Letter

Determination of the Collins-Soper kernel from cross-sections ratios

Armando Bermudez Martinez[1,2,*] and Alexey Vladimirov[3,4,†]

[1]*CERN, 1211 Geneva 23, Switzerland*
[2]*Deutsches Elektronen-Synchrotron DESY, Notkestrasse 85, 22607 Hamburg, Germany*
[3]*Institut für Theoretische Physik, Universität Regensburg, D-93040 Regensburg, Germany*
[4]*Departamento de Física Teórica & IPARCOS, Universidad Complutense de Madrid,*
E-28040 Madrid, Spain

The Collins-Soper kernel is a powerful tool for studying the properties of the QCD vacuum and an essential component of the transverse momentum dependent (TMD) factorization theorem. In this paper, we present a novel method for determining the Collins-Soper kernel directly from the comparison of differential cross sections measured at different energies. The method relies solely on the structure of the TMD factorization theorem and thus also provides a direct test of the theorem validity. With minor modifications, the procedure can be applied to the real measured data for Drell-Yan and SIDIS processes. As a demonstration, we analyze the pseudodata generated by the CASCADE event generator and determine the Collins-Soper kernel suggested by the parton branching model.

I. INTRODUCTION

The primary goal of the modern quantum chromo-dynamics (QCD) is to understand the inner structure of the nucleon and the forces that bind its constituents together. The confinement mechanism prevents any direct exploration of the hadron's insides, and thus one should employ indirect approaches. The main of these approaches is the analysis of the differential cross sections of particle scattering. The major tool for interpreting scattering cross sections is the factorization theorems [1], which are formulated in terms of universal parton distributions, each of which highlights a specific aspect of parton's dynamics. Among a variety of parton distributions the special role is played by the Collins-Soper (CS) kernel [2], which emerges from the factorization theorems for the trans-verse-momentum-differential cross sections [3–5].

Despite being a part of the factorization theorem, the CS kernel is conceptually different from other distributions. First and foremost, the CS kernel is not a characteristic of a hadron. It provides us with information about the long-range forces acting on quarks that are imposed solely by the nontrivial structure of the QCD vacuum [6]. In that sense,

the CS kernel is the most fundamental distribution in the modern framework of factorization theorems, and for that reason, there has been a continuous effort by many groups to determine the CS kernel.

At present, there are two approaches for the determination of the CS kernel. The most traditional one is the extraction of the CS kernel from the experimental data for TMD cross sections. The latest extractions are based on the global fits of Drell-Yan and Semi-Inclusive Deep-Inelastic scattering processes (SIDIS) [7–11]. Another approach uses the QCD lattice simulations. It has been suggested recently in Refs. [12–14], and the results of the first simulations are already available [15–18]. However, so far, the different extractions do not demonstrate a good agreement (see the comparison in Refs. [6,19], and also in Fig. 3), which is due to the large systematic uncertainties. The phenomenological extractions are biased by the fitting ansatz, while the lattice extractions are contaminated by large power corrections. In this paper, we propose a direct way of extracting the CS kernel from the scattering data. Our method relies on constructing proper ratios of cross sections in a way in which the hadron-dependent pieces exactly cancel leaving intact the hard and soft contributions. The method can be applied to the actual measurements of differential cross sections in Drell-Yan and SIDIS processes.

To demonstrate the power of the proposed approach, we use pseudodata generated by the CASCADE event generator [20]. The CASCADE event generator provides a very good approximation of real data for Drell-Yan transverse momentum spectra in a wide range of mass and center-of-mass energies [21–25].

