JYX
JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.

Author(s): Beuf, G.; Hänninen, H.; Lappi, T.; Mulian, Y.; Mäntysaari, H.

Title: Diffractive deep inelastic scattering at NLO in the dipole picture : The $q\bar{q}g$ contribution

Year: 2022

Version: Published version

Copyright: © Authors, 2022

Rights: CC BY 4.0

Rights url: https://creativecommons.org/licenses/by/4.0/

Please cite the original version:

Beuf, G., Hänninen, H., Lappi, T., Mulian, Y., & Mäntysaari, H. (2022). Diffractive deep inelastic scattering at NLO in the dipole picture : The $q\bar{q}g$ contribution. Physical Review D, 106(9), Article 094014. https://doi.org/10.1103/PhysRevD.106.094014

Diffractive deep inelastic scattering at NLO in the dipole picture: The $q\bar{q}g$ contribution

G. Beuf,[1] H. Hänninen,[2,3] T. Lappi,[2,3] Y. Mulian,[4] and H. Mäntysaari[2,3]

[1]*National Centre for Nuclear Research, 02-093 Warsaw, Poland*
[2]*Department of Physics, University of Jyväskylä, P.O. Box 35, 40014 University of Jyväskylä, Finland*
[3]*Helsinki Institute of Physics, P.O. Box 64, 00014 University of Helsinki, Finland*
[4]*Instituto Galego de Física de Altas Enerxías IGFAE, Universidade de Santiago de Compostela, 15782 Santiago de Compostela, Galicia-Spain*

(Received 1 July 2022; accepted 23 September 2022; published 9 November 2022)

We calculate the contribution from the $q\bar{q}g$ state production to the diffractive cross sections in deep inelastic scattering at high energy. The obtained cross section is finite by itself and a part of the full next-to-leading order result for the diffractive structure functions. We perform the calculation in exact kinematics in the eikonal limit, and show that the previously known high-Q^2 and large M_X^2 results for the structure functions can be extracted from our results in the appropriate limits. We furthermore discuss the steps required to obtain the full next-to-leading order results for the structure functions.

DOI: 10.1103/PhysRevD.106.094014

I. INTRODUCTION

In collisions with high scattering energy, one is measuring degrees of freedom of hadronic and nuclear states that only have a small fraction of the full momentum of the state, small-x degrees of freedom. The large amount of phase space available at high collision energies leads to an exponentially cascading emission of gluons. At some point, however, this cascade must be limited by unitarity requirements on scattering amplitudes. Thus gluon mergings eventually start to be equally important, even at transverse resolution scales where a weak coupling description is appropriate. The kinematical region where these two effects balance each other is referred to as the gluon saturation regime, and understanding it is the topic of many theoretical and experimental efforts. Experimentally the saturation regime is relevant for understanding hadronic collision processes and the formation of quark gluon plasma at RHIC and the LHC. A particularly precise and clean way to access the small-x degrees of freedom is, however, provided by high energy deep inelastic scattering (DIS), both in the HERA experiments, and at the future electron-ion collider (EIC) [1–3] and LHeC [4,5]. Theoretically, a convenient way to discuss the physics of gluon saturation is provided by the color glass condensate (CGC) [6–8] effective field theory, where the nonlinear gluon system is described as a classical color field. For the DIS process, the CGC framework naturally leads to the dipole picture [9–13].

In the dipole picture one factorizes the DIS process of a virtual photon off a hadronic target into two ingredients. First, the perturbative part of the process is the development of the photon into a partonic state, to leading order a color neutral quark-antiquark dipole. The second ingredient is the scattering of this partonic state with the gluonic target, which in the high collision energy limit can be treated as an eikonal interaction with the classical color field. With the prospect of higher luminosities and the availability of nuclear targets in future DIS experiments, there has been a systematic push in the field to improve the perturbative accuracy of the dipole picture by going to higher orders in perturbation theory. In recent years the dipole picture has been extended to NLO accuracy for the high energy BK/JIMWLK evolution [14–27] and the inclusive DIS cross section [28–40].

Exclusive or diffractive DIS is expected to be even more sensitive to gluon saturation than inclusive cross sections [3,41–43]. One way to understand this is to note that, due to the optical theorem, the total cross section is proportional to the elastic dipole-target amplitude, proportional to the gluon distribution in the target. Exclusive cross sections, on the other hand, are calculated as the square of the amplitude, and are thus much more sensitive to the large amplitudes, a signature of the saturation regime. Correspondingly, the recent work on inclusive scattering has been accompanied by several calculations of exclusive vector meson and diffractive dijet production at NLO in the dipole picture [44–53]. While these processes are an

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.

extremely important part of the coming experimental program, they both have some drawbacks for the purpose of understanding gluon saturation. Exclusive vector meson production requires some knowledge or modeling of the bound state physics of the meson. While there are systematical ways to do this perturbatively, e.g., by a nonrelativistic QCD approach as in Ref. [48] or by using universal parton distribution amplitudes that can independently be measured in other processes [46,50,54], this unavoidably adds an additional source of uncertainty. Dijets, on the other hand, are well defined perturbative objects in a given jet algorithm, but only if the jet transverse momenta are sufficiently large. At realistic collider energies this has a tendency to push jet measurements to larger x and thus outside the saturation regime.

In this paper we will focus on a process that has gathered somewhat less attention in the work to push the dipole picture to NLO accuracy; namely, inclusive diffractive DIS. Here the experimental signature is a large rapidity gap between the diffractive system (X) consisting of the photon remnants, and the target or its remnants. In the dipole picture the photon fluctuates into a variety of partonic states, which scatter off the target without exchanging color. In this sense the rapidity gap makes the process fundamentally an exclusive one, with a cross section given by the square of an amplitude. On the other hand, the measurement is inclusive in the sense that one sums over all of the different final states of the diffractive system, measuring the cross section differentially only in its total invariant mass M_X. This latter inclusive aspect makes it possible to extend the perturbative description to much lower invariant masses and to lower x than for diffractive dijets, even if the parton level final states are the same. The cross sections for such processes are expressed in terms of the diffractive structure functions $F_2^{\mathrm{D}(3)}(\beta, Q^2, x_{\mathbb{P}})$ and $F_L^{\mathrm{D}(3)}(\beta, Q^2, x_{\mathbb{P}})$ or, equivalently, the diffractive virtual photon cross sections $\mathrm{d}\sigma^{\gamma^*+A\to A+n}/\mathrm{d}[\mathcal{PS}]_n$, which are the quantities that we will calculate here.

At leading order in α_s, the diffractive final state only consists of a quark-antiquark dipole. This already provides a good description of the general features of the experimental measurements at small $M_X^2 \sim Q^2$ [41] (see also work in Refs. [55–57]). However, a strict leading order picture fails to describe the rise of the cross section towards larger M_X where, in order to make a high invariant mass partonic state, additional gluon radiation is required. The phenomenologically most successful approach has been to use the "Wüsthoff result" [58], which includes the radiation of one extra gluon into the final state [59–64] in a large Q^2 kinematical approximation. In our terminology, this tree-level gluon emission is already a part of the NLO result, being explicitly proportional to α_s. In this paper we will calculate the same contributions as in the Wüsthoff result at what we call the exact kinematics in the eikonal limit. This means that the kinematics within a diffractive system is treated exactly without a large Q^2 approximation, while the interaction with the target is eikonal.

Our result presented in this paper corresponds to a subset of the NLO results that is finite by itself, and suited for explicit numerical evaluation. The completion of the full NLO calculation requires the inclusion of virtual corrections with gluons that are not produced in the final state. We plan to return to these contributions in future work. Here we will merely outline steps that are needed to calculate these virtual contributions in our formalism. We have also here opted to calculate only the contributions where the gluon is emitted before the shockwave, not combining them with emissions after. This avoids issues with collinear and soft divergences in final state emissions, which eventually cancel against virtual corrections.

This paper is structured in the following way. We will start by introducing the experimental observable, the diffractive structure function, in Sec. II and the dipole picture formulation in terms of LCPT in Sec. III. Before moving to specific diagrams we will then discuss in Sec. V the general strategy to calculate phase space integrals for 2- and 3-particle final states of a fixed invariant mass M_X in the context of an eikonal scattering picture where the interaction with the target happens at a fixed transverse coordinate. We will rederive the known result for the leading $q\bar{q}$ component of the wave function in our notations in Sec. VI. We then move to the main new result of this paper, the calculation of the $q\bar{q}g$ component of the cross section in full kinematics in Sec. VII. A more detailed exposition of intermediate stages of the calculation has been presented earlier in Ref. [65]. We check in Sec. VIII that our calculation reduces to known results in the kinematical limits of large Q^2 (Wüsthoff [58]) and large M_X (e.g., in Ref. [66]), before concluding in Sec. IX. The results in this paper cover a finite, self-contained subset of the NLO corrections to the diffractive results, generalizing earlier calculations to the full kinematics. We plan to return to the calculation of the remaining parts in a future publication, as outlined in Sec. III.

II. DIFFRACTIVE STRUCTURE FUNCTIONS

The diffractive cross section $\sigma^{\mathrm{D}}_{e+A\to M_X+p}$ in electron-nucleus (or electron-proton) DIS integrated over the squared momentum transfer t is usually expressed in terms of the diffractive structure functions F_2^{D} and F_L^{D} defined as

$$\frac{\mathrm{d}\sigma^{\mathrm{D}}_{e+A\to M_X+p}}{\mathrm{d}\beta \mathrm{d}Q^2 \mathrm{d}x_{\mathbb{P}}} = \frac{2\pi\alpha_{\mathrm{em}}^2}{\beta Q^4}[1+(1-y)^2]\left[F_2^{\mathrm{D}(3)}(\beta, Q^2, x_{\mathbb{P}}) - \frac{y^2}{1+(1-y)^2}F_L^{\mathrm{D}(3)}(\beta, Q^2, x_{\mathbb{P}})\right]. \tag{1}$$

The superscript (3) refers to the structure functions that depend on three variables, in this case β, Q^2, and $x_{\mathbb{P}}$

discussed above. One can also consider structure functions differentially in the squared momentum transfer t, in which case one has the structure functions $F_{2,L}^{\mathrm{D}(4)}(\beta, Q^2, x_{\mathbb{P}}, t)$. In this work we consider both t-differential and t-integrated cross sections. For simplicity we focus here on coherent diffraction that corresponds to the events where the target does not dissociate, but our results are straightforward to generalize to dissociative events in the Good-Walker [67] picture (see, e.g., Refs. [68–71]).

The diffractive structure functions are related to the total diffractive cross sections in $\gamma^* + A$ scattering as

$$x_{\mathbb{P}} F_{T,L}^{\mathrm{D}(4)} = \frac{Q^2}{4\pi^2 \alpha_{\mathrm{em}}} \frac{Q^2}{\beta} \frac{\mathrm{d}\sigma^{\mathrm{D}}_{\gamma^*_{T,L}+A\to M_X+A}}{\mathrm{d}M_X^2 \mathrm{d}t}, \tag{2}$$

where T and L refer to transversely and longitudinally polarized photons. The experimentally measured [72–75] total diffractive cross sections are usually reported in terms of the diffractive reduced cross section

$$\sigma_r^{\mathrm{D}(3)}(\beta, Q^2 x_{\mathbb{P}}) = F_2^{\mathrm{D}(3)}(\beta, Q^2 x_{\mathbb{P}}) - \frac{y^2}{1+(1-y)^2} F_L^{\mathrm{D}(3)}(\beta, Q^2 x_{\mathbb{P}}). \tag{3}$$

Here the Lorentz-invariant quantities describing the kinematics are the virtuality of the photon $-Q^2$ and the fraction of the target longitudinal momentum $x_{\mathbb{P}}$ carried by the pomeron (exchanged in the scattering process) in the frame where the target has a large longitudinal momentum, defined as

$$x_{\mathbb{P}} \equiv \frac{(P-P')\cdot q}{P\cdot q} = \frac{M^2+Q^2-t}{W^2+Q^2-m_N^2} \approx \frac{M_X^2+Q^2}{W^2+Q^2}. \tag{4}$$

The invariant mass of the diffractively produced system is denoted by M_X^2 and the nucleon mass by m_N^2. The variable $\beta = Q^2/(2q\cdot(P-P')) \approx Q^2/(M_X^2+Q^2)$ has, in the frame where the target momentum is large, an interpretation as the fraction of the pomeron momentum carried by the struck quark. The four vectors P, P' are the target nucleon momenta before and after the scattering, respectively, see Fig. 1. Finally $y = (P\cdot q)/(P\cdot l)$ is the inelasticity describing the energy transfer from the lepton with initial momentum l, and q is the photon momentum.

We are working in the dipole picture, where one develops the virtual photon state in a series of partonic Fock states. Let us first consider the general case of an n-parton Fock state, for which we denote the phase space element as $[\mathcal{PS}]_n$, At leading order only the $n = q\bar{q}$ state contributes, and in this work we focus on including the $n = q\bar{q}g$ contribution, which is actually the dominant component at high M_X^2 (small β) and at high Q^2 [62]. This tree-level contribution is also a necessary ingredient for the future full calculation of the diffractive structure

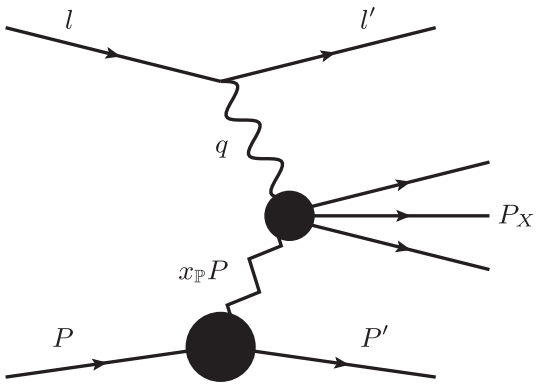

FIG. 1. Kinematics of inclusive diffractive DIS.

functions at NLO accuracy. The diffractive cross section can be written as

$$\frac{\mathrm{d}\sigma^{\mathrm{D}}_{\gamma^*_{T,L}+A\to M_X+A}}{\mathrm{d}M_X^2} = \sum_n \int \mathrm{d}[\mathcal{PS}]_n \frac{\mathrm{d}\sigma^{\gamma^*_{T,L}+A\to A+n}}{\mathrm{d}[\mathcal{PS}]_n} \delta(M_X^2 - M_n^2), \tag{5}$$

where M_n^2 is the invariant mass of the Fock state n. The cross section $\frac{\mathrm{d}\sigma^{\gamma^*A\to A+n}}{\mathrm{d}[\mathcal{PS}]_n}$ for a production of a color-singlet state n in photon-nucleus or photon-proton scattering is expressed in terms of the scattering amplitudes (see [76], except we now normalize with a $2q^+$ in a different place) as

$$\mathrm{d}\sigma^{\mathrm{D}}_{\gamma^*+A\to A+n} = 2q^+(2\pi)\delta(q^+ - q_n^+) \prod_{i\in \mathrm{F.s.}n} \widetilde{\mathrm{d}p_i} \, |\mathcal{M}_{\gamma\to n}|^2. \tag{6}$$

Here $i \in \mathrm{F.s.}n$ means iterating over all the particles i in the Fock state n and q_n^+ is the total plus momentum of the partons in this Fock state. The one particle phase space element reads

$$\widetilde{\mathrm{d}p_i} \equiv \frac{\mathrm{d}^2\mathbf{p}_i \mathrm{d}p_i^+}{2p_i^+(2\pi)^3}. \tag{7}$$

The scattering amplitude is obtained from the matrix elements of dressed, interacting, states by leaving out a momentum conservation delta function

$${}_D\langle \mathrm{F.S.}n|\hat{S}-1|\gamma\rangle_D = 2q^+(2\pi)\delta(q_\gamma^+ - q_n^+)\mathcal{M}_{\gamma\to n}, \tag{8}$$

where in the diffractive scattering the final state F.S.n is a color singlet.

At high energies, the scattering amplitudes $\mathcal{M}_{\gamma\to n}$ can be calculated by considering the $\gamma \to n$ process, and inserting an interaction with the shockwave (target color field) in all possible ways. At high energies the transverse coordinates of the partons are fixed when they propagate in the color field of the target and as such the interaction can be straightforwardly described in terms of Wilson lines at

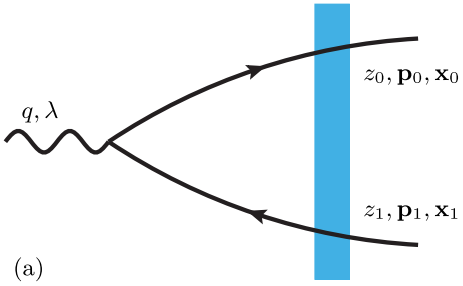

FIG. 2. Leading order amplitude. The blue band represents the interaction with the target color field.

fixed transverse coordinates as discussed in more detail in Sec. IV.

III. OUTLINE OF NLO CALCULATION

We will in this paper compute a part of the NLO correction to the diffractive structure functions that is finite by itself. However, let us first discuss the overall structure of the full NLO contribution in terms of the contributing diagrams. This discussion will make it more clear which parts of the NLO contributions are included in our result here, and what still needs to be done to calculate the rest. We are calculating, and drawing diagrams, for an exclusive amplitude for a virtual photon-target shockwave scattering. Thus a single diagram includes both the Fock state expansion of the incoming dressed virtual photon state $|\gamma\rangle_D$, and that of the outgoing dressed multiparton state ${}_D\langle \mathrm{F.s.}n|$ in the amplitude (8). In the diagrams the shockwave is represented by a blue band, with time progressing from left to right. The state furthest to the left is the asymptotic incoming state (the dressed virtual photon), which then develops into a superposition of bare parton states, corresponding to the LFWFs $\gamma \to n$. The interaction with the shockwave then is given in terms of the bare states [77]. On the other side of the shockwave, furthest to the right, is the asymptotic final state ${}_D\langle \mathrm{F.s.}n|$, which develops into a superposition of bare states, going leftward in the figure. Thus the part of the figure to the right of the shockwave corresponds to the complex conjugate of the LFWF of the final state, in particular with energy denominators calculated with respect to the final state.

At leading order, the only diagram contributing to the scattering amplitude is diagram (a) shown in Fig. 2, where the photon first splits to a $q\bar{q}$ dipole, and then subsequently the quarks scatter off the target color field with no net color charge transfer to the target. In momentum space, we denote the quark and antiquark transverse momenta as $\mathbf{p}_0$ and $\mathbf{p}_1$, respectively, and in transverse coordinate space use the coordinates $\mathbf{x}_0$, $\mathbf{x}_1$. Similarly the fractions of the photon plus momentum carried by the quark and the antiquark are $z_i = p_i^+/q^+$ with $z_0 + z_1 = 1$.

