cambridge University Press, Cambridge, England, 2013), ISBN 9781107645257, 9781107645257, 9780521855334, 9781139097826, http://www.cambridge.org/de/knowledge/ isbn/item5756723.

^[4] T. Becher and M. Neubert, Eur. Phys. J. C 71, 1665 (2011).

^[5] M. G. Echevarria, A. Idilbi, and I. Scimemi, J. High Energy Phys. 07 (2012) 002.

^[6] A. Vladimirov, J. High Energy Phys. 04 (2018) 045.

^[7] S. Tafat, J. High Energy Phys. 05 (2001) 004.

^[8] C. Lee and G. F. Sterman, Phys. Rev. D 75, 014022 (2007).

^[9] P. Schweitzer, M. Strikman, and C. Weiss, J. High Energy Phys. 01 (2013) 163.

^[10] T. Becher and G. Bell, Phys. Rev. Lett. **112**, 182002 (2014).

^[11] J. Collins and T. Rogers, Phys. Rev. D 91, 074020 (2015).

^[12] I. Scimemi and A. Vladimirov, J. High Energy Phys. 03

^[13] S. M. Aybat and T. C. Rogers, Phys. Rev. D **83**, 114042 (2011).

- [14] I. Scimemi and A. Vladimirov, J. High Energy Phys. 08 (2018) 003.
- [15] A. A. Vladimirov, Phys. Rev. Lett. 118, 062001 (2017).
- [16] A. Bacchetta, F. Delcarro, C. Pisano, M. Radici, and A. Signori, J. High Energy Phys. 06 (2017) 081.
- [17] I. Scimemi and A. Vladimirov, Eur. Phys. J. C 78, 89 (2018).
- [18] V. Bertone, I. Scimemi, and A. Vladimirov, J. High Energy Phys. 06 (2019) 028.
- [19] I. Scimemi and A. Vladimirov, J. High Energy Phys. 06 (2020) 137.
- [20] A. Bacchetta, V. Bertone, C. Bissolotti, G. Bozzi, F. Delcarro, F. Piacenza, and M. Radici, J. High Energy Phys. 07 (2020) 117.
- [21] M. A. Ebert, I. W. Stewart, and Y. Zhao, Phys. Rev. D 99, 034505 (2019).
- [22] X. Ji, Y. Liu, and Y.-S. Liu, Nucl. Phys. B955, 115054 (2020).
- [23] A. A. Vladimirov and A. Schäfer, Phys. Rev. D 101, 074517 (2020).
- [24] M. G. Echevarria, A. Idilbi, and I. Scimemi, Phys. Lett. B 726, 795 (2013).
- [25] J.-Y. Chiu, A. Jain, D. Neill, and I.Z. Rothstein, J. High Energy Phys. 05 (2012) 084.
- [26] M. G. Echevarria, I. Scimemi, and A. Vladimirov, Phys. Rev. D 93, 054004 (2016).
- [27] Y. Li, D. Neill, and H. X. Zhu, arXiv:1604.00392 [Phys. Rev. D (to be published)].

- [28] Y. Li and H. X. Zhu, Phys. Rev. Lett. 118, 022004 (2017).
- [29] M. G. Echevarria, I. Scimemi, and A. Vladimirov, J. High Energy Phys. 09 (2016) 004.
- [30] G. Korchemsky and G. Marchesini, Nucl. Phys. **B406**, 225 (1993).
- [31] I. I. Balitsky and V. M. Braun, Nucl. Phys. B311, 541 (1989).
- [32] I. Scimemi, A. Tarasov, and A. Vladimirov, J. High Energy Phys. 05 (2019) 125.
- [33] M. A. Shifman, A. Vainshtein, and V. I. Zakharov, Nucl. Phys. **B147**, 385 (1979).
- [34] S. Narison, Int. J. Mod. Phys. A 33, 1850045 (2018).
- [35] F. Landry, R. Brock, P.M. Nadolsky, and C.P. Yuan, Phys. Rev. D 67, 073016 (2003).
- [36] P. Sun, J. Isaacson, C. P. Yuan, and F. Yuan, Int. J. Mod. Phys. A 33, 1841006 (2018).
- [37] A. Di Giacomo, H. G. Dosch, V. Shevchenko, and Y. Simonov, Phys. Rep. **372**, 319 (2002).
- [38] G. S. Bali, N. Brambilla, and A. Vairo, Phys. Lett. B 421, 265 (1998).
- [39] E. Meggiolaro, Phys. Lett. B 451, 414 (1999).
- [40] Y. A. Simonov, Phys. Rev. D 99, 056012 (2019).
- [41] N. Brambilla and A. Vairo, Phys. Rev. D 55, 3974 (1997).
- [42] F. Hautmann, I. Scimemi, and A. Vladimirov, Phys. Lett. B 806, 135478 (2020).