---

*[*]armando.bermudez.martinez@cern.ch
[†]alexeyvl@ucm.es

II. METHOD

The method is founded on the leading power transverse momentum dependent (TMD) factorization theorem [3,4]. For the concreteness, we consider the Drell-Yan pair production process $h_1 + h_2 \to \gamma^*(\to \ell^+\ell^-) + X$. Other processes, for which the TMD factorization is formulated, can be analyzed in an analogous way. The TMD factorization theorem for the unpolarized Drell-Yan process reads [1,3–5]

$$\frac{d\sigma}{dQ^2 dy dq_T^2} = \frac{2\pi}{9}\frac{\alpha_{\rm em}^2(Q)}{sQ^2}|C_V(Q,\mu_Q)|^2 \int_0^\infty dbbJ_0(bq_T)$$
$$\times \sum_q e_q^2 f_{q,h_1}(x_1,b;\mu_Q,Q^2)$$
$$\times f_{\bar{q},h_2}(x_2,b;\mu_Q,Q^2), \qquad (1)$$

where Q, y, and q_T are the invariant mass, rapidity and transverse momentum of the virtual photon, $\mu_Q \sim Q$ is the factorization scale, s is the invariant mass squared of the initial state, e_q are the electric charges of quarks, $\alpha_{\rm em}$ is the fine-structure constant and J_0 is the Bessel function of the first kind. The variables x_1 and x_2 are longitudinal momentum fractions

$$x_1 = \frac{Q}{\sqrt{s}}e^y, \qquad x_2 = \frac{Q}{\sqrt{s}}e^{-y}. \qquad (2)$$

The hard coefficient function C_V is entirely perturbative and known up to next-to-next-to-next-to-leading order (N³LO) [26]. The functions f are nonperturbative unpolarized TMD distributions.

The CS kernel is hidden in the Q-dependence of TMD distributions that is described by a pair of evolution equations [27–29]

$$\frac{df_{q,h}(x,b;\mu,\zeta)}{d\ln\mu^2} = \frac{\gamma_F(\mu,\zeta)}{2}f_{q,h}(x,b;\mu,\zeta),$$
$$\frac{df_{q,h}(x,b;\mu,\zeta)}{d\ln\zeta} = -\mathcal{D}(b,\mu)f_{q,h}(x,b;\mu,\zeta). \qquad (3)$$

Here, γ_F is the TMD anomalous dimension which is perturbative and known up N³LO [26], and $\mathcal{D}$ is the nonperturbative CS kernel [30]. Thus, to extract the CS kernel one must explore the Q-dependence of the cross section.

The essential complication of any phenomenological analysis with the TMD factorization is that all nonperturbative functions are defined in the position space. We perform the inverse Hankel transform of the cross section

$$\Sigma(s,y,Q,b) = \int_0^\infty dq_T q_T J_0(q_T b)\frac{d\sigma}{dQ^2 dy dq_T^2}. \qquad (4)$$

The formula (1) is valid at small-q_T/Q, and the corrections to it are estimated as $\sim 1\%$ at $q_T = 0.1Q$ [10]. Consequently, Σ is accurately (up to 1%) described in the terms of TMD distributions for $b \gtrsim (0.1Q)^{-1}$.

The main idea of the method is to compare Σ's measured at different Q's (Q_1 and Q_2), such that the TMD distributions f exactly cancel in the ratio. To perform the cancellation we adjust the values of s such that the variables $x_{1,2}$ are identical. We compute

$$\frac{\Sigma(s_1,y,Q_1,b)}{\Sigma(s_2,y,Q_2,b)} = \left(\frac{Q_2}{Q_1}\right)^4 Z(Q_1,Q_2)e^{2\Delta(b;Q_1\to Q_2)}, \quad (5)$$

where $s_1/s_2 = Q_1^2/Q_2^2$. The function Z is entirely perturbative

$$Z(Q_1,Q_2) = \frac{\alpha_{\rm em}^2(Q_1)|C_V(Q_1,\mu_{Q_1})|^2}{\alpha_{\rm em}^2(Q_2)|C_V(Q_2,\mu_{Q_2})|^2}. \qquad (6)$$

The function Δ is resulted from the evolution of TMD distribution to the same scale by Eqs. (3),

$$\Delta(b,Q_1 \to Q_2) = \int_P \left(\gamma_F(\mu,\zeta)\frac{d\mu}{\mu} - \mathcal{D}(b,\mu)\frac{d\zeta}{\zeta}\right), \qquad (7)$$

where P is a path connecting points (μ_{Q_1},Q_1^2) and (μ_{Q_2},Q_2^2) in the (μ,ζ)-plane [29]. Thus, the only nonperturbative content in the formula (5) is the CS kernel.