A. Radiative corrections

The purpose of this paper is to calculate the gluon emission part of the next-to-leading order contributions to the diffractive structure functions, which dominates at large M_X^2. In the CGC, if one integrates the transverse momentum of several final state particles without restriction, one can encounter spurious UV divergences in real higher order corrections, associated with the breakdown of the eikonal approximation. This occurs when the light-cone momentum p^- scales associated with the produced system become comparable or larger than that of the target. In the case of diffractive structure functions considered here, the fixed invariant mass of the produced $q\bar{q}g$ state ensures that the eikonal approximation stays valid for the whole integration range, by constraining the p^- scale of the diffractive system. For that reason, no UV divergence can arise in real NLO corrections to diffractive structure functions. This is to be compared to the case of inclusive DIS [31–33] where one uses the optical theorem and thus the final state is completely fixed to be the same as the initial one. In dijet production [36,37,44,45,52], on the other hand, one typically fixes the momenta of some of the final state particles and integrates over the others. This can lead to a different pattern of cancellations between diagrams. The calculation of the loop corrections is left for future work, but for completeness we list all the relevant diagrams in the following Sec. III B.

In order to calculate the $n = q\bar{q}g$ contribution to the diffractive scattering amplitude we include gluon emission contributions from both the quark and the antiquark. There can be a regular gluon emission before the shockwave shown in diagrams in Figs. 3(b) and 3(c). In light cone

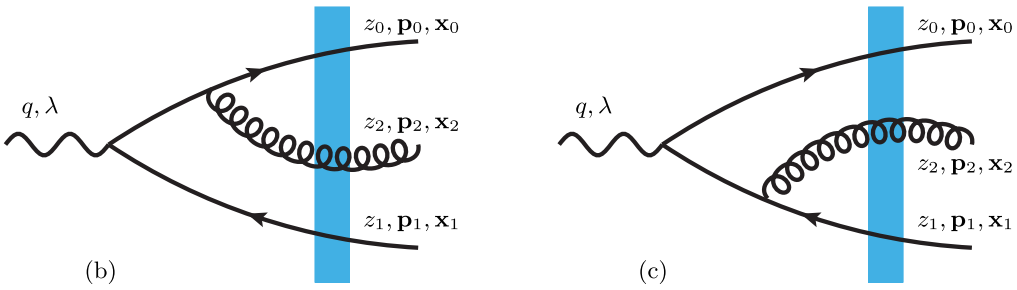

FIG. 3. Gluon emission before the shock.

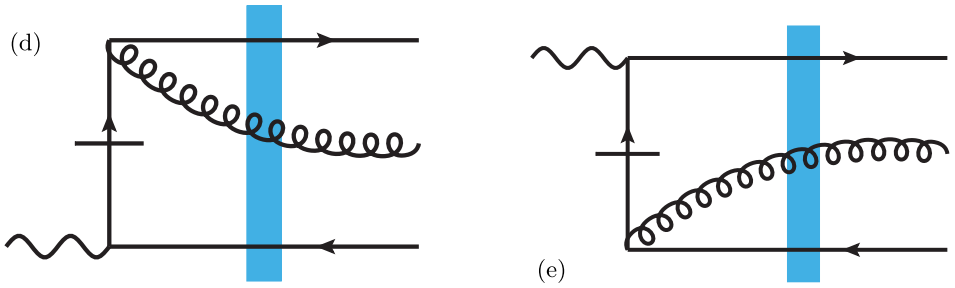

FIG. 4. Instantaneous diagram gluon emission wave function with emission before the cut.

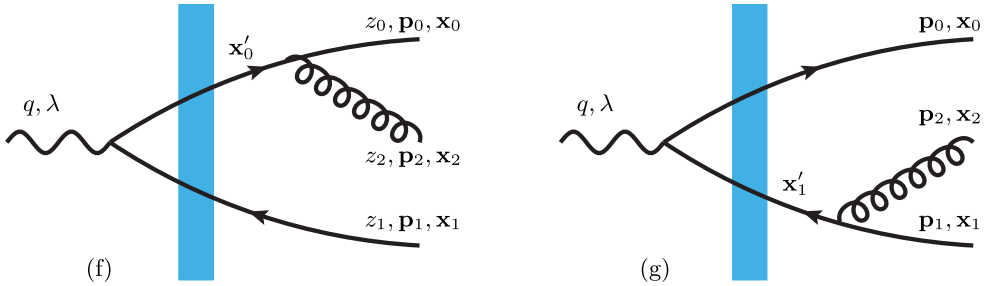

FIG. 5. Gluon emission after shock. Note that the intermediate transverse coordinate of the quark (antiquark) before the gluon emission is not the same as the coordinate at the cut. They are defined by $\mathbf{x}_0' := \frac{z_0\mathbf{x}_0+z_2\mathbf{x}_2}{z_0+z_2}$ and $\mathbf{x}_1' := \frac{z_1\mathbf{x}_1+z_2\mathbf{x}_2}{z_1+z_2}$.

perturbation theory there is also an instantaneous $\gamma \to q\bar{q}g$ vertex resulting in instantaneous gluon emission diagrams in Figs. 4(d) and 4(e). These are the contributions that we will calculate in detail in this paper. Similarly the gluon can be emitted after the shock, see diagrams in Figs. 5(f) and 5(g). There are several relations between these diagrams. First, since the virtual photon as a whole is color neutral, there is a destructive interference between emissions from the quark and from the antiquark. This cancels the leading small transverse momentum (collinear) divergence and serves as a useful check of the relative sign between the contributions.

The contributions with emissions after the shockwave, pictured in Fig. 5, can be conveniently obtained from the corresponding ones in Fig. 3 where the emission happens before by taking a specific coordinate limit, in a procedure developed in Ref. [78]. Here one first separates out from the coordinate space $\gamma^* \to q\bar{q}g$ wave function a piece corresponding to the final gluon emission. In terms of equations this means that one writes the $\gamma^* \to q\bar{q}g$ wave function obtained from diagrams in Figs. 3(b), 3(c), 4(d), and 4(e) in a factorized form as[1]

$$\tilde{\psi}_{\gamma_\lambda^*\to q_0\bar{q}_1 g_2} = \tilde{\psi}_{\gamma_\lambda^*\to q_0\bar{q}_1;q_0\to q_0g_2}\tilde{\psi}_{q_0\to q_0g_2} + \tilde{\psi}_{\gamma_\lambda^*\to q_0\bar{q}_1;\bar{q}_1\to\bar{q}_1g_2}\tilde{\psi}_{\bar{q}_1\to\bar{q}_1g_2}. \quad (9)$$

[1]Here $\tilde{\psi}$ denotes a reduced coordinate space wave function. See Eqs. (16) and (17) for the normalization in the $\gamma^* \to q\bar{q}$, $\gamma^* \to q\bar{q}g$ case; the convention is trivially extended to the $1 \to 2$ gluon emission case.

Here $\tilde{\psi}_{q_0\to q_0g_2}$ and $\tilde{\psi}_{\bar{q}_1\to\bar{q}_1g_2}$ are the $1 \to 2$ particle gluon emission wave functions. Equation (9) should be understood as the *definition* of the remaining parts $\tilde{\psi}_{\gamma_\lambda^*\to q_0\bar{q}_1;q_0\to q_0g_2}$ and $\tilde{\psi}_{\gamma_\lambda^*\to q_0\bar{q}_1;\bar{q}_1\to\bar{q}_1g_2}$. Here the notation refers to these being the parts of the wave function that are associated (e.g., by the helicity structure) with the first $\gamma^* \to q\bar{q}$ splitting ($\gamma_\lambda^* \to q_0\bar{q}_1$), but depend on the fact that the (anti)quark will later emit a gluon ($q_0 \to q_0g_2$, $\bar{q}_1 \to \bar{q}_1g_2$). Thus $\tilde{\psi}_{\gamma_\lambda^*\to q_0\bar{q}_1;q_0\to q_0g_2}$ and $\tilde{\psi}_{\gamma_\lambda^*\to q_0\bar{q}_1;\bar{q}_1\to\bar{q}_1g_2}$ depend on the coordinates of all three particles in the final state. The calculation of the gluon emission diagrams after the shock wave requires the $q\bar{q}$ component in the dressed ${}_D\langle q\bar{q}g|$ state. This, in turn, requires the $(q\bar{q}g \to q\bar{q})^\dagger$ merging wave function, which is given by (minus) the Hermitian conjugates of the corresponding emission wave functions $\tilde{\psi}_{q_0\to q_0g_2}$ and $\tilde{\psi}_{\bar{q}_1\to\bar{q}_1g_2}$. In other words, one is here factoring out from the gluon emission before the shockwave the gluon emission wave function that appears when the gluon is emitted after. A look at the transverse coordinate space $\gamma \to q\bar{q}g$ wave function [see, e.g., Eqs. (C14) and (C19) in Ref. [33] for explicit expressions] shows that indeed the structure of the regular emission wave functions naturally factorizes like this.

Using the factorized notation (9), the procedure of Ref. [78] for obtaining amplitudes for the emission-after-the-shock contributions in Fig. 5 is the following. One evaluates both the Wilson line operators and the $\gamma^* \to q\bar{q}$ parts of the wave functions $\tilde{\psi}_{\gamma_\lambda^*\to q_0\bar{q}_1;q_0\to q_0g_2}$, $\tilde{\psi}_{\gamma_\lambda^*\to q_0\bar{q}_1;\bar{q}_1\to\bar{q}_1g_2}$ with the transverse coordinates of the gluon and its parent

(anti)quark replaced by the coordinate of the parent before the emission, and changes the sign. Thus, for the emission from the quark, diagram in Fig. 3(b), one replaces $\mathbf{x}_0 \to \mathbf{x}_0'$ and $\mathbf{x}_2 \to \mathbf{x}_0'$, in both the Wilson line operator and in $\tilde{\psi}_{\gamma_\lambda^* \to q_0\bar{q}_1; q_0 \to q_0 g_2}$, with the coordinate defined as $\mathbf{x}_0' := \frac{z_0\mathbf{x}_0 + z_2\mathbf{x}_2}{z_0+z_2}$. Correspondingly, for the emission from the antiquark, diagram in Fig. 3(c), one replaces $\mathbf{x}_1 \to \mathbf{x}_1'$ and $\mathbf{x}_2 \to \mathbf{x}_1'$, with $\mathbf{x}_1' := \frac{z_1\mathbf{x}_1 + z_2\mathbf{x}_2}{z_1+z_2}$. It is clear that this corresponds to the Wilson line operator being evaluated at the correct coordinate for the diagrams in Fig. 5(f) and 5(g). There is no "emission after the shockwave" contribution for the instantaneous diagrams in Fig. 4(d) and 4(e). This is, however, built into the formalism of Ref. [78], since it turns out that the parts of $\tilde{\psi}_{\gamma_\lambda^* \to q_0\bar{q}_1; q_0 \to q_0 g_2}$, $\tilde{\psi}_{\gamma_\lambda^* \to q_0\bar{q}_1; \bar{q}_1 \to \bar{q}_1 g_2}$ corresponding to the instantaneous diagrams vanish in the coordinate limits $\mathbf{x}_0, \mathbf{x}_2 \to \mathbf{x}_0'$ and $\mathbf{x}_1, \mathbf{x}_2 \to \mathbf{x}_1'$, respectively.

One can arrive at this procedure for constructing the final state emissions in multiple ways. In Ref. [78] it is derived explicitly by looking at the expressions and noticing a relation between the energy denominators of the different diagrams. More generally, using the orthogonality of the $|\gamma\rangle_D$ and $|q\bar{q}g\rangle_D$ states one can derive the corresponding relation between the wavefunctions for $(q\bar{q}g \to q\bar{q})^\dagger$, $\gamma \to q\bar{q}$ [diagrams in Fig. 5(f) and 5(g)], $\gamma \to q\bar{q}g$ [diagrams in Figs. 3(b), 3(c), 4(d), and 4(e)] and the process $(q\bar{q}g \to \gamma)^\dagger$, corresponding to the photon crossing the shockwave first and all the emissions happening after (i.e., the bare photon state in the final ${}_D\langle q\bar{q}g|$). Since the last one does not contribute to the amplitude because the photon is color neutral, one obtains a linear relation for the contributions of the emission diagrams of Figs. 3 and 5. As discussed above, one can also see this relation directly by looking at the coordinate space wave functions, using the Fourier transforms from, e.g., Appendix C of [33].

In an inclusive observable where one integrates over the momenta of the final state gluon and of its parent without any restrictions, it would be natural to always keep the initial and final state gluon emissions, Figs. 3 and 5, together because they have a tendency to cancel each other in the UV region where the gluon is at the same coordinate as the emitting quark. Thus, they are often evaluated together such as in Refs. [78,79]. However, for the case of the diffractive structure function, the restriction on the diffractive system mass M_X cuts out contributions of large transverse momenta. Thus it is quite natural to evaluate the contributions of the emissions before and after the shockwave separately. On the other hand, the final state gluon emissions are associated with the wave function renormalization constants for, and gluon exchanges between, the outgoing quarks. The relation to these contributions which are, in our language, a part of the $q\bar{q}$ part of the cross section is especially important for the kinematical region when the gluon becomes collinear to the quark, where corresponding IR divergences must cancel. Thus it would not be natural here to consider the diagrams with gluon emission after the shockwave, before taking into account all the loop corrections. In conclusion, for a final state with a fixed M_X, the natural way to group diagrams together is different from some other observables. Since we are here leaving the NLO $q\bar{q}$ contribution overall to a future paper, we will also not calculate the final state emission contributions here. The exception to this is in Sec. VIII A, where one works in the $M_X \to \infty$ limit neglecting the restriction on final state momenta, and thus only the inclusion of the final state emissions allows one to get a finite result.

The 3-jet cross section in diffractive DIS has been computed earlier in Ref. [44], using the shockwave formalism that should be equivalent to our result here. The calculation includes emissions both before and after the shockwave, as is appropriate for the case of a fully differential 3-jet cross section. The IR divergences associated with wave function renormalization of the outgoing quarks would appear only after integrating over the phase space of the gluon, which is not done in Ref. [44], but is done here. Checking the equivalence of the result at the final cross section level would require a significant amount of algebra which we have not performed here. However, it was found in Ref. [44] that the result is compatible with the $\gamma^* \to q\bar{q}$ wave function of Ref. [30], which is the starting point of our calculation, as discussed in more detail in Sec. VII.

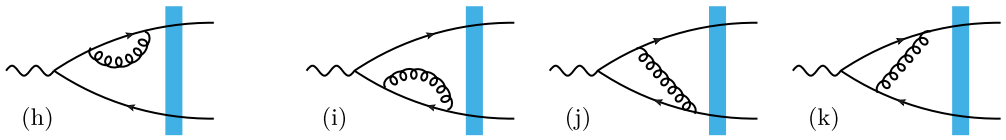

FIG. 6. Propagator and normal vertex correction diagrams calculated in [32,33].

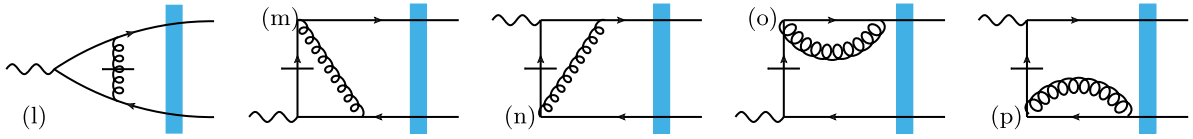

FIG. 7. Instantaneous vertex correction diagrams calculated in [32,33].

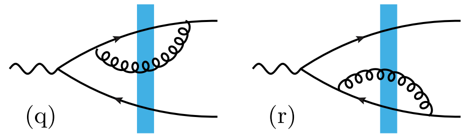

FIG. 8. Propagator correction diagrams where the gluon crosses the shockwave, but is not produced.

B. Loop corrections

In addition to the radiative contributions discussed in this paper, there are also loop corrections at NLO. We will return to them in more detail in a future paper, but let us make a few remarks here. First, there are the one-loop corrections to the $\gamma^* \to q\bar{q}$ wave function, depicted in Figs. 6 and 7. These include a quark propagator correction before the shockwave shown in diagrams in Fig. 6(h) and 6(i), and corrections to the $\gamma^* \to q\bar{q}$ vertex (including regular and instantaneous gluon or quark exchange) shown in diagrams in Figs. 6(j), 6(k), 7(l), 7(m), 7(n), 7(o), and 7(p). These one-loop wave functions have already been calculated as a part of the virtual photon NLO wave function [32,33,38–40], and the results for the loop calculations can be directly taken from these references.

In diffractive scattering there are also additional diagrams that are not needed for total (inclusive) cross section and as such are not available in the literature. Now it will be necessary to add propagator correction diagrams where the gluon crosses the shockwave, diagrams in Fig. 8(q) and 8(r). These exhibit UV divergences in the limit when the gluon coordinate becomes equal to the emitting (anti)quark. Similarly to the calculation of the inclusive cross section these will have to partially cancel UV divergences in the vertex correction diagrams in Figs. 6 and 7. This cancellation is the reason why the one-loop $\gamma^* \to q\bar{q}$ wave function alone is not sufficient to directly achieve the full NLO result. In addition to the propagator correction type diagrams, similar normal and instantaneous vertex correction diagrams in Fig. 9(s)–9(v) are also needed.

In our formalism (see [31,80] for more detailed discussions, based on the seminal work of [76]) we specifically exclude diagrams containing self-energy corrections inserted on the external asymptotic particles. Thus we do not explicitly have the diagrams in Fig. 10(w) and 10(x) in our calculation. Instead, one must attach to the amplitude a wave function renormalization constant $\sqrt{Z_{q/\bar{q}}}$ [again see Eq. (11)], which includes the same physical contribution, and is determined by the unitarity of the evolution operator.

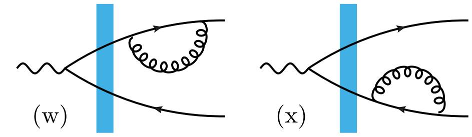

FIG. 10. Diagrams with propagator correction in final state. These are not included, but instead there is a wave function renormalization constant for the outgoing states.

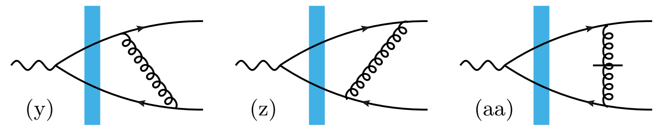

FIG. 11. Final state interaction diagrams.

The outgoing quark and antiquark wave function renormalization constants also have UV divergences, which should cancel the rest of the UV divergences from the vertex corrections. Finally there are also diagrams with a gluon exchange in the final state [diagrams in Fig. 11(y), 11(z), and 11(aa)]. Naively, one could think that these corrections correspond to a renormalization of the outgoing $q\bar{q}$ state, but they cannot be absorbed into just a constant, since the quark and antiquark actually exchange momentum in the exchange. A proper discussion of how to define the dressed $q\bar{q}$ outgoing state and treat the interactions between the outgoing particles is a major part of the discussion of the full NLO result, which we will return to in future work.

IV. INITIAL AND FINAL FOCK STATES

We will from now on focus on the leading order $q\bar{q}$ part and the radiative $q\bar{q}g$ part of the cross section. To begin, let us define explicitly our notations and normalization for the Fock states. The Fock expansion of the incoming virtual photon state in terms of bare partonic states is

$$
\begin{aligned}
|\gamma^*_\lambda(q^+,\mathbf{q};Q^2)\rangle_D = \sqrt{Z_{\gamma^*_\lambda}}\Big\{ & \text{Non-QCD Fock states} \\
& + \widetilde{\sum_{q_0\bar{q}_1\ \mathrm{F.s.}}} \tilde{\Psi}_{\gamma^*_\lambda\to q_0\bar{q}_1} \tilde{b}_0^\dagger \tilde{d}_1^\dagger |0\rangle \\
& + \widetilde{\sum_{q_0\bar{q}_1 g_2\ \mathrm{F.s.}}} \tilde{\Psi}_{\gamma^*_\lambda\to q_0\bar{q}_1 g_2} \tilde{b}_0^\dagger \tilde{d}_1^\dagger \tilde{a}_2^\dagger |0\rangle + \cdots \Big\},
\end{aligned}
\tag{10}
$$

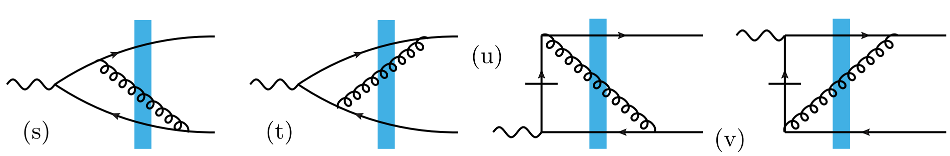

FIG. 9. Gluon emission diagrams where the gluon crosses the shockwave, but is not produced.

where F. s. stands for "Fock states." For the outgoing partonic states the corresponding expansions are

$$ {}_D\langle \bar{q}_1 q_0| = \sqrt{Z_q}\sqrt{Z_{\bar{q}}}\left\{ \langle 0|\tilde{d}_1\tilde{b}_0 + \widetilde{\sum_{q_{0'}\bar{q}_{1'}\,\mathrm{F.s.}}}{}' \langle 0|\tilde{d}_{1'}\tilde{b}_{0'}\tilde{\Psi}^{\dagger}_{q_0\bar{q}_1\to q_{0'}\bar{q}_{1'}} + \widetilde{\sum_{q_{0'}\bar{q}_{1'}g_{2'}\,\mathrm{F.s.}}} \langle 0|\tilde{a}_{2'}\tilde{d}_{1'}\tilde{b}_{0'}\tilde{\Psi}^{\dagger}_{q_0\bar{q}_1\to q_{0'}\bar{q}_{1'}g_{2'}} + \cdots \right\}, \tag{11} $$

$$ {}_D\langle g_2\bar{q}_1 q_0| = \sqrt{Z_g}\sqrt{Z_q}\sqrt{Z_{\bar{q}}}\left\{ \langle 0|\tilde{a}_2\tilde{d}_1\tilde{b}_0 + \widetilde{\sum_{q_{0'}\bar{q}_{1'}\,\mathrm{F.s.}}} \langle 0|\tilde{d}_{1'}\tilde{b}_{0'}\tilde{\Psi}^{\dagger}_{q_0\bar{q}_1 g_2\to q_{0'}\bar{q}_{1'}} + \cdots \right\}. \tag{12} $$

Here, we have only written out states that are needed for the full NLO cross section. In addition, marked with …, there are other states needed at higher orders and non-QCD Fock states, including the photon. In fact, for the radiative corrections that we calculate in this paper, only the leading order terms in Eqs. (11) and (12) are needed. Strictly speaking the final state wave functions are not exactly hermitian conjugates of initial state ones, but differ by the sign of the $i\varepsilon$ in the energy denominators [see Eqs. (4.2) and (4.3) of Ref. [76]]. This difference does not play a role in the calculation of this paper, but will be crucial in the calculation of the full NLO corrections to diffractive DIS, that we leave for a future publication. The notation $\tilde{\Sigma}$ denotes the sum over the quantum numbers of each parton in the Fock state and a mixed space phase-space integration for each parton [31,65]. Here the prime $'$ in the sum denotes the fact that the original state is not included in the sum [76].[2] For brevity the transverse coordinates and flavor (f) and helicity (h) indices are not written down explicitly here. We will discuss the color structure of the final state explicitly below. The γ^*_λ-state renormalization coefficient is $Z_{\gamma^*_\lambda} = 1 + \mathcal{O}(e^2)$ and so it can be dropped in this work. The renormalization coefficients $Z_{q,\bar{q},g}$ of the partonic states are of the order of $1 + \mathcal{O}(g^2)$, and as such do not affect the tree-level NLO corrections that we discuss in this paper.

In the mixed space $\mathbf{x}_i$ are the transverse coordinates and k_i^+ the longitudinal momenta of the partons, and indices $i = 0, 1$ refer to the quark and the antiquark, and $i = 2$ to the gluon. The quark, antiquark, and gluon creation and annihilation operators satisfy the (anti-)commutation relations

$$ [a(k_0^+,\mathbf{x}_0,\lambda_0,a_0), a^\dagger(k_1^+,\mathbf{x}_1,\lambda_1,a_1)] = (2k_0^+)(2\pi)\delta(k_0^+ - k_1^+)\delta^{(2)}(\mathbf{x}_0-\mathbf{x}_1)\delta_{\lambda_0,\lambda_1}\delta_{a_0,a_1}, \tag{13} $$

$$ \{b(k_0^+,\mathbf{x}_0,h_0,\alpha_0), b^\dagger(k_1^+,\mathbf{x}_1,h_1,\alpha_1)\} = (2k_0^+)(2\pi)\delta(k_0^+ - k_1^+)\delta^{(2)}(\mathbf{x}_0-\mathbf{x}_1)\delta_{h_0,h_1}\delta_{\alpha_0,\alpha_1}, \tag{14} $$

$$ \{d(k_0^+,\mathbf{x}_0,h_0,\alpha_0), d^\dagger(k_1^+,\mathbf{x}_1,h_1,\alpha_1)\} = (2k_0^+)(2\pi)\delta(k_0^+ - k_1^+)\delta^{(2)}(\mathbf{x}_0-\mathbf{x}_1)\delta_{h_0,h_1}\delta_{\alpha_0,\alpha_1}. \tag{15} $$

Here a_i and λ_i refer to the gluon color and polarization, respectively.

In Eq. (10) the functions $\tilde{\Psi}_{\gamma^*_\lambda\to q_0\bar{q}_1}$ and $\tilde{\Psi}_{\gamma^*_\lambda\to q_0\bar{q}_1 g_2}$ are the light front wave functions (LFWFs) describing the perturbative $\gamma^*\to q\bar{q}$ and $\gamma^*\to q\bar{q}g$ splittings. Furthermore it is convenient to factor out the overall color factor, momentum conservation and dependence on the photon transverse momentum $\mathbf{q}$, and define the reduced wave functions $\tilde{\psi}_{\gamma^*_\lambda\to q_0\bar{q}_1}$ and $\tilde{\psi}_{\gamma^*_\lambda\to q_0\bar{q}_1 g_2}$ as

$$ \tilde{\Psi}_{\gamma^*_\lambda\to q_0\bar{q}_1} = (2q^+)2\pi\delta(k_0^+ + k_1^+ - q^+) e^{i\frac{\mathbf{q}}{q^+}\cdot(k_0^+\mathbf{x}_0 + k_1^+\mathbf{x}_1)} \times \mathbf{1}_{\alpha_0\alpha_1}\tilde{\psi}_{\gamma^*_\lambda\to q_0\bar{q}_1}, \tag{16} $$

$$ \tilde{\Psi}_{\gamma^*_\lambda\to q_0\bar{q}_1 g_2} = (2q^+)2\pi\delta(k_0^+ + k_1^+ + k_2^+ - q^+) \times e^{i\frac{\mathbf{q}}{q^+}(k_0^+\mathbf{x}_0 + k_1^+\mathbf{x}_1 + k_2^+\mathbf{x}_2)} t^a_{\alpha_0\alpha_1}\tilde{\psi}_{\gamma^*_\lambda\to q_0\bar{q}_1 g_2}. \tag{17} $$

These LFWFs are currently available in the literature. The lowest order $\tilde{\psi}_{\gamma^*_\lambda\to q_0\bar{q}_1}$ is a standard result [81]. Loop corrections to it, as well as tree-level wave function describing the $\gamma^*\to q\bar{q}g$ splitting have been recently calculated in mixed space (and d dimensions) in Refs. [30,32,38–40]; other results derived in momentum space include Refs. [82–85], and some are compatible with BFKL evolution [28,29] but not the gluon saturation regime.

[2]What exactly counts as the "same state" requires a more detailed discussion in the case of a two-particle state than for one particle; we plan to return to this in a future paper in the context of the full NLO calculation where this term is needed. In practice, only the diagrams including self-energy corrections on asymptotic external legs need to be excluded, since they are already taken into account thanks to the overall renormalization constants.

In this work we consider diffractive scattering, which requires that the final state partons must be in a color singlet state (see also Ref. [86]). For both $q\bar{q}$ and $q\bar{q}g$ systems there exists exactly one such a configuration. These states are

$$|q_0\bar{q}_1\rangle_D^{\text{singlet}} = \frac{\delta_{\beta_0,\beta_1}}{\sqrt{N_c}}|q_0(\beta_0)\bar{q}_1(\beta_1)\rangle_D, \quad \text{and} \tag{18}$$

$$|q_0\bar{q}_1 g_2\rangle_D^{\text{singlet}} = \frac{t^b_{\beta_0\beta_1}}{\sqrt{C_F N_c}}|q_0(\beta_0)\bar{q}_1(\beta_1)g_2(b)\rangle_D. \tag{19}$$

Here β_0, β_1 are the quark and antiquark colors and b is the gluon color, and $\sqrt{N_c}$ and $\sqrt{C_F N_c}$ are normalization factors.

Using the virtual photon Fock states it now becomes possible to express the matrix elements written in Eq. (8) in terms of the eikonal scattering operators $\hat{S}_E$ describing a color rotation of a quark or a gluon in the target color field. For the $n = q\bar{q}$ Fock state we have

$$\begin{aligned}&{}_D^{\text{singlet}}\langle\bar{q}_1 q_0|(\hat{S}_E - \mathbf{1})|\gamma^*_\lambda(q^+,\mathbf{q};Q^2)\rangle_D \\ &\quad = (2q^+)2\pi\delta(p_0^+ + p_1^+ - q^+)i\mathcal{M}^{\text{LO}}_{\gamma^*\to q\bar{q}},\end{aligned} \tag{20}$$

where the superscript LO refers to the fact that we do not include loop corrections to the $\gamma^* \to q\bar{q}$ splitting in this work. Similarly for $n = q\bar{q}g$ we can write

$$\begin{aligned}&{}_D^{\text{singlet}}\langle g_2\bar{q}_1 q_0|(\hat{S}_E - \mathbf{1})|\gamma^*_\lambda(q^+,\mathbf{q};Q^2)\rangle_D \\ &\quad = (2q^+)2\pi\delta(p_0^+ + p_1^+ + p_2^+ - q^+)i\mathcal{M}^{\text{NLO}}_{\gamma^*\to q\bar{q}g}.\end{aligned} \tag{21}$$

The eikonal scattering operator $\hat{S}_E$ acts on the bare quark, antiquark, and gluon creation operators as

$$\hat{S}_E\tilde{a}^\dagger(k^+,\mathbf{x},\lambda,a) = U_A(\mathbf{x})_{ba}\tilde{a}^\dagger(k^+,\mathbf{x},\lambda,b)\hat{S}_E, \tag{22}$$

$$\hat{S}_E\tilde{b}^\dagger(k^+,\mathbf{x},h,\alpha) = U_F(\mathbf{x})_{\beta\alpha}\tilde{b}^\dagger(k^+,\mathbf{x},h,\beta)\hat{S}_E, \tag{23}$$

$$\hat{S}_E\tilde{d}^\dagger(k^+,\mathbf{x},h,\alpha) = [U_F^\dagger(\mathbf{x})]_{\alpha\beta}\tilde{d}^\dagger(k^+,\mathbf{x},h,\beta)\hat{S}_E. \tag{24}$$

Here α (a) is the quark (gluon) color before the shock and β (b) after, and $U_{F(A)}(\mathbf{x})$ refer to the Wilson lines at transverse coordinate $\mathbf{x}$ in the fundamental (adjoint) representation, describing a color rotation of the quark (gluon) state when it propagates eikonally through the shockwave.

Using Eq. (6) the leading order diffractive cross section can be written in terms of the scattering amplitude $\mathcal{M}^{\text{LO}}_{\gamma^*\to q\bar{q}}$

$$\begin{aligned}\mathrm{d}\sigma^{\text{D,LO}}_{\gamma^*_\lambda\to q\bar{q}} &:= (2q^+)2\pi\delta(p_0^+ + p_1^+ - q^+) \\ &\quad\times \widetilde{\mathrm{d}p_0}\,\widetilde{\mathrm{d}p_1}\sum_{h_0,f_0,h_1,f_1}|\mathcal{M}^{\text{LO}}_{\gamma^*\to q\bar{q}}|^2,\end{aligned} \tag{25}$$

where the summation is over the quantum numbers of the produced $q\bar{q}$ state (for which there is exactly one color singlet color configuration as discussed above). Similarly the cross section for diffractive $q\bar{q}g$ production can be written as

$$\begin{aligned}\mathrm{d}\sigma^{\text{D,NLO}}_{\gamma^*_\lambda\to q\bar{q}g\,\text{singlet}} &:= (2q^+)2\pi\delta(p_0^+ + p_1^+ + p_2^+ - q^+) \\ &\quad\times \widetilde{\mathrm{d}p_0}\,\widetilde{\mathrm{d}p_1}\,\widetilde{\mathrm{d}p_2}\sum_{h_0,f_0,h_1,f_1,\lambda_2}|\mathcal{M}_{\gamma^*\to q\bar{q}g}|^2,\end{aligned} \tag{26}$$

where we again sum over the final state quantum numbers with only one possible color configuration.

The Wilson line structure in the scattering amplitude (20) corresponding to diffractive $q\bar{q}$ production now reads

$$\begin{aligned}&\frac{\delta_{\beta_0\beta_1}}{\sqrt{N_c}}\delta_{\alpha_0\alpha_1}[U_F(\mathbf{x}_0)_{\beta_0\alpha_0}U_F^\dagger(\mathbf{x}_1)_{\alpha_1\beta_1} - \delta_{\beta_0\alpha_0}\delta_{\beta_1\alpha_1}] \\ &\quad = \frac{1}{\sqrt{N_c}}[\mathrm{Tr}(U_F(\mathbf{x}_0)U_F^\dagger(\mathbf{x}_1)) - N_c],\end{aligned} \tag{27}$$

where $\mathbf{x}_0$ and $\mathbf{x}_1$ are the quark and antiquark transverse coordinates, respectively.

In this work we consider coherent diffraction in which case the target nucleus does not dissociate and the average over the target color sources is taken at the amplitude level [67] (see also, e.g., Refs. [68–71] for a discussion of the averaging procedure). This target average gives

$$\frac{1}{\sqrt{N_c}}\langle\mathrm{Tr}[U_F(\mathbf{x}_0)U_F^\dagger(\mathbf{x}_1)] - N_c\rangle = \sqrt{N_c}(S_{01} - 1), \tag{28}$$

where $\langle\mathcal{O}\rangle$ denotes the average over the target configurations and we have defined

$$S_{01} \equiv \frac{1}{N_c}\langle\mathrm{Tr}[U_F(\mathbf{x}_0)U_F^\dagger(\mathbf{x}_1)]\rangle. \tag{29}$$

From the complex conjugate amplitude one obtains exactly the same structure but with the Wilson lines evaluated at different transverse coordinates $\bar{\mathbf{x}}_i$. The Wilson line structure at the cross section level then reads

$$N_c(S_{01} - 1)(S^\dagger_{\bar{0}\bar{1}} - 1). \tag{30}$$

The $q\bar{q}g$ production case, Eq. (21), can be considered similarly. The Wilson line structure in the amplitude (21) is

$$\begin{aligned}&\frac{(t^b_{\beta_0\beta_1})^*}{\sqrt{C_F N_c}}t^a_{\alpha_0\alpha_1}[U_A(\mathbf{x}_2)^{ba}U_F(\mathbf{x}_0)_{\beta_0\alpha_0}U_F^\dagger(\mathbf{x}_1)_{\alpha_1\beta_1} - \delta_{\alpha_0\beta_0}\delta_{\alpha_1\beta_1}\delta_{ab}] \\ &= \frac{1}{\sqrt{C_F N_c}}[U_A^{ba}(\mathbf{x}_2)\mathrm{Tr}(t^b U_F(\mathbf{x}_0)t^a U_F^\dagger(\mathbf{x}_1)) - C_F N_c].\end{aligned} \tag{31}$$

Note that the factor $t^a_{\alpha_0\alpha_1}$ originates from the $\gamma \to q\bar{q}g$ wave function (17).

Again this needs to be averaged over the target configurations at the amplitude level. Defining

$$S_{012} \equiv \frac{1}{C_F N_c}\langle U_A^{ba}(\mathbf{x}_2)\mathrm{Tr}[t^b U_F(\mathbf{x}_0)t^a U_F^\dagger(\mathbf{x}_1)]\rangle = \frac{N_c}{2C_F}\left(S_{02}S_{12} - \frac{1}{{N_c}^2}S_{01}\right) \tag{32}$$

the Wilson line structure at the cross section level can be written as

$$C_F N_c (S_{012} - 1)(S^\dagger_{\bar{0}\bar{1}\bar{2}} - 1). \tag{33}$$

Here the identity

$$U_A^{ba}(\mathbf{x}_2) = 2\mathrm{Tr}[U_F(\mathbf{x}_2)t^a U_F^\dagger(\mathbf{x}_2)t^b]. \tag{34}$$

was used to express the adjoint Wilson line in terms of the fundamental representation Wilson lines. In Eq. (32) the mean field limit $\langle \mathcal{O}_1\mathcal{O}_2\rangle = \langle \mathcal{O}_1\rangle\langle \mathcal{O}_2\rangle$ was used to express the expectation value of the Wilson lines in terms of the dipole correlators S_{ij}. In particular we note that no higher multipole functions (traces of $n > 2$ Wilson lines) appear even at finite N_c in the mean field limit when we evaluate the diffractive cross section, in contrast to the inclusive two or three jet production case [87]. However at finite N_c the BK [14,16] or JIMWLK [88–95] evolution would introduce an implicit dependence on such correlators.