To invert the formula (5) we use the rectangular contour for the integration path in Eq. (7) and find

$$\mathcal{D}(b,\mu_0) = \frac{\ln(\frac{\Sigma_1}{\Sigma_2}) - \ln Z(Q_1,Q_2) - 2\Delta_R(Q_1,Q_2;\mu_0)}{4\ln(Q_2/Q_1)} - 1,$$
$$(8)$$

where Σ_1/Σ_2 is shorthand notation for the ratio (5), and

$$\Delta_R(Q_1,Q_2;\mu_0) = \int_{\mu_{Q_2}}^{\mu_{Q_1}}\frac{d\mu}{\mu}\gamma_F(\mu,Q_1)$$
$$- 2\ln\left(\frac{Q_1}{Q_2}\right)\int_{\mu_0}^{\mu_{Q_2}}\frac{d\mu}{\mu}\Gamma_{\rm cusp}(\mu), \quad (9)$$

with $\Gamma_{\rm cusp}$ being the cusp anomalous dimension. The last term in Eq. (9) evolves CS kernel to the scale μ_0, which is used to compare different extractions. Apart of the ratio Σ_1/Σ_2 all terms in Eq. (8) are perturbative, and nowadays know up to N³LO. Therefore, the formula (8) can be used to extract CS kernel directly from the data without any further approximation.

Practically, the experimental measurements for differential cross sections are presented by a collection of points in bins of (Q,y,q_T). Therefore, the transformation (4) cannot be computed analytically but by the discrete Hankel

transform. Herewith, one should find a balance between the statistical precision of $d\Sigma$ (which usually decreases at low-q_T) and the range of b (larger b requires lower q_T). Alternatively, the experimental curve can be fit by an analytical form, and the transformation (4) is performed analytically. This path, however, introduces uncertainty due to the curve parametrization.

The integration over q_T leaves intact the dependence on Q and y, which can be used to increase the statistical precision. We introduce

$$\Sigma(s, Q, b) = \int_{Q-\delta Q}^{Q+\delta Q} dQ^2 \int_{-\delta y}^{\delta y} dy \, d\Sigma(s, y, Q, b), \quad (10)$$

where δQ and δy are sizes of the bin. These function can be also used in the ratio Σ_1 / Σ_2 with the only restriction that $\delta Q \ll Q$. In this case, the effects of variation of Q within the bin could be neglected. There is no limitation for δy, except that δy is the same for Σ_1 and Σ_2.

Equation (8) is valid for any kind of initial hadrons. This property can be used to cross-validate the extraction. The values of Q should be large enough to neglect the QCD power and target-mass corrections $Q \gg \Lambda_{QCD}$, and small enough to neglect the contribution of the Z-boson intermediate state $Q \ll M_Z$. It provides a very large window of available energies. The corrections for the Z-boson can be included in Eq. (8) by modifying expression for Z Eq. (6).

III. DETERMINATION OF THE CS KERNEL USING SIMULATED DATA

To demonstrate the proposed methodology, we study the pseudodata generated with the CASCADE event generator [20], which is based on the parton branching (PB) approach [31,32]. The predictions provided by CASCADE describe the Drell-Yan transverse momentum spectrum for both low-energy measurements from the PHENIX, NUSEA, R209, E605 experiments [21,22], as well as high-energy data from the LHC experiments CMS, and ATLAS [23–25]. The PB approach does not explicitly employ the TMD factorization theorem. In particular, there are no parameters specially dedicated to the CS kernel. Therefore, our determination of the CS kernel has the additional advantage of serving as a check of compatibility between the PB and the factorization approaches.

We have generated several sets of pseudodata. These sets span a wide range in phase space, namely, different Q, s and δy, as well as different hadron types. The CS kernel is independent of these parameters and thus, a priori all combinations should result in the same value. The kinematic setups for pseudodata are shown in Table I.