The dipole scattering amplitude $1 - S_{01}$ satisfies the small-x BK or JIMWLK evolution equation. The necessary nonperturbative input for this evolution (initial condition at moderate x) can be determined by performing a fit to the HERA inclusive structure function data [96–99] as, e.g., in Refs. [34,100–103]. Instead of the BK/JIMWLK evolution one can also use phenomenological parametrizations such as the IPsat [104] model, where again the model parameters can similarly be constrained by HERA data [105,106].

Although superficially different, our formulation here is equivalent to the "outgoing state" formulation used, e.g., in Refs. [78,107]. The generic amplitude (8) is given by a matrix element between two dressed states $\mathcal{M} \sim {}_D\langle \mathrm{out}|\hat{S} - 1|\mathrm{in}\rangle_D$. The outgoing state approach consists of first expressing the incoming dressed state $|\mathrm{in}\rangle_D$ in terms of the bare Fock states just like we do (formally expressed as a time evolution operator acting on a bare asymptotic state). One then passes through the shockwave, and obtains the state $(\hat{S} - 1)|\mathrm{in}\rangle_D$ also in terms of the bare states. In the outgoing state formulation one then, instead of taking a matrix element with the dressed state ${}_D\langle \mathrm{out}|$, first inverts its Fock state expansion, expressing the bare states at the shockwave in terms of the dressed states. Then inserting this inverse Fock state expression into the expression $(\hat{S} - 1)|\mathrm{in}\rangle_D$, one obtains the outgoing state $(\hat{S} - 1)|\mathrm{in}\rangle_D$ in terms of the dressed asymptotic (future) states. From here one can either read off the amplitudes, or first square them and think of the projection operator $|\mathrm{out}\rangle_{DD}\langle \mathrm{out}|$ as a particle number operator counting what are in Refs. [78,107] called bare particles at $x^+ = \infty$, which we would here call dressed particles. In other words, in the outgoing state formalism one is acting with the time evolution operator from $x^+ = 0$ to $x^+ \to \infty$ on the state $(\hat{S} - 1)|\mathrm{in}\rangle_D$ to express it in terms of the states $|\mathrm{out}\rangle_D$, whereas here we use the inverse time evolution operator from $x^+ \to \infty$ to $x^+ = 0$ to get ${}_D\langle \mathrm{out}|$ in terms of the bare states at $x^+ = 0$. The connection is easiest to see in terms of the diagrams for the amplitude, which are the same in both approaches and lead to the same expressions.[3]

V. FINAL STATE PHASE SPACE

The diffractive structure functions are measured at fixed invariant mass M_X^2. On the other hand, the diffractive 2- or 3-parton production cross sections in Eqs. (25) and (26) are written in terms of the three-momenta of the quarks and gluons. These momenta are, in turn, obtained by Fourier-transformation from coordinate space, which is how we understand the interaction with the target color field. The only place where the momenta of the final state particles appear is in the exponentials of this Fourier-transform and the delta function setting the invariant mass to M_X^2. In order to get the final diffractive structure function one needs to integrate over the final state momenta with the restriction on M_X. It can be convenient to do this before integrating over the coordinates of the particles. This results in generic "transfer functions" from a coordinate space squared amplitude to the final states with mass M_X. These functions, one for the two- and another one for the three-particle final states, are the same for all states with the same number of partons. Thus, it makes sense to calculate them separately. Here, we will consider the two- and three-particle phase space integrals separately in Secs. V A and V B.

A. Two-particle phase space

In terms of the reduced wave function $\tilde{\psi}_{\gamma^*_\lambda \to q_0\bar{q}_1} := \tilde{\psi}_{\gamma^*_\lambda \to q_0\bar{q}_1}(\mathbf{x}_0, \mathbf{x}_1, z_0, z_1)$ defined in Eq. (16) the diffractive $q\bar{q}$ production cross section (25) reads

$$\frac{\mathrm{d}\sigma^{\mathrm{D}}_{\gamma^*_\lambda \to q\bar{q}(\mathrm{LO})}}{\widetilde{\mathrm{d}p_0}\,\widetilde{\mathrm{d}p_1}} = 4\pi q^+ \delta(p_0^+ + p_1^+ - q^+) N_c \int \mathrm{d}^2\mathbf{x}_0 \int \mathrm{d}^2\mathbf{x}_1 \times \int \mathrm{d}^2\bar{\mathbf{x}}_0 \int \mathrm{d}^2\bar{\mathbf{x}}_1 e^{-i\mathbf{x}_{0\bar{0}}(\mathbf{p}_0 - z_0\mathbf{q})} e^{-i\mathbf{x}_{1\bar{1}}(\mathbf{p}_1 - z_1\mathbf{q})} \times \sum_{h_0,h_1,f} (\tilde{\psi}_{\gamma^*_\lambda \to q_0\bar{q}_{\bar{1}}})^\dagger (\tilde{\psi}_{\gamma^*_\lambda \to q_0\bar{q}_{\bar{1}}})[S^\dagger_{\bar{0}\bar{1}} - 1][S_{01} - 1], \tag{35}$$

[3]Note, however, that Refs. [78,107] use a different normalization for single particle states and for phase space integrals.

where we have written $S_{ij} = S(\mathbf{x}_i, \mathbf{x}_j)$ and $\mathbf{x}_{ij} = \mathbf{x}_i - \mathbf{x}_j$. The virtual photon polarization is denoted by λ, and the transverse coordinates in the complex conjugate amplitude are $\bar{\mathbf{x}}_i$. The overall color factor N_c is obtained when performing the color algebra in the final state requiring that the outgoing state is a color singlet, see Eq. (30).

In order to obtain the diffractive cross section at fixed invariant mass M_X and squared momentum transfer t, we need to integrate over the three-momenta $\vec{p}_0$ and $\vec{p}_1$ and introduce delta functions that enforce the required kinematics. Towards this goal we define the following transverse momentum variables

$$\boldsymbol{\Delta} \equiv \mathbf{p}_0 + \mathbf{p}_1 - \mathbf{q}, \tag{36}$$

$$\mathbf{l} \equiv z_1 \mathbf{p}_0 - z_0 \mathbf{p}_1, \tag{37}$$

which should be understood as the total momentum transfer from the target to the diffractive system, and the relative momentum of the quark-antiquark pair.[4] These satisfy

$$\boldsymbol{\Delta}^2 \approx -t, \tag{38}$$

$$\frac{\mathbf{l}^2}{z_0 z_1} \equiv M_X^2, \tag{39}$$

and neatly the Jacobian is unity, i.e., in (25) we may replace $\mathrm{d}^3\vec{p}_0 \mathrm{d}^3\vec{p}_1 \mapsto (q^+)^2 \mathrm{d}^2\boldsymbol{\Delta}\mathrm{d}^2\mathbf{l}\mathrm{d}z_0\mathrm{d}z_1$ with $z_i := p_i^+/q^+$. With this change of variables, and imposing the M_X and t constraints, the diffractive cross section (25) becomes

$$\begin{aligned}\frac{\mathrm{d}\sigma^{\mathrm{D}}_{\lambda,q\bar{q}}}{\mathrm{d}M_X^2 \mathrm{d}|t|} = \frac{N_c}{4\pi}\int_0^1 \mathrm{d}z_0 \int_0^1 \mathrm{d}z_1 \delta(z_0 + z_1 - 1) \\ \times \int \mathrm{d}^2\mathbf{x}_0 \int \mathrm{d}^2\mathbf{x}_1 \int \mathrm{d}^2\bar{\mathbf{x}}_0 \int \mathrm{d}^2\bar{\mathbf{x}}_1 \\ \times \mathcal{I}^{(2)}_{\boldsymbol{\Delta}} \mathcal{I}^{(2)}_{M_X} \sum_f \sum_{h_0,h_1} (\tilde{\psi}_{\gamma^*_\lambda \to q_0\bar{q}_{\bar{1}}})^\dagger (\tilde{\psi}_{\gamma^*_\lambda \to q_0\bar{q}_1}) \\ \times [S^\dagger_{\bar{0}\,\bar{1}} - 1][S_{01} - 1],\end{aligned} \tag{40}$$

where we defined

$$\mathcal{I}^{(2)}_{\boldsymbol{\Delta}} := \int \frac{\mathrm{d}^2\boldsymbol{\Delta}}{(2\pi)^2} \delta(\boldsymbol{\Delta}^2 - |t|) e^{i\boldsymbol{\Delta}\cdot(z_0\mathbf{x}_{\bar{0}0} - z_1\mathbf{x}_{\bar{1}1})} \tag{41}$$

and

$$\mathcal{I}^{(2)}_{M_X} := \int \frac{\mathrm{d}^2\mathbf{l}}{(2\pi)^2} \delta(\mathbf{l}^2 - z_0 z_1 M_X^2) e^{i\mathbf{l}\cdot(\mathbf{x}_{\bar{0}\,\bar{1}} - \mathbf{x}_{01})}. \tag{42}$$

[4]In fact, in the transverse plane the light cone coordinates correspond to a two-dimensional nonrelativistic system, where p^+ plays the role of a mass. The definition of the relative momentum can then be thought of as $\mathbf{l} \sim \mathbf{p}_0/z_0 - \mathbf{p}_1/z_1$, which is the nonrelativistic or Galileian *velocity* of the particles in the rest frame of the two-particle system.

As the reduced photon wave function $\tilde{\psi}_{\gamma^*_\lambda \to q_0\bar{q}_1}$ can only depend on the coordinate separation $\mathbf{r} := \mathbf{x}_0 - \mathbf{x}_1$, it is useful to make a change of variables from $\mathbf{x}_0$, $\mathbf{x}_1$ to the dipole size $\mathbf{r}$ and the impact parameter $\mathbf{b} = (\mathbf{x}_0 + \mathbf{x}_1)/2$ (using again coordinates with a bar for the complex conjugate amplitude). In these coordinates we get

$$\mathcal{I}^{(2)}_{\boldsymbol{\Delta}} = \int \frac{\mathrm{d}^2\boldsymbol{\Delta}}{(2\pi)^2} \delta(\boldsymbol{\Delta}^2 - |t|) e^{i\boldsymbol{\Delta}\cdot(\bar{\mathbf{b}} - \mathbf{b} + \frac{2z_0-1}{2}(\bar{\mathbf{r}} - \mathbf{r}))}, \tag{43}$$

$$\mathcal{I}^{(2)}_{M_X} = \int \frac{\mathrm{d}^2\mathbf{l}}{(2\pi)^2} \delta(\mathbf{l}^2 - z_0 z_1 M_X^2) e^{i\mathbf{l}\cdot(\bar{\mathbf{r}} - \mathbf{r})}. \tag{44}$$

In the most general case (i.e., without further assumptions), it is also possible to perform the $\mathbf{l}$ integral which gives

$$\mathcal{I}^{(2)}_{M_X} = \frac{1}{4\pi} \mathrm{J}_0(\sqrt{z_0 z_1} M_X \|\bar{\mathbf{r}} - \mathbf{r}\|). \tag{45}$$

This transfer function is related to the probability to form a final state with the given invariant mass M_X given the dipole sizes $\mathbf{r}$ and $\bar{\mathbf{r}}$ in the amplitude and conjugate amplitude with fixed longitudinal momentum fractions z_i for the quarks. The integral over $\boldsymbol{\Delta}$ is of the same form, and gives

$$\mathcal{I}^{(2)}_{\boldsymbol{\Delta}} = \frac{1}{4\pi} \mathrm{J}_0\left(\sqrt{|t|} \left\| \bar{\mathbf{b}} - \mathbf{b} + \frac{(2z_0 - 1)}{2}(\bar{\mathbf{r}} - \mathbf{r}) \right\|\right). \tag{46}$$

Eventually we are also interested in t-integrated diffractive cross sections and structure functions. Integrating over the squared momentum transfer t we find

$$\int_{-\infty}^{0} \mathrm{d}t\, \mathcal{I}^{(2)}_{\boldsymbol{\Delta}} = \delta^{(2)}\left(\bar{\mathbf{b}} - \mathbf{b} + \frac{2z_0 - 1}{2}(\bar{\mathbf{r}} - \mathbf{r})\right). \tag{47}$$

The most general result for the total diffractive cross section in the case where the final state consists of two particles is then given by Eq. (40) with the phase space integrals $\mathcal{I}^{(2)}_{\boldsymbol{\Delta}}$ and $\mathcal{I}^{(2)}_{M_X}$ given above. In particular, we emphasize that the cross section (40) cannot be, in general, written in a factorized form commonly used in the literature [61,62,108], where the result is expressed as a square of an integral over the transverse coordinates in the amplitude. In Sec. VI B we discuss in detail the approximations necessary to obtain a form for the diffractive cross sections where the dependence on impact parameter, amplitude coordinates, and conjugate amplitude coordinates factorize and enables one to write the result in the "squared integral" form.

We finally note that the "off-forward" phase in the amplitude coupling the dipole size and the momentum

transfer is $\exp\left(i\frac{2z_0-1}{2}\boldsymbol{\Delta}\cdot\mathbf{r}\right)$ as shown in Eq. (43), and not $\exp\left(i(1-z_0)\mathbf{r}\cdot\boldsymbol{\Delta}\right)$ as has been commonly used in the literature based on Refs. [109,110]. The correct phase factor has been discussed, e.g., in Refs. [111,112]. We furthermore note that if one used the center-of-mass $\mathbf{b}' = z_0\mathbf{x}_0 + z_1\mathbf{x}_1$ as an impact parameter (which would be a natural variable in light cone perturbation theory), no such off-forward phase would appear and the phase factor in Eq. (43) would be just $\exp\left(i\boldsymbol{\Delta}\cdot(\bar{\mathbf{b}}' - \mathbf{b}')\right)$. While $\mathbf{b}'$ is the conjugate variable to the momentum transfer, it depends on the probe. The definition $\mathbf{b}$ is more natural in a context where one wants to discuss the geometrical transverse structure of the target in a probe-independent manner. Thus it is often used in the literature.

B. Three-particle phase space

Let us next consider the phase space integral for the case where there are three particles in the final state, referring to the $q\bar{q}g$ system in this work. Analogously to the two-particle case discussed above, the starting point is the total diffractive cross section [see Eq. (26)] at fixed M_X^2 and t:

$$\begin{aligned}\frac{\mathrm{d}\sigma^{\mathrm{D}}_{\gamma^*_\lambda\to q\bar{q}g}}{\mathrm{d}M_X^2\mathrm{d}|t|} &= \frac{N_c C_F}{(4\pi)^2}\int\frac{\mathrm{d}^2\mathbf{p}_0}{(2\pi)^2}\int\frac{\mathrm{d}^2\mathbf{p}_1}{(2\pi)^2}\int\frac{\mathrm{d}^2\mathbf{p}_2}{(2\pi)^2}\int_0^1\frac{\mathrm{d}z_0}{z_0}\int_0^1\frac{\mathrm{d}z_1}{z_1}\int_0^1\frac{\mathrm{d}z_2}{z_2}\delta(z_0+z_1+z_2-1)\delta((\mathbf{p}_0+\mathbf{p}_1+\mathbf{p}_2-\mathbf{q})^2-|t|)\\
&\times\delta((p_0+p_1+p_2)^2-M_X^2)\int_{\mathbf{x}_0}\int_{\mathbf{x}_1}\int_{\mathbf{x}_2}\int_{\bar{\mathbf{x}}_0}\int_{\bar{\mathbf{x}}_1}\int_{\bar{\mathbf{x}}_2}(2\pi)^6 e^{i\mathbf{x}_{\bar{0}0}(\mathbf{p}_0-z_0\mathbf{q})}e^{i\mathbf{x}_{\bar{1}1}(\mathbf{p}_1-z_1\mathbf{q})}e^{i\mathbf{x}_{\bar{2}2}(\mathbf{p}_2-z_2\mathbf{q})}\\
&\times\sum_{f,h_0,h_1,\lambda_2}(\tilde{\psi}_{\gamma^*_\lambda\to q_{\bar{0}}\bar{q}_{\bar{1}}g_{\bar{2}}})^\dagger(\tilde{\psi}_{\gamma^*_\lambda\to q_0\bar{q}_1g_2})[S^\dagger_{\bar{0}\bar{1}\bar{2}}-1][S_{012}-1],\end{aligned}\tag{48}$$

where the color factor $\mathrm{tr}(t^a t^a) = N_c C_F$ is again obtained as shown in Eq. (33) in Sec. IV, p_0, p_1, p_2 are the four-momenta of the produced partons, and their transverse coordinates are again labeled as $\mathbf{x}_0$, $\mathbf{x}_1$, $\mathbf{x}_2$ in the amplitude, and $\bar{\mathbf{x}}_0, \bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2$ in the conjugate amplitude. The plus momentum fractions are again denoted by z_i, and the transverse integral normalization is defined as $\int_{\mathbf{x}} \coloneqq \int\frac{\mathrm{d}^2\mathbf{x}}{2\pi}$. Note that this introduces an explicit $(2\pi)^6$ in Eq. (48), but will lead to nicer expressions in the end.

Next we define the following transverse momenta

$$\mathbf{P}_i \coloneqq \mathbf{p}_i - z_i\mathbf{q},\tag{49}$$

$$\mathbf{K} \coloneqq \mathbf{P}_2 - \frac{z_2}{z_0}\mathbf{P}_0,\tag{50}$$

$$\mathbf{P} \coloneqq \mathbf{P}_0 + z_0\boldsymbol{\Delta} + \frac{z_0}{1-z_1}\mathbf{K},\tag{51}$$

$$\boldsymbol{\Delta} \coloneqq \mathbf{q} - \mathbf{p}_0 - \mathbf{p}_1 - \mathbf{p}_2 = -\mathbf{P}_0 - \mathbf{P}_1 - \mathbf{P}_2.\tag{52}$$

Here $\mathbf{P}_i$ could be interpreted as the momentum of the particle i with respect to the center of mass of the three-particle system before the scattering (i.e., the momentum $\mathbf{q}$). The momentum $\boldsymbol{\Delta}$ is then the total momentum transfer from the target to the scattering system, and $\mathbf{K}$ the relative momentum of the gluon with respect to the quark. The remaining $\mathbf{P}$ then is proportional to the relative momentum of the antiquark with respect to the quark-gluon system, which becomes more obvious if one writes it as $\mathbf{P} = z_0 z_1[(\mathbf{P}_0+\mathbf{P}_2)/(1-z_1) - \mathbf{P}_1/z_1]$. Using the above variables allows us to rewrite the invariant mass of the final state particles in a simple way as the sum of two squared momenta:

$$M^2_{q\bar{q}g} \coloneqq (p_0+p_1+p_2)^2 = \frac{\mathbf{P}_0^2}{z_0} + \frac{\mathbf{P}_1^2}{z_1} + \frac{\mathbf{P}_2^2}{z_2} - \boldsymbol{\Delta}^2 = \frac{1-z_1}{z_1 z_0^2}\mathbf{P}^2 + \frac{z_0}{z_2(1-z_1)}\mathbf{K}^2.\tag{53}$$

Now we need to apply the same changes to the exponential phases in the integral (48):

$$e^{i\mathbf{x}_{\bar{0}0}(\mathbf{p}_0-z_0\mathbf{q})}e^{i\mathbf{x}_{\bar{1}1}(\mathbf{p}_1-z_1\mathbf{q})}e^{i\mathbf{x}_{\bar{2}2}(\mathbf{p}_2-z_2\mathbf{q})} = e^{i(\mathbf{x}_{\bar{0}0}+\frac{z_2}{z_0}\mathbf{x}_{\bar{2}2}-\frac{z_0+z_2}{z_0}\mathbf{x}_{\bar{1}1})\cdot\mathbf{P}}e^{i\frac{z_0}{1-z_1}(\mathbf{x}_{\bar{2}2}-\mathbf{x}_{\bar{0}0})\cdot\mathbf{K}}e^{-i(z_0\mathbf{x}_{\bar{0}0}+z_1\mathbf{x}_{\bar{1}1}+z_2\mathbf{x}_{\bar{2}2})\cdot\boldsymbol{\Delta}}.\tag{54}$$

To obtain the cross section differentially in invariant mass and squared momentum transfer, we again integrate over all three-momenta and include delta functions that impose the required constraints. This gives

$$\frac{\mathrm{d}\sigma^{\mathrm{D}}_{\gamma^*_\lambda\to q\bar{q}g}}{\mathrm{d}M_X^2\mathrm{d}|t|} = 4\pi^4 N_\mathrm{c} C_\mathrm{F}\int_0^1\frac{\mathrm{d}z_0}{z_0}\int_0^1\frac{\mathrm{d}z_1}{z_1}\int_0^1\frac{\mathrm{d}z_2}{z_2}\delta(z_0+z_1+z_2-1)\mathcal{I}^{(3)}_{\boldsymbol{\Delta}}\mathcal{I}^{(3)}_{M_X} \times\int_{\mathbf{x}_0}\int_{\mathbf{x}_1}\int_{\mathbf{x}_2}\int_{\bar{\mathbf{x}}_0}\int_{\bar{\mathbf{x}}_1}\int_{\bar{\mathbf{x}}_2}\sum_{h_0,h_1,\lambda_2}(\tilde{\psi}_{\gamma^*_\lambda\to q_{\bar{0}}\bar{q}_{\bar{1}}g_{\bar{2}}})^\dagger(\tilde{\psi}_{\gamma^*_\lambda\to q_0\bar{q}_1 g_2})[1-S^\dagger_{\bar{0}\,\bar{1}\,\bar{2}}][1-S_{012}], \tag{55}$$

where we have again separated the transverse momentum integrals:

$$\mathcal{I}^{(3)}_{\boldsymbol{\Delta}} = \int\frac{\mathrm{d}^2\boldsymbol{\Delta}}{(2\pi)^2}\delta(\boldsymbol{\Delta}^2-|t|)e^{-i(z_0\mathbf{x}_{\bar{0}0}+z_1\mathbf{x}_{\bar{1}1}+z_2\mathbf{x}_{\bar{2}2})\cdot\boldsymbol{\Delta}}, \tag{56}$$

$$\mathcal{I}^{(3)}_{M_X} = \int\frac{\mathrm{d}^2\mathbf{P}}{(2\pi)^2}\int\frac{\mathrm{d}^2\mathbf{K}}{(2\pi)^2}\delta\!\left(\frac{1-z_1}{z_1z_0^2}\mathbf{P}^2+\frac{z_0}{z_2(1-z_1)}\mathbf{K}^2-M_X^2\right)e^{i(\mathbf{x}_{\bar{0}0}+\frac{z_2}{z_0}\mathbf{x}_{\bar{2}2}-\frac{z_0+z_2}{z_0}\mathbf{x}_{\bar{1}1})\cdot\mathbf{P}}e^{i\frac{z_0}{1-z_1}(\mathbf{x}_{\bar{2}2}-\mathbf{x}_{\bar{0}0})\cdot\mathbf{K}}. \tag{57}$$

The integral in $\mathcal{I}^{(3)}_{\boldsymbol{\Delta}}$ can be evaluated using standard methods in spherical coordinates, yielding

$$\mathcal{I}^{(3)}_{\boldsymbol{\Delta}} = \frac{1}{4\pi}\mathrm{J}_0(\sqrt{-t}\|z_0\mathbf{x}_{\bar{0}0}+z_1\mathbf{x}_{\bar{1}1}+z_2\mathbf{x}_{\bar{2}2}\|). \tag{58}$$

Eventually we want to calculate t integrated diffractive structure functions. The integration over the squared momentum transfer t gives

$$\int_{-\infty}^0\mathrm{d}t\mathcal{I}^{(3)}_{\boldsymbol{\Delta}} = \int_{-\infty}^0\mathrm{d}t\int\frac{\mathrm{d}^2\boldsymbol{\Delta}}{(2\pi)^2}\delta(\boldsymbol{\Delta}^2-|t|)e^{-i(z_0\mathbf{x}_{\bar{0}0}+z_1\mathbf{x}_{\bar{1}1}+z_2\mathbf{x}_{\bar{2}2})\cdot\boldsymbol{\Delta}} = \delta^{(2)}(z_0\mathbf{x}_{\bar{0}0}+z_1\mathbf{x}_{\bar{1}1}+z_2\mathbf{x}_{\bar{2}2})\equiv\delta^{(2)}(\bar{\mathbf{b}}-\mathbf{b}), \tag{59}$$

where $\mathbf{b}\coloneqq z_0\mathbf{x}_0+z_1\mathbf{x}_1+z_2\mathbf{x}_2$ is the center of mass of the $q\bar{q}g$ system, and $\bar{\mathbf{b}}$ is the respective coordinate in the conjugate amplitude.[5] The calculation of $\mathcal{I}^{(3)}_{M_X}$ is more involved, and proceeds by Fourier transforming the δ function:

$$\delta\!\left(\frac{1-z_1}{z_1z_0^2}\mathbf{P}^2+\frac{z_0}{z_2(1-z_1)}\mathbf{K}^2-M_X^2\right) = \int_{\mathbb{R}}\frac{\mathrm{d}\eta}{2\pi}e^{i\eta\left(\frac{1-z_1}{z_1z_0^2}\mathbf{P}^2+\frac{z_0}{z_2(1-z_1)}\mathbf{K}^2-M_X^2\right)}. \tag{60}$$

To simplify the notation, we define the following transverse coordinates:

$$\mathbf{z}\coloneqq\mathbf{x}_{\bar{0}0}+\frac{z_2}{z_0}\mathbf{x}_{\bar{2}2}-\frac{z_0+z_2}{z_0}\mathbf{x}_{\bar{1}1}, \tag{61}$$

$$\mathbf{y}\coloneqq\mathbf{x}_{\bar{2}2}-\mathbf{x}_{\bar{0}0}. \tag{62}$$

With these and Eq. (60)—and completing some squares—we can write

$$\mathcal{I}^{(3)}_{M_X} = \int_{\mathbb{R}}\frac{\mathrm{d}\eta}{2\pi}e^{-i\eta M_X^2}e^{-i\eta\frac{z_0^2z_1}{1-z_1}(\frac{\mathbf{z}}{2\eta})^2}e^{-i\eta\frac{z_0z_2}{1-z_1}(\frac{\mathbf{y}}{2\eta})^2}\int\frac{\mathrm{d}^2\mathbf{P}}{(2\pi)^2}\int\frac{\mathrm{d}^2\mathbf{K}}{(2\pi)^2}e^{i\eta\frac{1-z_1}{z_0^2z_1}\left(\mathbf{P}+\frac{z_0^2z_1}{1-z_1}\frac{1}{2\eta}\mathbf{z}\right)^2}e^{i\eta\frac{z_0}{(1-z_1)z_2}(\mathbf{K}+\frac{z_2}{2\eta}\mathbf{y})^2}. \tag{63}$$

Computing the—now Gaussian—transverse momentum integrals requires shifting $\eta\to\eta+i\epsilon$ in the complex plane. With this we have $\int_0^\infty\mathrm{d}z e^{i\eta z} = i/(\eta+i\epsilon)$, and so

[5]Here we directly defined the impact parameter $\mathbf{b}$ as the "true" momentum-weighted impact parameter and not just the average of the coordinates, cf. the discussion below Eq. (47).

$$\int \frac{\mathrm{d}^2\mathbf{P}}{(2\pi)^2}\int \frac{\mathrm{d}^2\mathbf{K}}{(2\pi)^2} e^{i\eta\frac{1-z_1}{z_0^2 z_1}\left(\mathbf{P}+\frac{z_0^2 z_1}{1-z_1}\frac{1}{2\eta}\mathbf{z}\right)^2} e^{i\eta\frac{z_0}{(1-z_1)z_2}(\mathbf{K}+\frac{z_2}{2\eta}\mathbf{y})^2} = \frac{1}{(4\pi)^2}\frac{i\left(\frac{z_0^2 z_1}{1-z_1}\right)}{\eta+i\epsilon}\frac{i\left(\frac{(1-z_1)z_2}{z_0}\right)}{\eta+i\epsilon}. \tag{64}$$

This allows us to write the remaining η integral as a residue in the lower half of the complex plane:

$$\mathcal{I}^{(3)}_{M_X} = i\frac{z_0 z_1 z_2}{(4\pi)^2}\mathrm{Res}\left(\frac{e^{-i\eta M_X^2} e^{-i\eta\frac{z_0^2 z_1}{1-z_1}(\frac{\mathbf{z}}{2\eta})^2} e^{-i\eta\frac{z_0 z_2}{1-z_1}(\frac{\mathbf{y}}{2\eta})^2}}{(\eta+i\epsilon)^2}, \eta\to -i\epsilon\right), \tag{65}$$

where we can now take $\epsilon\to 0$. The singularity at $\eta=0$ is an essential singularity, which means that we can read the above residue as the coefficient of the $\frac{1}{\eta}$ term in the series expansion of the residue function. Defining $\mathbf{Y}_{012}^2 \coloneqq \frac{z_0^2 z_1}{1-z_1}\mathbf{z}^2 + \frac{z_0 z_2}{1-z_1}\mathbf{y}^2$, the expansion is

$$\frac{1}{\eta^2} e^{-i\eta M_X^2} e^{-i\frac{\mathbf{Y}_{012}^2}{4\eta}} = \sum_{n=0}^{\infty}\frac{(-iM_X^2)^n}{n!}\sum_{m=0}^{\infty}\frac{(-i\mathbf{Y}_{012}^2)^m}{m!}\frac{1}{\eta^{2+m-n}}, \tag{66}$$

and so the residue is found at $2+m-n=1\Rightarrow n=m+1$. Thus, we finally have

$$\begin{aligned}\mathcal{I}^{(3)}_{M_X} &= i\frac{z_0 z_1 z_2}{(4\pi)^2}\sum_{m=0}^{\infty}\frac{(-i)^{2m+1}}{m!(m+1)!}(M_X^2)^{m+1}\left(\frac{\mathbf{Y}_{012}^2}{4}\right)^m \\ &= 2\frac{z_0 z_1 z_2}{(4\pi)^2}\frac{M_X}{Y_{012}}\mathrm{J}_1(M_X Y_{012}),\end{aligned} \tag{67}$$

where $Y_{012}\coloneqq\|\mathbf{Y}_{012}\|$, and the series expansion of the Bessel function of the first kind was recognized:

$$\mathrm{J}_1(x) = \sum_{n=0}^{\infty}\frac{(-1)^n}{n!(n+1)!}\left(\frac{x}{2}\right)^{2n+1}. \tag{68}$$

In terms of the quark, antiquark, and gluon coordinates, the transverse distance scale appearing as a conjugate to the invariant mass reads

$$\begin{aligned}\mathbf{Y}_{012}^2 = {}& z_0 z_1(\mathbf{x}_{\bar{0}0}-\mathbf{x}_{\bar{1}1})^2 + z_1 z_2(\mathbf{x}_{\bar{2}2}-\mathbf{x}_{\bar{1}1})^2 \\ &+ z_0 z_2(\mathbf{x}_{\bar{2}2}-\mathbf{x}_{\bar{0}0})^2.\end{aligned} \tag{69}$$

Note that $\mathbf{Y}_{012}^2$ does not depend on the center-of-mass of the $q\bar{q}g$ system $\mathbf{b}=z_0\mathbf{x}_0+z_1\mathbf{x}_1+z_2\mathbf{x}_2$ (or on $\bar{\mathbf{b}}$), which is the Fourier conjugate to the momentum transfer $\mathbf{\Delta}$. Consequently, if the impact parameter dependence factorizes from the Wilson lines as $S_{012}-1=T(\mathbf{b})(S_{012}-1)$, then the t-integrated diffractive cross section is proportional to $\int\mathrm{d}^2\mathbf{b}|T(\mathbf{b})|^2$.

The transverse momentum integrals $\mathcal{I}^{(3)}_{M_X}$ and $\mathcal{I}^{(3)}_{\mathbf{\Delta}}$ combined with the virtual photon wave functions and $q\bar{q}g$-target scattering amplitudes can now be directly used to calculate the total diffractive cross section at fixed invariant mass M_X^2 using Eq. (55).

VI. LEADING ORDER DIFFRACTIVE CROSS SECTION

In this section we present for completeness a derivation for the $q\bar{q}$ contribution to the leading order diffractive cross section. The calculation is organized as follows. First in Sec. VI A we review the leading order photon wave function describing the $\gamma\to q\bar{q}$ dipole transition, and show the squared wave function needed in the case of DDIS. In Sec. VI B we derive the general leading order result for the diffractive cross sections, after which we discuss in detail what approximations are necessary in order to derive the form commonly used in the literature.

A. Coordinate space wave function

The wave function describing the tree-level $\gamma^*\to q\bar{q}$ splitting is required to evaluate the leading order cross section Eq. (40). In $D=4$ dimensions the virtual photon wave functions in transverse coordinate space read [81] (see also Refs. [31,32])

$$\tilde{\psi}_{\gamma_L^*\to q_0\bar{q}_1} = -\frac{ee_f}{2\pi}\delta_{h_1,-h_0} z_0^{3/2} z_1^{3/2} 2Q\mathrm{K}_0(x_{01}\bar{Q}) \tag{70}$$

for the longitudinal photon, and

$$\begin{aligned}\tilde{\psi}_{\gamma_\lambda^*\to q_0\bar{q}_1} = {}& -i\frac{ee_f}{2\pi}\sqrt{z_0 z_1}[(z_0-z_1)\delta^{ij}+2(-h_1)i\epsilon^{ij}]\delta_{h_1,-h_0} \\ &\times\frac{\epsilon_\lambda^i \mathbf{x}_{01}^j}{x_{01}}\bar{Q}\mathrm{K}_1(x_{01}\bar{Q})\end{aligned} \tag{71}$$

for a transversely polarized photon with polarization λ, with $x_{01}\coloneqq\|\mathbf{x}_{01}\|$ and $\bar{Q}\coloneqq\sqrt{z_0 z_1}Q$. The quark fractional charge is denoted by e_f. Note that our convention to pull out a normalization factor $(2q^+)$ from the definition of the reduced wave function in Eq. (16) allows us to write the wave functions in terms of the longitudinal momentum fractions z_i with no explicit dependence on q^+.

In order to calculate the diffractive cross sections we need the squared wave functions in the case where the quark transverse coordinates are different in the amplitude and in the conjugate amplitude. Summing over the quark helicities these squares read

$$\sum_{h_0,h_1}(\tilde{\psi}_{\gamma_L^*\to q_0\bar{q}_{\bar{1}}})^\dagger(\tilde{\psi}_{\gamma_L^*\to q_0\bar{q}_1}) = 2\frac{e^2e_f^2}{(2\pi)^2}z_0^3z_1^3 4Q^2\mathrm{K}_0(x_{01}\bar{Q})\mathrm{K}_0(x_{\bar{0}\,\bar{1}}\bar{Q}) \tag{72}$$

and

$$\frac{1}{2}\sum_{T\ \mathrm{pol}}\sum_{\lambda}\sum_{h_0,h_1}(\tilde{\psi}_{\gamma_\lambda^*\to q_0\bar{q}_{\bar{1}}})^\dagger(\tilde{\psi}_{\gamma_\lambda^*\to q_0\bar{q}_1}) = \frac{e^2e_f^2}{(2\pi)^2}z_0^2z_1^2((z_0-z_1)^2+1)\frac{\mathbf{x}_{01}\cdot\mathbf{x}_{\bar{0}\,\bar{1}}}{x_{01}x_{\bar{0}\,\bar{1}}}Q^2\mathrm{K}_1(x_{01}\bar{Q})\mathrm{K}_1(x_{\bar{0}\,\bar{1}}\bar{Q}). \tag{73}$$

A more familiar form of the momentum fraction dependence can be obtained noting that $(z_0-z_1)^2+1\equiv 2(z_0^2+z_1^2)$, which holds under the plus-momentum conservation $z_0+z_1=1$.

B. From wave function to diffractive cross section

The total diffractive cross section at leading order can be obtained by substituting the virtual photon wave function in Eq. (40), and using the phase space integrals $\mathcal{I}^{(2)}_{\mathbf{\Delta}}$ and $\mathcal{I}^{(2)}_{M_X}$ given in Eqs. (43) and (45). Let us first consider the case where the virtual photon is longitudinally polarized. We note that although the squared wave function (72) factorizes as $\mathrm{K}_0(x_{01}\bar{Q})\mathrm{K}_0(x_{\bar{0}\,\bar{1}}\bar{Q})$, the diffractive cross section cannot be written simply as a square of an amplitude, since the phase space integral $\mathcal{I}^{(2)}_{M_X}$ mixes the transverse coordinates in the amplitude and in the conjugate amplitude even after an integral over the total momentum transfer, see Eq. (45).