The inverse Hankel transform has been performed using the algorithm [33]. The algorithm expects that the input function vanishes beyond the presented domain. It is a good approximation since the cross section for the Drell-Yan process drops rapidly at large-q_T. For the considered cases,

TABLE I. Parameters of the generated events. For each case $\delta Q = 5\% \cdot Q$.

Init. state	Q [GeV]	$\sqrt{s}$ [GeV]	δy
pp	12	655.2	4
	16	873.6	
	20	1092.	
	24	1310.	
pp	12	88.7	2
	16	118.2	
pp	12	241.0	3
	16	321.4	
pp	12	1781.	5
	16	2375.	
p$\bar{p}$	12	655.2	4
	16	873.6	
	20	1092.	

the relative numerical uncertainty due to the truncation is 10^{-5}–10^{-6} and can be safely neglected.

The effective range and accuracy of the discrete Hankel transform depend on the density and range of the input cross section. So, to obtain a stable curve at the large-b, one needs a large number of points at small-q_T. In particular, we collect events into narrow q_T-bins with the size 0.05 GeV, which allows us to reach $b \sim 4$–5 GeV^{-1}. At larger values of b, the inverse function is sensitive to the finite-bin effects and becomes unstable. The examples of the cross section in momentum and position spaces are shown in Fig. 1. We employ the bootstrap method to estimate the propagation of statistical uncertainty from the momentum to the position space. During the re-sampling, we also vary the central value of q_T within a bin, which allows us to estimate the uncertainty due to finite bin size at large-b.

Using the sets listed in Table I, we compose twelve ratios Σ_1 / Σ_2. Each ratio provides an independent value of CS kernel (8). The collection of resulting curves at $\mu = 2$ GeV is shown in Fig. 2 (top). Apparently all curves are in a perfect agreement for $b > 0.4$ GeV^{-1}. It manifests that the CASCADE event generator supports the TMD factorization theorem. Note that three extractions are made with proton-antiproton collision, and they are indistinguishable from the proton-proton cases. It confirms the universality of the CS kernel.

For an extra test of the universality, we have simulated the pseudodata (pp collisions at $Q = 12$ and 16 GeV, $\delta y = 4$) using different collinear PDFs. Namely, the CT18 [34], MSHT20 [35] and NNPDF3.1 [36]. The comparison of these cases is shown in Fig. 2 (top). These curves are also in the perfect agreement, which shows that the TMD evolution within CASCADE is independent of the DGLAP evolution. For a demonstration of a possible misbehavior, we show [in the gray-dashed line in Fig. 2 (bottom)] the CS kernel extracted from the ratio with different collinear PDFs (CASCADE at $Q = 12$ GeV to
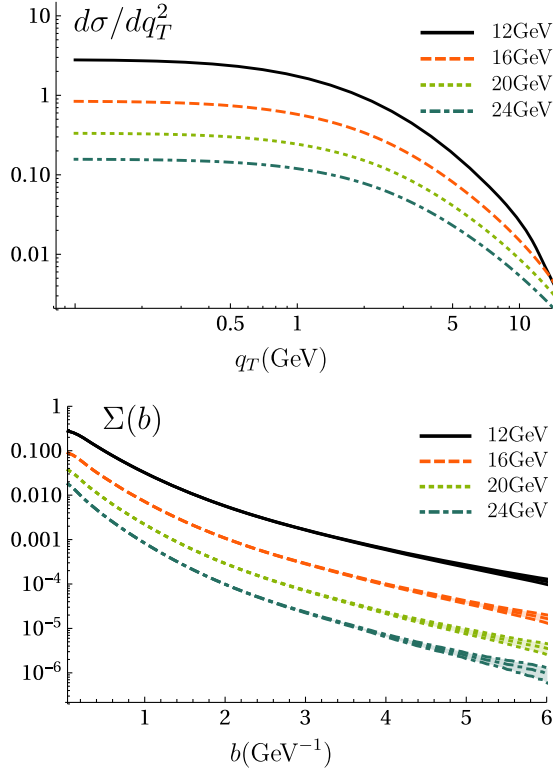
FIG. 1. Example of pseudodata in the momentum (top plot) and position (bottom plot) spaces. The statistical uncertainty is shown by the width of the line. The shown cases correspond to pp-scattering integrated in $|y| < 4$.
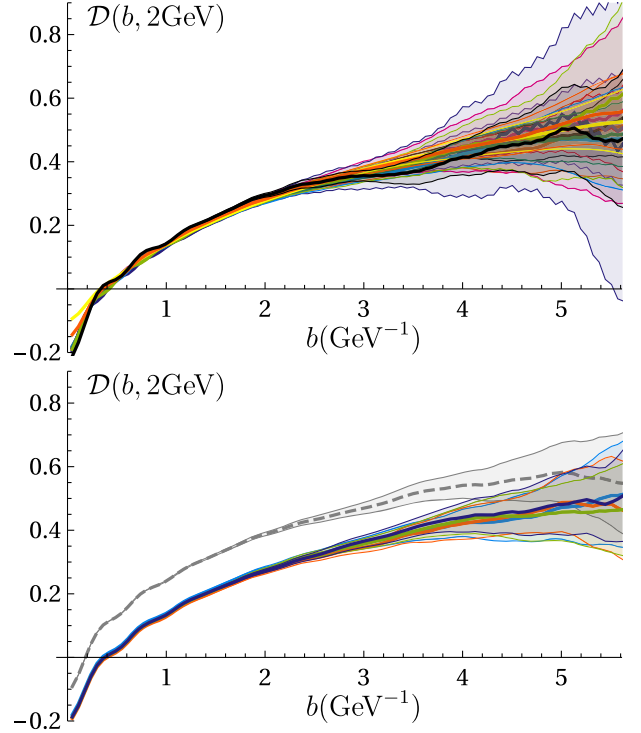