In order to derive the leading order results for the $q\bar{q}$ contribution to the diffractive structure functions commonly used in the literature [61], further approximations are required. In particular, we assume that

(i) The invariant mass M_X^2 or virtuality Q^2 is so large that $\exp[-i\frac{2z-1}{2}\mathbf{\Delta}\cdot(\mathbf{r}-\bar{\mathbf{r}})]\approx 1$ (note that $\mathbf{r}^2,\bar{\mathbf{r}}^2\lesssim 1/Q^2, 1/M_X^2$). In this case, the momentum transfer integral of $\mathcal{I}^{(2)}_{\mathbf{\Delta}}$ gives only $\delta^{(2)}(\mathbf{b}-\bar{\mathbf{b}})$, see Eq. (47). This is also the case if the dipole-proton scattering amplitude depends on the center of mass of the $q\bar{q}$ system $\mathbf{b}'=z_0\mathbf{x}_0+z_1\mathbf{x}_1$ and not on the impact parameter $\mathbf{b}=(\mathbf{x}_0+\mathbf{x}_1)/2$ (see discussion in Sec. V A).

(ii) The dipole-target interaction does not depend on the orientation of the dipole or that of the impact parameter, i.e., $S_{01}\equiv S(\|\mathbf{x}_0-\mathbf{x}_1\|,\|\mathbf{b}\|)=:S_{rb}$. The angular dependence is commonly neglected when the initial condition for the BK evolution is determined by fitting the HERA data [34,100–103] and in parametrizations such as IPsat [104]. However, in general such a (probably weak) angular dependence should exist, see, e.g., Refs. [112–116].

Under these assumptions, the only dependence on the angle $\theta_{\mathbf{r},\bar{\mathbf{r}}}$ between $\mathbf{r}:=\mathbf{x}_{01}$ and $\bar{\mathbf{r}}:=\mathbf{x}_{\bar{0}\,\bar{1}}$ is in the phase space integral $\mathcal{I}^{(2)}_{M_X}$. This integral then gives

$$\int \mathrm{d}\theta_{\mathbf{r},\bar{\mathbf{r}}}\mathcal{I}^{(2)}_{M_X} = \int \mathrm{d}\theta_{\mathbf{r},\bar{\mathbf{r}}}\mathrm{J}_0(\sqrt{z_0z_1}M_X\|\bar{\mathbf{r}}-\mathbf{r}\|) = 2\pi\mathrm{J}_0(\sqrt{z_0z_1}M_X\|\mathbf{r}\|)\mathrm{J}_0(\sqrt{z_0z_1}M_X\|\bar{\mathbf{r}}\|), \tag{74}$$

and as such it factorizes into parts that only depend on transverse coordinates in the amplitude or in the conjugate amplitude. Integrating over t results in a delta function forcing $\mathbf{b}=\bar{\mathbf{b}}$ [see Eq. (47)], and the cross section becomes

$$\frac{\mathrm{d}\sigma^{\mathrm{D}}_{L,q\bar{q}}}{\mathrm{d}M_X^2} = N_{\mathrm{c}}\frac{e^2}{(2\pi)^2}\sum_f e_f^2\int_0^1 \mathrm{d}z_0\, z_0^3(1-z_0)^3\int \mathrm{d}^2\mathbf{b}\left\{\int \mathrm{d}r\, r\mathrm{J}_0(\sqrt{z_0(1-z_0)}M_X r)Q\mathrm{K}_0(r\bar{Q})(S_{rb}-1)\right\}^2, \tag{75}$$

where we have substituted the longitudinal photon wave function summed over helicities shown in Eq. (72). Note that if the impact parameter dependence factorizes from the dipole scattering amplitude, i.e., $(S_{rb}-1)\equiv T(\mathbf{b})[S_r-1]$, then the impact parameter integral completely factorizes and gives $\int \mathrm{d}^2\mathbf{b}|T(\mathbf{b})|^2$.

On the other hand, for the transversely polarized photons, the leading order reduced wave function squared depends on the angle between $\mathbf{r}$ and $\bar{\mathbf{r}}$ as shown in Eq. (73):

$$\sum_{T\mathrm{pol}.\lambda}\sum_{h_0,h_1}(\tilde{\psi}_{\gamma_\lambda^*\to q_0\bar{q}_{\bar{1}}})^\dagger(\tilde{\psi}_{\gamma_\lambda^*\to q_0\bar{q}_1}) \propto \frac{\mathbf{r}\cdot\bar{\mathbf{r}}}{\|\mathbf{r}\|\|\bar{\mathbf{r}}\|}, \tag{76}$$

which means that the part that depends on the dipole sizes $\mathbf{r}$ and $\bar{\mathbf{r}}$—omitting the dipole amplitudes for now—reads

$$\int \mathrm{d}^2\mathbf{r} \int \mathrm{d}^2\bar{\mathbf{r}} \mathcal{I}^{(2)}_{M_X} \frac{\mathbf{r}\cdot\bar{\mathbf{r}}}{\|\mathbf{r}\|\|\bar{\mathbf{r}}\|} = \int \mathrm{d}^2\mathbf{r} \int \mathrm{d}^2\bar{\mathbf{r}} \int \frac{\mathrm{d}^2\mathbf{l}}{(2\pi)^2} \delta(\mathbf{l}^2 - z_0(1-z_0)M_X^2) \frac{\mathbf{r}\cdot\bar{\mathbf{r}}}{\|\mathbf{r}\|\|\bar{\mathbf{r}}\|} e^{i\mathbf{l}\cdot(\bar{\mathbf{r}}-\mathbf{r})}. \tag{77}$$

Parametrizing the angles as $\angle(\mathbf{l},\mathbf{r}) =: \theta$ and $\angle(\mathbf{l},\bar{\mathbf{r}}) =: \bar{\theta}$, we have for the dot product: $\mathbf{r}\cdot\bar{\mathbf{r}} = r\bar{r}(\cos\theta\cos\bar{\theta} + \sin\theta\sin\bar{\theta})$, where the sine term vanishes in the integration. Thus we are left with

$$\begin{aligned} &\int r\,\mathrm{d}r\,\mathrm{d}\theta \int \bar{r}\,\mathrm{d}\bar{r}\,\mathrm{d}\bar{\theta} \int \frac{\mathrm{d}^2\mathbf{l}}{(2\pi)^2} \delta(\mathbf{l}^2 - z_0(1-z_0)M_X^2) \cos\theta\cos\bar{\theta} e^{il\bar{r}\cos\bar{\theta}} e^{-ilr\cos\theta} \\ &= \pi \int r\mathrm{d}r \int \bar{r}\,\mathrm{d}\bar{r}\,\mathrm{J}_1(\sqrt{z_0(1-z_0)}M_X r)\mathrm{J}_1(\sqrt{z_0(1-z_0)}M_X\bar{r}). \end{aligned} \tag{78}$$

Consequently, we find that (only) under the assumptions listed at the beginning of this subsection the diffractive cross section can be written in a factorized form independently of the photon polarization. Otherwise the transverse coordinates in the amplitude and conjugate amplitude are mixed. In the future it will be interesting to study numerically the effect of these assumptions, that were used, e.g., in Ref. [62] where a good description of the HERA diffractive structure function data was obtained. Both cross sections can now under these assumptions be expressed in terms of an auxiliary function, denoting now $z_0 = z$ with the integral over z_1 is performed using the δ function

$$\Phi_n(z,\beta,Q,\mathbf{b}) = \left[\int \mathrm{d}r\, r\mathrm{J}_n(\sqrt{z(1-z)}M_X r)\mathrm{K}_n(\sqrt{z(1-z)}Qr)(S_{rb}-1)\right]^2. \tag{79}$$

Using this, we may write the diffractive structure functions as

$$x_{\mathbb{P}}F^{\mathrm{D}}_{L,q\bar{q}}(\beta,x_{\mathbb{P}},Q^2) = \frac{N_c Q^4}{2\pi^3\beta}\sum e_f^2 \int \mathrm{d}^2\mathbf{b} \int_0^1 \mathrm{d}z\, z^3(1-z)^3 Q^2 \Phi_0(z,\beta,Q,\mathbf{b}), \tag{80}$$

$$x_{\mathbb{P}}F^{\mathrm{D}}_{T,q\bar{q}}(\beta,x_{\mathbb{P}},Q^2) = \frac{N_c Q^4}{8\pi^3\beta}\sum e_f^2 \int \mathrm{d}^2\mathbf{b} \int_0^1 \mathrm{d}z\, z^2(1-z)^2(z^2+(1-z)^2) Q^2 \Phi_1(z,\beta,Q,\mathbf{b}), \tag{81}$$

which is in agreement with Ref. [62], once one accounts for the different normalization of the dipole amplitude and the different integration domain of z.

VII. TREE-LEVEL $q\bar{q}g$ CONTRIBUTION TO THE DIFFRACTIVE STRUCTURE FUNCTIONS

In this section we present the main result of this paper: the tree-level calculation of diffractive $q\bar{q}g$ production as a function of M_X^2 and t. We consider the case where the gluon is emitted before the shockwave and the $q\bar{q}g$ system then interacts with the target, corresponding to the diagrams in Figs. 3(b), 3(c), 4(d), and 4(e). The emission-after-shock contribution could then in principle be obtained by taking the appropriate coordinate limits following the method developed in Refs. [78,107]. As discussed in Sec. III, we will however leave it to a future publication. For simplicity we only consider the massless quark limit in this work.

The diffractive $q\bar{q}g$ production cross section was written in Sec. V B [see Eq. (55)] in terms of the phase space integrals $\mathcal{I}^{(3)}_{M_X}$ and $\mathcal{I}^{(3)}_{\mathbf{\Delta}}$ given in Eqs. (67) and (59) as

$$\begin{aligned} \frac{\mathrm{d}\sigma^{\mathrm{D}}_{\lambda,q\bar{q}g}}{\mathrm{d}M_X^2\mathrm{d}|t|} &= 4\pi^4 N_c C_F \int_0^1 \frac{\mathrm{d}z_0}{z_0}\int_0^1 \frac{\mathrm{d}z_1}{z_1}\int_0^1 \frac{\mathrm{d}z_2}{z_2}\delta(z_0+z_1+z_2-1)\int_{\mathbf{x}_0}\int_{\mathbf{x}_1}\int_{\mathbf{x}_2}\int_{\bar{\mathbf{x}}_0}\int_{\bar{\mathbf{x}}_1}\int_{\bar{\mathbf{x}}_2} \mathcal{I}^{(3)}_{\mathbf{\Delta}}\mathcal{I}^{(3)}_{M_X} \\ &\times \sum_{h_0,h_1,\lambda_2} (\tilde{\psi}_{\gamma^*_\lambda\to q_0\bar{q}_1 g_2})^\dagger (\tilde{\psi}_{\gamma^*_\lambda\to q_0\bar{q}_1 g_2})[S^\dagger_{\bar{0}\bar{1}\bar{2}} - 1][S_{012}-1]. \end{aligned} \tag{82}$$

The only part missing from the diffractive $q\bar{q}g$ production cross section is thus the calculation of the square of the tree-level wave function $\tilde{\psi}_{\gamma^*_\lambda\to q_0\bar{q}_1 g_2}$, with different transverse coordinates in the amplitude ($\mathbf{x}_i$) and in the conjugate amplitude ($\bar{\mathbf{x}}_i$), summed over helicities. The plus momentum fractions z_i are external kinematical variables and therefore the same in the direct and complex conjugate amplitude. We calculate this square using the wave functions for the longitudinally and transversely polarized photons in four dimensions from Ref. [32] (see also Refs. [30,38–40]), with the modification that the factor of $2q^+$ has been taken out of the reduced wave functions $\tilde{\psi}$ in the definition (17). The reduced LFWF for the longitudinal photon reads

$$\tilde{\psi}^{\mathrm{Tree}}_{\gamma^*_L\to q_0\bar{q}_1g_2} = ee_f g\frac{i}{(2\pi)^2}\epsilon^{j*}_{\lambda_2}2Q\mathrm{K}_0(QX_{012})\sqrt{z_0}\sqrt{z_1}\delta_{h_1,-h_0} \times\left\{z_1[(2z_0+z_2)\delta^{jm}-i(2h_0)z_2\epsilon^{jm}]\left(\frac{\mathbf{x}^m_{20}}{\mathbf{x}^2_{20}}\right)-z_0[(2z_1+z_2)\delta^{jm}+i(2h_0)z_2\epsilon^{jm}]\left(\frac{\mathbf{x}^m_{21}}{\mathbf{x}^2_{21}}\right)\right\}, \tag{83}$$

whereas for the transverse photon we have

$$\begin{aligned}\tilde{\psi}^{\mathrm{Tree}}_{\gamma^*_\lambda\to q_0\bar{q}_1g_2} = {}& \frac{ee_f g}{(2\pi)^2}\epsilon^i_\lambda\epsilon^{j*}_{\lambda_2}\sqrt{z_0}\sqrt{z_1}\delta_{h_1,-h_0}\frac{Q}{X_{012}}\mathrm{K}_1(QX_{012})\\ &\times\left\{z_1[(2z_0+z_2)\delta^{jm}-i(2h_0)z_2\epsilon^{jm}][(2z_1-1)\delta^{il}-i(2h_0)\epsilon^{il}]\mathbf{x}^l_{0+2;1}\left(\frac{\mathbf{x}^m_{20}}{\mathbf{x}^2_{20}}\right)\right.\\ &+z_0[(2z_1+z_2)\delta^{jm}+i(2h_0)z_2\epsilon^{jm}][(2z_0-1)\delta^{il}+i(2h_0)\epsilon^{il}]\mathbf{x}^l_{0;1+2}\left(\frac{\mathbf{x}^m_{21}}{\mathbf{x}^2_{21}}\right)\\ &\left.-\frac{z_0z_1z_2}{z_0+z_2}[\delta^{ij}-i(2h_0)\epsilon^{ij}]+\frac{z_0z_1z_2}{z_1+z_2}[\delta^{ij}+i(2h_0)\epsilon^{ij}]\right\},\end{aligned} \tag{84}$$

where X_{012}, $\mathbf{x}_{0+2;1}$ and $\mathbf{x}_{0;1+2}$ are defined as

$$X^2_{012} \coloneqq z_0z_1\mathbf{x}^2_{01}+z_0z_2\mathbf{x}^2_{02}+z_1z_2\mathbf{x}^2_{12}, \tag{85}$$

$$\mathbf{x}_{0+2;1} \coloneqq -\frac{z_0}{z_0+z_2}\mathbf{x}_{20}+\mathbf{x}_{21} = \mathbf{x}_{01}+\frac{z_2}{z_0+z_2}\mathbf{x}_{20}, \tag{86}$$

$$\mathbf{x}_{0;1+2} \coloneqq -\mathbf{x}_{20}+\frac{z_1}{z_1+z_2}\mathbf{x}_{21} = \mathbf{x}_{01}-\frac{z_2}{z_1+z_2}\mathbf{x}_{21}. \tag{87}$$

The quantity $Q^2X^2_{012}$ corresponds to the $q\bar{q}g$ formation time divided by the lifetime of the virtual photon that forms the $q\bar{q}g$ system, as discussed in more detail in Ref. [30]. Configurations with large $Q^2X^2_{012}$ are exponentially suppressed, which enforces the restriction that the $q\bar{q}g$ state must develop within a formation time that is less than the lifetime of the virtual photon.

The calculation of the squared wave functions $\sum_{h_0,h_1,\lambda_2}(\tilde{\psi}_{\gamma^*_\lambda\to q_{\bar{0}}\bar{q}_{\bar{1}}g_{\bar{2}}})^\dagger(\tilde{\psi}_{\gamma^*_\lambda\to q_0\bar{q}_1g_2})$ is cumbersome but straightforward. More technical details are given in Ref. [65]. After a lot of algebra, we obtain the diffractive structure functions

$$\begin{aligned}x_{\mathbb{P}}F^{\mathrm{D(4)NLO}}_{L,q\bar{q}g}(x_{\mathrm{Bj}},Q^2,\beta,t) = {}& 4\frac{\alpha_{\mathrm{s}}N_{\mathrm{c}}C_{\mathrm{F}}Q^4}{\beta}\sum_f e_f^2\int_0^1\frac{\mathrm{d}z_0}{z_0}\int_0^1\frac{\mathrm{d}z_1}{z_1}\int_0^1\frac{\mathrm{d}z_2}{z_2}\delta(z_0+z_1+z_2-1)\\ &\times\int_{\mathbf{x}_0}\int_{\mathbf{x}_1}\int_{\mathbf{x}_2}\int_{\bar{\mathbf{x}}_0}\int_{\bar{\mathbf{x}}_1}\int_{\bar{\mathbf{x}}_2}\mathcal{I}^{(3)}_{M_X}\mathcal{I}^{(3)}_{\boldsymbol{\Delta}}4z_0z_1Q^2\mathrm{K}_0(QX_{012})\mathrm{K}_0(QX_{\bar{0}\bar{1}\bar{2}})\\ &\times\left\{z_1^2\left[(2z_0(z_0+z_2)+z_2^2)\left(\frac{\mathbf{x}_{20}}{\mathbf{x}^2_{20}}\cdot\left(\frac{\mathbf{x}_{\bar{2}\bar{0}}}{\mathbf{x}^2_{\bar{2}\bar{0}}}-\frac{1}{2}\frac{\mathbf{x}_{\bar{2}\bar{1}}}{\mathbf{x}^2_{\bar{2}\bar{1}}}\right)-\frac{1}{2}\frac{\mathbf{x}_{\bar{2}\bar{0}}\cdot\mathbf{x}_{21}}{\mathbf{x}^2_{\bar{2}\bar{0}}\mathbf{x}^2_{21}}\right)+\frac{z_2^2}{2}\left(\frac{\mathbf{x}_{\bar{2}\bar{0}}\cdot\mathbf{x}_{21}}{\mathbf{x}^2_{\bar{2}\bar{0}}\mathbf{x}^2_{21}}+\frac{\mathbf{x}_{20}\cdot\mathbf{x}_{\bar{2}\bar{1}}}{\mathbf{x}^2_{20}\mathbf{x}^2_{\bar{2}\bar{1}}}\right)\right]\right.\\ &\left.+z_0^2\left[(2z_1(z_1+z_2)+z_2^2)\left(\frac{\mathbf{x}_{21}}{\mathbf{x}^2_{21}}\cdot\left(\frac{\mathbf{x}_{\bar{2}\bar{1}}}{\mathbf{x}^2_{\bar{2}\bar{1}}}-\frac{1}{2}\frac{\mathbf{x}_{\bar{2}\bar{0}}}{\mathbf{x}^2_{\bar{2}\bar{0}}}\right)-\frac{1}{2}\frac{\mathbf{x}_{20}\cdot\mathbf{x}_{\bar{2}\bar{1}}}{\mathbf{x}^2_{20}\mathbf{x}^2_{\bar{2}\bar{1}}}\right)+\frac{z_2^2}{2}\left(\frac{\mathbf{x}_{\bar{2}\bar{0}}\cdot\mathbf{x}_{21}}{\mathbf{x}^2_{\bar{2}\bar{0}}\mathbf{x}^2_{21}}+\frac{\mathbf{x}_{20}\cdot\mathbf{x}_{\bar{2}\bar{1}}}{\mathbf{x}^2_{20}\mathbf{x}^2_{\bar{2}\bar{1}}}\right)\right]\right\}\\ &\times[1-S^\dagger_{\bar{0}\bar{1}\bar{2}}][1-S_{012}],\end{aligned} \tag{88}$$