FIG. 2. Comparison of CS kernels extracted from different combinations of the pseudodata. The top plot shows all possible (twelve) combinations of pseudodata with different kinematics, listed in the Table I. The bottom plot show extractions made with different input collinear PDFs. The solid lines are the central values. The shaded areas are the statistical uncertainty. The oscillations at $b \sim 4$–6 GeV^{-1} are due to the finite bin size in the q_T-space. The gray dashed line in the lower plot shows the effect of incomplete cancellation of parton's momentum if PDFs in the comparing cross section are different (here, CT18 vs CASCADE).

CT18 at $Q = 16$ GeV), where the parton distributions do not cancel exactly.

All extractions are made using $\mu_Q = Q$, and N^{3}LO perturbative accuracy for functions Z and Δ_R. The perturbative expansion is very stable, which we test by varying the scale $\mu_Q \in [Q/2, 2Q]$. The size of the scale variation band is constant in b. The maximum variation among all extractions is $(-0.0016, +0.0008)$ in the absolute value, which is smaller than the statistical uncertainty.

The final curve for the CS kernel predicted by the CASCADE event generator is obtained by combining all twelve extractions. The comparison with other extractions of CS kernel is shown in Fig. 3. The CASCADE extraction lightly disagrees with the perturbative curve ($b < 1$ GeV^{-1}), but in agreement with the SV19 [10] and Pavia17 [7] for $1 < b < 3$ GeV^{-1}.

The fit of the large-b part by a polynomial gives

$$\mathcal{D}(b, \mu) \sim [(0.069 \pm 0.031) \text{ GeV}] \times b, \qquad (11)$$

with a negligible quadratic part. We conclude that the CASCADE suggests a linear asymptotic, which was also used in the SV19 series of fits [9,10,41], and supported by theoretical estimations [14,42]
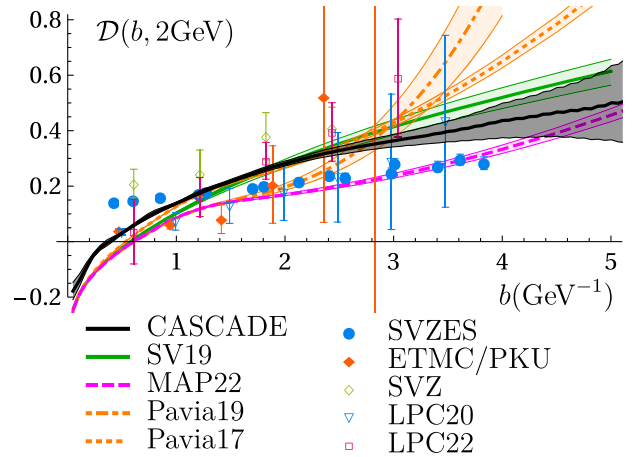

FIG. 3. Comparison of the CS kernels obtained in different approaches. The black curve is the one obtained in this work. The curves SV19, MAP22, Pavia19, and Pavia17 are obtained from the fits of Drell-Yan and SIDIS data in Refs. [7,10,11,37], correspondingly. Dots represent the computations of CS kernel on the lattice, with SVZES, ETMC/PKU, SVZ, LPC20 and LPC22 corresponding to Refs. [16,17,38–40].

IV. CONCLUSIONS

We have presented a novel method for direct determination of the CS kernel from data, by using the proper combination of cross sections in different kinematic regions. For explicitness, we considered the case of the Drell-Yan process, nevertheless the method can be easily generalized to other processes such as SIDIS, semi-inclusive annihilation, and their polarized versions.

The method is tested using pseudodata generated with the CASCADE event generator, and the corresponding CS kernel is determined. As a by-product, we have confirmed the compatibility between the TMD factorization and the PB approach, since all expected properties of the CS kernel (such as universality) are observed in the CASCADE generator. This solves an old-stated problem of comparison between factorization theorem and PB approaches [43,44].

The procedure can be applied to the real experimental data without modifications. In this case, the quality of extraction will be essentially affected by the resolution and finite-bin effects. On the other hand, there is an explicit advantage in the reduction of sensitivity to various sources of systematic experimental uncertainty, which partially cancel in ratios. Thus the method is feasible for modern and future experiments, such JLab [45,46], LHC [47], and EIC [48,49].

We stress that the procedure is model-independent and excludes any adaptation of parameters. The factorization theorem guaranties the same result independently on the used reaction and kinematic range of data. A deviation from this rule indicates the violation of the factorization theorem. Therefore, the method gives an ultimate test of the TMD factorization hypothesis, which is essential for experiments with lower energies (such as JLab, COMPASS, EIC) where factorization is concerned.

[1] J. C. Collins, D. E. Soper, and G. F. Sterman, Adv. Ser. Dir. High Energy Phys. **5**, 1 (1989).

[2] J. C. Collins, D. E. Soper, and G. F. Sterman, Nucl. Phys. **B250**, 199 (1985).

[3] J. Collins, *Foundations of Perturbative QCD* (Cambridge University Press, Cambridge, England, 2013), Vol. 32.

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

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

[6] A. A. Vladimirov, Phys. Rev. Lett. **125**, 192002 (2020).

[7] A. Bacchetta, F. Delcarro, C. Pisano, M. Radici, and A. Signori, J. High Energy Phys. 06 (2017) 081; 06 (2019) 051(E).

[8] I. Scimemi and A. Vladimirov, Eur. Phys. J. C **78**, 89 (2018).

[9] V. Bertone, I. Scimemi, and A. Vladimirov, J. High Energy Phys. 06 (2019) 028.

[10] I. Scimemi and A. Vladimirov, J. High Energy Phys. 06 (2020) 137.

[11] A. Bacchetta, V. Bertone, C. Bissolotti, G. Bozzi, F. Delcarro, F. Piacenza, and M. Radici, J. High Energy Phys. 07 (2020) 117.

[12] M. A. Ebert, I. W. Stewart, and Y. Zhao, Phys. Rev. D **99**, 034505 (2019).

[13] X. Ji, Y. Liu, and Y.-S. Liu, Phys. Lett. B **811**, 135946 (2020).

[14] A. A. Vladimirov and A. Schäfer, Phys. Rev. D **101**, 074517 (2020).

[15] P. Shanahan, M. Wagman, and Y. Zhao, Phys. Rev. D **102**, 014511 (2020).

[16] M. Schlemmer, A. Vladimirov, C. Zimmermann, M. Engelhardt, and A. Schäfer, J. High Energy Phys. 08 (2021) 004.