for the longitudinal structure function, and

$$x_{\mathbb{P}}F_{T,q\bar{q}g}^{\mathrm{D(4)NLO}}(x_{\mathrm{Bj}},Q^2,\beta,t)=2\frac{\alpha_{\mathrm{s}}N_{\mathrm{c}}C_{\mathrm{F}}Q^4}{\beta}\sum_f e_f^2\int_0^1\frac{\mathrm{d}z_0}{z_0}\int_0^1\frac{\mathrm{d}z_1}{z_1}\int_0^1\frac{\mathrm{d}z_2}{z_2}\delta(z_0+z_1+z_2-1)$$
$$\times\int_{\mathbf{x}_0}\int_{\mathbf{x}_1}\int_{\mathbf{x}_2}\int_{\bar{\mathbf{x}}_0}\int_{\bar{\mathbf{x}}_1}\int_{\bar{\mathbf{x}}_2}\mathcal{I}_{M_X}^{(3)}\mathcal{I}_{\boldsymbol{\Delta}}^{(3)}\frac{z_0z_1Q^2}{X_{012}X_{\bar{0}\,\bar{1}\,\bar{2}}}\mathrm{K}_1(QX_{012})\mathrm{K}_1(QX_{\bar{0}\,\bar{1}\,\bar{2}})$$
$$\times\{\Upsilon_{\mathrm{reg}}^{(|(b)|^2)}+\Upsilon_{\mathrm{reg}}^{(|(c)|^2)}+\Upsilon_{\mathrm{inst}}^{(d)}+\Upsilon_{\mathrm{inst}}^{(e)}+\Upsilon_{\mathrm{interf}}^{(b)\times(c)}\}[1-S_{\bar{0}\,\bar{1}\,\bar{2}}^{\dagger}][1-S_{012}] \tag{89}$$

for the transverse structure function.

The Υ terms of the squared virtual photon light-front wave function are

$$\Upsilon_{\mathrm{reg}}^{(|(b)|^2)}=z_1^2\Bigg[(2z_0(z_0+z_2)+z_2^2)(1-2z_1(1-z_1))(\mathbf{x}_{\bar{0}+\bar{2};\bar{1}}\cdot\mathbf{x}_{0+2;1})\frac{(\mathbf{x}_{\bar{2}\,\bar{0}}\cdot\mathbf{x}_{20})}{\mathbf{x}_{\bar{2}\,\bar{0}}^2\mathbf{x}_{20}^2}$$
$$-z_2(2z_0+z_2)(2z_1-1)\frac{(\mathbf{x}_{\bar{0}+\bar{2};\bar{1}}\cdot\mathbf{x}_{\bar{2}\,\bar{0}})(\mathbf{x}_{0+2;1}\cdot\mathbf{x}_{20})-(\mathbf{x}_{\bar{0}+\bar{2};\bar{1}}\cdot\mathbf{x}_{20})(\mathbf{x}_{0+2;1}\cdot\mathbf{x}_{\bar{2}\,\bar{0}})}{\mathbf{x}_{\bar{2}\,\bar{0}}^2\mathbf{x}_{20}^2}\Bigg], \tag{90}$$

$$\Upsilon_{\mathrm{reg}}^{(|(c)|^2)}=z_0^2\Bigg[(2z_1(z_1+z_2)+z_2^2)(1-2z_0(1-z_0))(\mathbf{x}_{\bar{0};\bar{1}+\bar{2}}\cdot\mathbf{x}_{0;1+2})\frac{(\mathbf{x}_{\bar{2}\,\bar{1}}\cdot\mathbf{x}_{21})}{\mathbf{x}_{\bar{2}\,\bar{1}}^2\mathbf{x}_{21}^2}$$
$$-z_2(2z_1+z_2)(2z_0-1)\frac{(\mathbf{x}_{\bar{0};\bar{1}+\bar{2}}\cdot\mathbf{x}_{\bar{2}\,\bar{1}})(\mathbf{x}_{0;1+2}\cdot\mathbf{x}_{21})-(\mathbf{x}_{\bar{0};\bar{1}+\bar{2}}\cdot\mathbf{x}_{21})(\mathbf{x}_{0;1+2}\cdot\mathbf{x}_{\bar{2}\,\bar{1}})}{\mathbf{x}_{\bar{2}\,\bar{1}}^2\mathbf{x}_{21}^2}\Bigg], \tag{91}$$

$$\Upsilon_{\mathrm{inst}}^{(d)}=\frac{z_0^2z_1^2z_2^2}{(z_0+z_2)^2}-\frac{z_0^2z_1^3z_2}{z_0+z_2}\left(\frac{\mathbf{x}_{0+2;1}\cdot\mathbf{x}_{20}}{\mathbf{x}_{20}^2}+\frac{\mathbf{x}_{\bar{0}+\bar{2};\bar{1}}\cdot\mathbf{x}_{\bar{2}\,\bar{0}}}{\mathbf{x}_{\bar{2}\,\bar{0}}^2}\right)$$
$$+\frac{z_0^2z_1(z_1+z_2)^2z_2}{z_0+z_2}\left(\frac{\mathbf{x}_{0;1+2}\cdot\mathbf{x}_{21}}{\mathbf{x}_{21}^2}+\frac{\mathbf{x}_{\bar{0};\bar{1}+\bar{2}}\cdot\mathbf{x}_{\bar{2}\,\bar{1}}}{\mathbf{x}_{\bar{2}\,\bar{1}}^2}\right), \tag{92}$$

$$\Upsilon_{\mathrm{inst}}^{(e)}=\frac{z_0^2z_1^2z_2^2}{(z_1+z_2)^2}-\frac{z_0z_1^2(z_0+z_2)^2z_2}{z_1+z_2}\left(\frac{\mathbf{x}_{0+2;1}\cdot\mathbf{x}_{20}}{\mathbf{x}_{20}^2}+\frac{\mathbf{x}_{\bar{0}+\bar{2};\bar{1}}\cdot\mathbf{x}_{\bar{2}\,\bar{0}}}{\mathbf{x}_{\bar{2}\,\bar{0}}^2}\right)$$
$$+\frac{z_0^3z_1^2z_2}{z_1+z_2}\left(\frac{\mathbf{x}_{0;1+2}\cdot\mathbf{x}_{21}}{\mathbf{x}_{21}^2}+\frac{\mathbf{x}_{\bar{0};\bar{1}+\bar{2}}\cdot\mathbf{x}_{\bar{2}\,\bar{1}}}{\mathbf{x}_{\bar{2}\,\bar{1}}^2}\right), \tag{93}$$

$$\Upsilon_{\mathrm{interf}}^{(b)\times(c)}=-z_0z_1[z_1(z_0+z_2)+z_0(z_1+z_2)][z_0(z_0+z_2)+z_1(z_1+z_2)]$$
$$\times\left[(\mathbf{x}_{\bar{0}+\bar{2};\bar{1}}\cdot\mathbf{x}_{0;1+2})\frac{(\mathbf{x}_{\bar{2}\,\bar{0}}\cdot\mathbf{x}_{21})}{\mathbf{x}_{\bar{2}\,\bar{0}}^2\mathbf{x}_{21}^2}+(\mathbf{x}_{\bar{0};\bar{1}+\bar{2}}\cdot\mathbf{x}_{0+2;1})\frac{(\mathbf{x}_{\bar{2}\,\bar{1}}\cdot\mathbf{x}_{20})}{\mathbf{x}_{\bar{2}\,\bar{1}}^2\mathbf{x}_{20}^2}\right]+z_0z_1z_2(z_0-z_1)^2$$
$$\times\Bigg[\frac{(\mathbf{x}_{\bar{0}+\bar{2};\bar{1}}\cdot\mathbf{x}_{\bar{2}\,\bar{0}})(\mathbf{x}_{0;1+2}\cdot\mathbf{x}_{21})-(\mathbf{x}_{\bar{0}+\bar{2};\bar{1}}\cdot\mathbf{x}_{21})(\mathbf{x}_{0;1+2}\cdot\mathbf{x}_{\bar{2}\,\bar{0}})}{\mathbf{x}_{\bar{2}\,\bar{0}}^2\mathbf{x}_{21}^2}$$
$$+\frac{(\mathbf{x}_{\bar{0};\bar{1}+\bar{2}}\cdot\mathbf{x}_{\bar{2}\,\bar{1}})(\mathbf{x}_{0+2;1}\cdot\mathbf{x}_{20})-(\mathbf{x}_{\bar{0};\bar{1}+\bar{2}}\cdot\mathbf{x}_{20})(\mathbf{x}_{0+2;1}\cdot\mathbf{x}_{\bar{2}\,\bar{1}})}{\mathbf{x}_{\bar{2}\,\bar{1}}^2\mathbf{x}_{20}^2}\Bigg]. \tag{94}$$

It might be elucidating to note that the above expressions satisfy the following symmetries in particle exchanges:

$$\Upsilon_{\mathrm{reg}}^{(|(c)|^2)}\equiv\Upsilon_{\mathrm{reg}}^{(|(b)|^2)}(z_0\leftrightarrow z_1,\mathbf{x}_0\leftrightarrow\mathbf{x}_1,\bar{\mathbf{x}}_0\leftrightarrow\bar{\mathbf{x}}_1) \tag{95}$$

$$\Upsilon_{\mathrm{inst}}^{(e)}\equiv\Upsilon_{\mathrm{inst}}^{(d)}(z_0\leftrightarrow z_1,\mathbf{x}_0\leftrightarrow\mathbf{x}_1,\bar{\mathbf{x}}_0\leftrightarrow\bar{\mathbf{x}}_1), \tag{96}$$

making their sum symmetric under the exchange of the quark and antiquark. Meanwhile the $(b)\times(c)$ interference term is already a sum of terms that can be obtained by this exchange, and is thus symmetric by itself.

The results as expressed in Eqs. (88) and (89) are written for the most general transverse coordinate space dependence, and contain a total of six two-dimensional tranverse coordinate integrations. In practical applications, the Wilson line operators are usually available in coordinate space from solutions of the BK equation, thus one way or the other such coordinate integrations have to be present; if the cross section formula was expressed in momentum space, a similar number of transverse integrals would be needed in the Fourier-transformation of the Wilson line

operators to momentum space. More generally, with a 3-parton configuration passing through the shockwave one ultimately has to integrate over either the coordinates of 6 Wilson lines, or over 6 tranverse momenta transferred by the Wilson lines from the target to the scattering partonic system.

Fully evaluating the cross section expressions numerically will be a nontrivial task. One simplification can be provided by an additional assumption of a factorized impact parameter dependence for the Wilson line operator $1-S_{012}$, an approximation that is often made in phenomenological studies. The remaining integrals are nevertheless more demanding than in the previously commonly used large Q^2 limit that we will discuss in Sec. VIII B. The ultimate reason is that the large Q^2 limit reduces the problem to an "effective adjoint dipole" with only two Wilson lines in the amplitude, instead of three. Then further kinematical approximations allow one to write the cross section in terms of the square of a single coordinate integral. In the full case the helicity sums over the produced particles, performed analytically, result in expressions that couple the coordinates in the amplitude and its complex conjugate. Thus a similar factorization is not possible.

The dipole amplitude $N_{ij}=1-S_{ij}$ satisfies the BK or JIMWLK high-energy evolution equation, where the evolution is parametrized in terms of the evolution rapidity determined from the kinematics of the process. Different versions of the evolution equations are factorized in terms of different variables, associated with either light cone energy or momentum [34], and require a consistent treatment in the cross section. We emphasize that the contribution considered in this section is one where the gluon momentum fraction z_2 cannot become small, since it is limited by the finite M_X. This contribution is thus not associated with the high energy evolution of the target. We will have to return to the question of evolution rapidity in more detail when addressing the full NLO diffractive structure function.

These contributions to the diffractive structure functions are finite without requiring any additional cancellations with other diagrams. This is because the invariant mass of the final state is fixed, and thus ultraviolet divergences do not appear. This also means that the divergences in the loop contributions that we are not calculating here should cancel against each other, see discussion in Sec. III. Note that without the M_X^2 restriction the integration over the final state momenta would set $\bar{\mathbf{x}}_i\to\mathbf{x}_i$ and an ultraviolet divergence $\sim\int \mathrm{d}^2\mathbf{x}_{02}/\mathbf{x}_{02}^2$ or $\sim\int \mathrm{d}^2\mathbf{x}_{12}/\mathbf{x}_{12}^2$ would appear. In the diffractive structure functions the corresponding structure is $\sim\int \mathrm{d}^2\mathbf{x}_{i2}\mathbf{x}_{i2}/\mathbf{x}_{i2}^2$ (with $i=0,1$) which is UV finite. The integration over the momentum transfer sets the center of mass of the $q\bar{q}g$ system $\mathbf{b}$ to be the same in the amplitude and in the conjugate amplitude, $\mathbf{b}=\bar{\mathbf{b}}$, but this does not affect the behavior of these integrals in the ultraviolet region (note that $\mathcal{I}^{(3)}_{M_X}$ does not depend on $\mathbf{b}$ or on $\bar{\mathbf{b}}$).

A potential (transverse) infrared divergence is removed, for a gluon emitted before the shockwave, by the dipole amplitude part vanishing for $|\mathbf{x}_{02}|\sim|\mathbf{x}_{12}|\to\infty$. For the emissions after the shockwave that we are not calculating here, these configurations are not suppressed by the Wilson line correlator, and need to cancel against the wave function renormalization of the outgoing quarks. Similarly there is no soft gluon divergence in the limit $z_2\to 0$, as the invariant mass, which we keep finite, gives a lower bound for the integral $z_2\gtrsim 1/M_X^2$. For a parametrically large M_X our result would give a large logarithm $\sim\ln M_X^2$ from the lower limit of the z_2 integration. While such contributions could be resummed [117], they are not easily accessible at EIC energies and we will not consider this resummation further here. The BK/JIMWLK evolution of the target, on the other hand, is associated with the $z_2\to 0$ limit of contributions where the gluon crosses the shockwave, but is reabsorbed and not measured in the final state, which we are leaving to future work.

The interpretation of the cumbersome Υ terms is actually straightforward. First, $\Upsilon_{\mathrm{reg}}^{(|(b)|^2)}$ describes the contribution where the gluon is emitted by the quark in the amplitude and absorbed by the same quark in the conjugate amplitude. Similarly, $\Upsilon_{\mathrm{reg}}^{(|(c)|^2)}$ corresponds to the case where the antiquark emits and absorbs the gluon. Furthermore, the instantaneous gluon emission and absorption by the quark (antiquark) is described by $\Upsilon_{\mathrm{inst}}^{(d)}$ ($\Upsilon_{\mathrm{inst}}^{(e)}$). Finally, gluon emission by the quark and absorption by the antiquark (or vice versa) contributes the term $\Upsilon_{\mathrm{interf}}^{(b)\times(c)}$. Note that the instantaneous contribution only appears as a part of the transverse cross section. The interference between the regular and instantaneous gluon emissions is included in the terms $\Upsilon_{\mathrm{inst}}^{(d)}$ and $\Upsilon_{\mathrm{inst}}^{(e)}$, with the former including the interference contributions containing the (d) diagram, and similarly the latter those of the (e) diagram.

We emphasize that prior to this work the $q\bar{q}g$ contribution to the diffractive cross section has only been known in approximate kinematics and for a transverse photon only (which dominates at high Q^2), see discussion in Sec. VIII. For the longitudinal polarization even approximative results have been missing from the literature. The cross sections (88) and (89), that are the main results of this work, are finite and can be straightforwardly implemented in phenomenological applications. In a future work we plan to apply these results to describe the HERA diffractive structure function data.

VIII. RECOVERING KNOWN LIMITS

A. The large-M_X limit

As a verification of our calculation, here we will extract the $q\bar{q}g$ contribution for F_T^{D} in the limit of large M_X, i.e., in the limit when the emitted gluon is soft. This limit has been considered by several authors, e.g., in Refs. [66,117–121].

To be specific, we will use the result as it is written in Ref. [66] and also used in the phenomenological studies [61,62]. The derivation simplifies considerably in the $\beta \to 0$ limit, i.e., for final states with arbitrarily large invariant masses M_X in comparison to the virtuality of the photon: $M_X^2 \gg Q^2$. The result of Ref. [66] for the $q\bar{q}g$ contribution—including gluon emissions from the quark-antiquark dipole both *before* and *after* the scattering off the shockwave[6]—is, converted to our notations,

$$x_{\mathbb{P}} F_{T,q\bar{q}g}^{\mathrm{D(MS)}}(x_{\mathbb{P}}, \beta = 0, Q^2) = \frac{\alpha_{\mathrm{s}} N_{\mathrm{c}} C_{\mathrm{F}} Q^2}{16\pi^5 \alpha_{\mathrm{em}}} \int \mathrm{d}^2\mathbf{x}_0 \int \mathrm{d}^2\mathbf{x}_1 \int \mathrm{d}^2\mathbf{x}_2 \int_0^1 \frac{\mathrm{d}z}{z(1-z)} |\tilde{\psi}^{\mathrm{LO}}_{\gamma_\lambda^* \to q_0 \bar{q}_{\bar{1}}}|^2 \times \frac{\mathbf{x}_{01}^2}{\mathbf{x}_{02}^2 \mathbf{x}_{12}^2} [N_{02} + N_{12} - N_{01} - N_{02} N_{12}]^2. \tag{97}$$

A key feature of this result is that the $q\bar{q}g$ LFWF has been factorized to the leading order $q\bar{q}$ LFWF, the BK kernel describing the gluon emission, and the scattering of the tripole state in terms of scatterings of daughter dipoles.