[17] P. Shanahan, M. Wagman, and Y. Zhao, Phys. Rev. D **104**, 114502 (2021).

[18] M.-H. Chu *et al.*, arXiv:2204.00200.

[19] F. Hautmann, I. Scimemi, and A. Vladimirov, Phys. Lett. B **806**, 135478 (2020).

[20] S. Baranov *et al.*, Eur. Phys. J. C **81**, 425 (2021).

[21] A. Bermudez Martinez *et al.*, Eur. Phys. J. C **80**, 598 (2020).

[22] A. B. Martinez, L. I. Estevez Banos, H. Jung, J. Lidrych, M. Mendizabal, S. T. Monfared, Q. Wang, and H. Yang, Proc. Sci., EPS-HEP2021 (**2022**) 386 [arXiv:2111.03582].

[23] A. Bermudez Martinez *et al.*, Phys. Rev. D **100**, 074027 (2019).

[24] A. M. Sirunyan *et al.* (CMS Collaboration), J. High Energy Phys. 12 (2019) 061.

[25] A. B. Martinez, F. Hautmann, and M. L. Mangano, Phys. Lett. B **822**, 136700 (2021).

[26] T. Gehrmann, E. W. N. Glover, T. Huber, N. Ikizlerli, and C. Studerus, J. High Energy Phys. 06 (2010) 094.

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

[28] J.-Y. Chiu, A. Jain, D. Neill, and I. Z. Rothstein, J. High Energy Phys. 05 (2012) 084.

[29] I. Scimemi and A. Vladimirov, J. High Energy Phys. 08 (2018) 003.

[30] There is no common notation for CS kernel. In the literature, it is also denoted as $\tilde{K} = \gamma_\nu^f = -F_{f\bar{f}} = -2\mathcal{D}$ [5,27–29].

[31] F. Hautmann, H. Jung, A. Lelek, V. Radescu, and R. Zlebcik, Phys. Lett. B **772**, 446 (2017).

[32] F. Hautmann, H. Jung, A. Lelek, V. Radescu, and R. Zlebcik, J. High Energy Phys. 01 (2018) 070.

[33] D. Lemoine, J. Chem. Phys. **101**, 3936 (1994).

[34] T.-J. Hou *et al.*, Phys. Rev. D **103**, 014013 (2021).

[35] S. Bailey, T. Cridge, L. A. Harland-Lang, A. D. Martin, and R. S. Thorne, Eur. Phys. J. C **81**, 341 (2021).

[36] R. D. Ball *et al.* (NNPDF Collaboration), Eur. Phys. J. C **77**, 663 (2017).

[37] A. Bacchetta, V. Bertone, C. Bissolotti, G. Bozzi, M. Cerutti, F. Piacenza, M. Radici, and A. Signori, J. High Energy Phys. 10 (2022) 127.

[38] Y. Li *et al.*, Phys. Rev. Lett. **128**, 062002 (2022).

[39] Q.-A. Zhang *et al.* (Lattice Parton Collaboration), Phys. Rev. Lett. **125**, 192001 (2020).

[40] M.-H. Chu *et al.* (LPC Collaboration), Phys. Rev. D **106**, 034509 (2022).

[41] M. Bury, F. Hautmann, S. Leal-Gomez, I. Scimemi, A. Vladimirov, and P. Zurita, arXiv:2201.07114.

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

[43] N. A. Abdulov *et al.*, Eur. Phys. J. C **81**, 752 (2021).

[44] R. Angeles-Martinez *et al.*, Acta Phys. Pol. B **46**, 2501 (2015).

[45] J. Dudek *et al.*, Eur. Phys. J. A **48**, 187 (2012).

[46] A. Accardi *et al.*, Eur. Phys. J. A **57**, 261 (2021).

[47] S. Amoroso *et al.*, arXiv:2203.13923.

[48] R. Abdul Khalek *et al.*, Nucl. Phys. **A1026**, 122447 (2022).

[49] R. A. Khalek *et al.*, in 2022 Snowmass Summer Study (2022), arXiv:2203.13199.