Our starting point to rederive the large-M_X result is Eq. (48), with the final state momentum integrals still undone. First we note that the leading process to create high invariant mass final states is caused by the emissions of soft gluons, with $z_2 \ll 1$, and at the limit $z_2 \to 0$ we have $M_X^2 \approx \mathbf{P}_2^2/z_2$. We consider a t integrated cross section, whose momentum dependence in this limit reads

$$\begin{aligned} \frac{\mathrm{d}\sigma^{\mathrm{D}}_{\gamma_\lambda^* \to q\bar{q}g}}{\mathrm{d}M_X^2} &\sim \int \frac{\mathrm{d}z_2}{z_2} \int \frac{\mathrm{d}^2\mathbf{p}_0}{(2\pi)^2} \int \frac{\mathrm{d}^2\mathbf{p}_1}{(2\pi)^2} \\ &\times \int \frac{\mathrm{d}^2\mathbf{p}_2}{(2\pi)^2} e^{i\mathbf{x}_{\bar{0}0}\mathbf{p}_0 + i\mathbf{x}_{\bar{1}1}\mathbf{p}_1 + i\mathbf{x}_{\bar{2}2}\mathbf{p}_2} \delta\left(\frac{\mathbf{p}_2^2}{z_2} - M_X^2\right) \\ &= \delta^{(2)}(\mathbf{x}_0 - \mathbf{x}_{\bar{0}}) \delta^{(2)}(\mathbf{x}_1 - \mathbf{x}_{\bar{1}}) \int \mathrm{d}z_2 \\ &\times \int \frac{\mathrm{d}^2\mathbf{p}_2}{(2\pi)^2} e^{i\mathbf{x}_{\bar{2}2}\mathbf{p}_2} \frac{z_2}{\mathbf{p}_2^2} \delta\left(z_2 - \frac{\mathbf{p}_2^2}{M_X^2}\right) \\ &= \frac{1}{M_X^2} \delta^{(2)}(\mathbf{x}_0 - \mathbf{x}_{\bar{0}}) \delta^{(2)}(\mathbf{x}_1 - \mathbf{x}_{\bar{1}}) \delta^{(2)}(\mathbf{x}_2 - \mathbf{x}_{\bar{2}}), \end{aligned} \tag{98}$$

where we assumed that M_X^2 is dominated by the gluon light cone energy $\mathbf{p}_2^2/z_2$. We have here integrated over z_2, assuming that M_X is large enough so that it is always possible to find a solution to the delta function constraint $z_2 = \mathbf{p}_2^2/M_X^2$ with $0 < z_2 < 1$. In reality this is not possible for arbitrarily large $\mathbf{p}_2$. Thus our approximation has rendered the $\mathbf{p}_2$ integral unbounded, resulting in a delta function setting $\mathbf{x}_2 - \mathbf{x}_{\bar{2}}$. This approximation, as discussed in Sec. III and in Sec. VII, would make the cross section UV divergent unless one also includes the diagrams with emission after the shockwave, using the procedure of subtracting from the "emission before" term the appropriate coordinate limit. Thus we will include these contributions here, unlike in the rest of this paper.

At this point we can integrate over the coordinates in the conjugate amplitude using the delta functions. Starting from the equation (48) we now get

$$\begin{aligned} \frac{\mathrm{d}\sigma^{\mathrm{D}}_{\gamma_\lambda^* \to q\bar{q}g}}{\mathrm{d}M_X^2 \mathrm{d}x_{\mathbb{P}}} &= \frac{N_{\mathrm{c}} C_{\mathrm{F}}}{(4\pi)^2} \int_0^1 \frac{\mathrm{d}z_0}{z_0} \int_0^1 \frac{\mathrm{d}z_1}{z_1} \delta(z_0 + z_1 - 1) \\ &\times \int \mathrm{d}^2\mathbf{x}_0 \mathrm{d}^2\mathbf{x}_1 \mathrm{d}^2\mathbf{x}_2 \frac{1}{M_X^2} \\ &\times \sum_{f, h_0, h_1, \lambda_2} |\tilde{\psi}_{\gamma_\lambda^* \to q_{\bar{0}} \bar{q}_{\bar{1}} g_{\bar{2}}} (S^\dagger_{012} - 1) \\ &- [\text{emission after}]|^2. \end{aligned} \tag{99}$$

For calculating the subtraction terms correctly, we need to decompose the $\gamma \to q\bar{q}g$ wave function into gluon emission and effective $\gamma \to q\bar{q}$ parts according to Eq. (9). Here the limit of $z_2 \to 0$ simplifies things dramatically. In the soft gluon limit the wave function is given by [note that in our normalization conventions for the reduced wave functions, see Eqs. (16) and (17), this relation is true without additional coefficients]:

$$\tilde{\psi}_{\gamma_\lambda^* \to q_0 \bar{q}_1 g_2} \underset{z_2 \to 0}{\approx} \tilde{\psi}^{\mathrm{LO}}_{\gamma_\lambda^* \to q_0 \bar{q}_1} [\tilde{\psi}_{q_0 \to q_0 g_2} + \tilde{\psi}_{\bar{q}_1 \to \bar{q}_1 g_2}], \tag{100}$$

where the effective $\gamma \to q\bar{q}$ wave functions $\tilde{\psi}_{\gamma_\lambda^* \to q_0 \bar{q}_1; q_0 \to q_0 g_2}$ and $\tilde{\psi}_{\gamma_\lambda^* \to q_0 \bar{q}_1; \bar{q}_1 \to \bar{q}_1 g_2}$ defined by Eq. (9) have become independent of the gluon emission in the limit $z_2 \to 0$. This simplification can be understood in several equivalent ways. In coordinate space, the argument is that for a soft emission $z_2 \to 0$ the coordinate of the emitting particle does not change, and therefore the original $\gamma \to q\bar{q}$ splitting is independent of the later gluon emission. For momentum space wave functions the reason is that in the limit $z_2 \to 0$ the gluon light cone energy blows up and thus energy denominators for gluon emission become independent of the state emitting the gluon,

$$k^-_{q\bar{q}g} - k^-_\gamma \approx k^-_g \approx k^-_{qg} - k^-_q \approx k^-_{\bar{q}g} - k^-_{\bar{q}}. \tag{101}$$

This leads to a factorization of the $\gamma \to q\bar{q}g$ wave function into leading order $\gamma \to q\bar{q}$ and $q \to qg$, $\bar{q} \to \bar{q}g$ wave functions in momentum space, which is carried over into coordinate space. We recall from Sec. III A that the

[6]The authors of Ref. [66] justify the contribution from the emission after the shockwave as resulting from the normalization of the photon state, which is possible in the soft gluon limit. However, this contribution is better understood as resulting from emissions after the shockwave; see discussion in Ref. [121]. The normalization condition for the photon state can only be used to obtain such contributions in the soft gluon limit where the $\gamma \to q\bar{q}g$ wave function factorizes into $\gamma \to q\bar{q}$ and gluon emission wave functions, but not in general kinematics [31,32].

subtraction procedure consists in pulling out the gluon emission wave function and then taking the limit of the gluon coordinate being equal to its emitter in the remaining factors. In this case we can also factor out the common $\gamma \to q\bar{q}$ part. Thus we get, restoring the coordinates for clarity,

$$\begin{aligned}&[\tilde{\psi}_{\gamma^*_\lambda \to q_0\bar{q}_1 g_2}(\mathbf{x}_0,\mathbf{x}_1,\mathbf{x}_2)(S_{012}-1) - [\text{emission after}]]\\ &= \tilde{\psi}^{\mathrm{LO}}_{\gamma^*_\lambda \to q_0\bar{q}_1}(\mathbf{x}_0,\mathbf{x}_1)\{\tilde{\psi}_{q_0\to q_0 g_2}(\mathbf{x}_0,\mathbf{x}_2)\\ &\quad\times[(S_{012}-1)-(S_{012}-1)|_{\mathbf{x}_2\to\mathbf{x}_0}]\\ &\quad+\tilde{\psi}_{\bar{q}_1\to\bar{q}_1 g_2}(\mathbf{x}_1,\mathbf{x}_2)[(S_{012}-1)-(S_{012}-1)|_{\mathbf{x}_2\to\mathbf{x}_1}]\}.\end{aligned} \tag{102}$$

We now use the coordinate limits of the "tripole" (see, e.g., [32,33]) operators

$$S_{012}|_{\mathbf{x}_2\to\mathbf{x}_0} = S_{012}|_{\mathbf{x}_2\to\mathbf{x}_1} = S_{01} \tag{103}$$

and the relation (see also, e.g., [32,33])

$$\begin{aligned}S_{012}-S_{01} &= \frac{{N_c}^2}{{N_c}^2-1}(S_{02}S_{21}-S_{01}) \overset{N_c\to\infty}{\approx} S_{02}S_{21}-S_{01}\\ &= N_{02}N_{21}-N_{02}-N_{21}+N_{01}\end{aligned} \tag{104}$$

with $N_{ij} = 1 - S_{ij}$ to get

$$\begin{aligned}&[\tilde{\psi}_{\gamma^*_\lambda \to q_0\bar{q}_1 g_2}(\mathbf{x}_0,\mathbf{x}_1,\mathbf{x}_2)(S_{012}-1) - [\text{emission after}]]\\ &= \tilde{\psi}^{\mathrm{LO}}_{\gamma^*_\lambda \to q_0\bar{q}_1}(\mathbf{x}_0,\mathbf{x}_1)\{\tilde{\psi}_{q_0\to q_0 g_2}(\mathbf{x}_0,\mathbf{x}_2)\\ &\quad+\tilde{\psi}_{\bar{q}_1\to\bar{q}_1 g_2}(\mathbf{x}_1,\mathbf{x}_2)\}[N_{02}N_{21}-N_{02}-N_{21}+N_{01}].\end{aligned} \tag{105}$$

We can now use the gluon emission wave function, which in our conventions [see Eq. (37) of Ref. [80] and recall the normalization of the reduced wavefunction from (16), (17)] reads

$$\begin{aligned}\tilde{\psi}_{q_0\to q_0 g_2}(\mathbf{x}_0,\mathbf{x}_2) &\underset{z_2\to 0}{\approx} 2g\int\frac{\mathrm{d}^2\mathbf{p}_2}{(2\pi)^2}\frac{\mathbf{p}_2\cdot\boldsymbol{\varepsilon}^*_{\lambda_2}}{\mathbf{p}_2^2}e^{i\mathbf{p}_2\cdot(\mathbf{x}_2-\mathbf{x}_0)}\\ &= \frac{2gi}{2\pi}\frac{(\mathbf{x}_2-\mathbf{x}_0)\cdot\boldsymbol{\varepsilon}^*_{\lambda_2}}{(\mathbf{x}_2-\mathbf{x}_0)^2},\end{aligned} \tag{106}$$

and similarly for the antiquark. Using this relation in Eq. (99) enables the familiar calculation of the BK kernel

$$\begin{aligned}&\sum_{\lambda_2}|\tilde{\psi}_{q_0\to q_0 g_2}(\mathbf{x}_0,\mathbf{x}_2)+\tilde{\psi}_{\bar{q}_1\to\bar{q}_1 g_2}(\mathbf{x}_1,\mathbf{x}_2)|^2\\ &= \frac{g^2}{\pi^2}\frac{\mathbf{x}_{01}^2}{\mathbf{x}_{02}^2\mathbf{x}_{21}^2} = \frac{4\alpha_s}{\pi}\frac{\mathbf{x}_{01}^2}{\mathbf{x}_{02}^2\mathbf{x}_{21}^2}.\end{aligned} \tag{107}$$

Inserting the squared wave functions from Eq. (107) and (105) into (99) one then directly recovers the result (97).

B. The large-Q^2 limit

In this section we will rederive the Wüsthoff result for the $q\bar{q}g$ contribution to F_T^D [58,60,108], which is an approximate result at the large-Q^2 limit for this next-to-leading order contribution. Specifically we will verify that the Wüsthoff result emerges from the NLO result calculated in exact kinematics when one takes the large-Q^2 limit. The Wüsthoff result for the $q\bar{q}g$ contribution has been extensively used in phenomenology [60–62,64], and it has some special features we seek to understand in depth. It can be written in coordinate space in the appealing short form[7]:

$$\begin{aligned}x_{\mathbb{P}}F^{\mathrm{D(GBW)}}_{T,q\bar{q}g}(x_{\mathbb{P}},\beta,Q^2) &= \frac{\alpha_s\beta}{8\pi^4}\sum_f e_f^2\int\mathrm{d}^2\mathbf{b}\int_0^{Q^2}\mathrm{d}k^2\int_\beta^1\mathrm{d}z\Big\{k^4\ln\frac{Q^2}{k^2}\left[\left(1-\frac{\beta}{z}\right)^2+\left(\frac{\beta}{z}\right)^2\right]\\ &\quad\times\left[\int_0^\infty\mathrm{d}r\,r\frac{\mathrm{d}\tilde{\sigma}_{\mathrm{dip}}}{\mathrm{d}^2\mathbf{b}}(\mathbf{b},\mathbf{r},x_{\mathbb{P}})\mathrm{K}_2(\sqrt{z}kr)\mathrm{J}_2(\sqrt{1-z}kr)\right]^2\Big\},\end{aligned} \tag{108}$$

while a momentum space version can be found, e.g., from Ref. [60].

An essential feature of the large-Q^2 structure function is the manifestation of the DGLAP splitting function for $g \to q\bar{q}$ splitting. This is associated with the DGLAP evolution of the parton distributions of the pomeron, and is written in terms of the target minus momentum fractions β and z (recall that we work in a frame where the photon has a large plus momentum). In the frame where the target has a large longitudinal momentum β can be interpreted as the fraction of the pomeron momentum carried by the stuck parton, and it is also related to the invariant mass of the final state as $\beta = Q^2/(Q^2+M_X^2)$. The fraction z is the fraction of the pomeron minus momentum transferred to the $q\bar{q}$ system. On the one hand, the presence of the splitting function in the large Q^2 limit is unavoidable in QCD. On the other hand, since we are using light cone perturbation theory to quantize

[7]We use the explicit form from Ref. [62], which is written in color glass condensate formalism. Its connection to the original two-gluon exchange formulation [58,60] is discussed in Refs. [61,65].

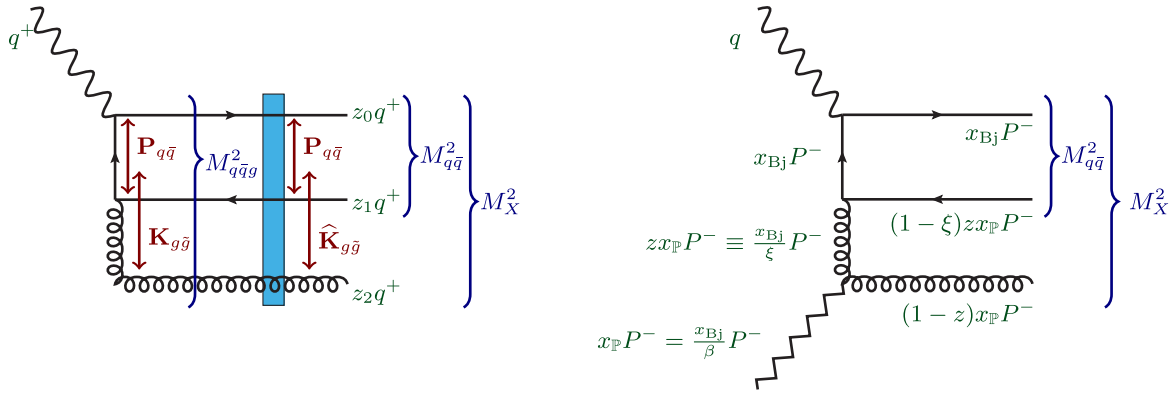

FIG. 12. Kinematics for the diffractive $q\bar{q}g$-production diagram. Left: dipole picture frame, where the probe has a large q^+ which is conserved in the interaction with the target. We show the plus and transverse momenta, with $\mathbf{P}_{q\bar{q}}$ the relative transverse momentum of the $q\bar{q}$ pair and $\mathbf{K}_{g\tilde{g}}$ of the gluon-effective gluon dipole. Right: infinite target momentum frame, where one tracks the minus component of the momentum, i.e., the target momentum fraction. The zigzag line refers to the pomeron emitted from the target, and is used to illustrate a generic diffractive interaction between the virtual photon and target. The two scattering pictures are connected by the invariants of the scattering. In the GBW limit the $q\bar{q}$ pair forms a hard system that is seen as pointlike by the target color field, thus $M^2_{q\bar{q}}$ and $\mathbf{P}_{q\bar{q}}$ are the same before and after the shockwave. In the aligned jet limit $z_2 \ll z_1 \ll z_0 \approx 1$ we have $M^2_{q\bar{q}} \approx \mathbf{P}^2_{q\bar{q}}/z_1$, and from the collinear factorization picture on the right: $M^2_{q\bar{q}} = 2(x_{\mathrm{Bj}}/\xi)q^+P^- + q^2 = (1/\xi - 1)Q^2$. The invariant mass of the $q\bar{q}g$ system and $\mathbf{K}_{g\tilde{g}}$, on the other hand, are affected by the interaction with the shockwave. In the dipole picture, before the shockwave, we have $M^2_{q\bar{q}g} \approx M^2_{q\bar{q}} + \mathbf{K}^2_{g\tilde{g}}/z_2$. The invariant mass after the shockwave is $M^2_X = (1/\beta - 1)Q^2 \approx M^2_{q\bar{q}} + \hat{\mathbf{K}}^2_{g\tilde{g}}/z_2$.

the projectile photon, the light cone momentum fraction $\xi = \beta/z$ does not appear spontaneously in the calculation. These kinematical variables are illustrated in Fig. 12.

The three-particle phase space has been treated only approximately in the Wüsthoff result. The result is written in terms of the size r of an "effective adjoint dipole" formed by the gluon and the quark-antiquark pair. Thus the interaction with the target color field appears in the form of an adjoint representation dipole cross section $\tilde{\sigma}_{\mathrm{dip}}$. There is also an explicit $\log Q^2$ resulting from the integral over some of the "internal" kinematics of the $q\bar{q}$ pair as we will demonstrate explicitly below.

The contribution corresponding to the DGLAP splitting function $[(1-\beta/z)^2 + (\beta/z)^2]$ originates from the part of phase space where emissions are strongly ordered in transverse momenta. Since we want to obtain the DGLAP splitting function with fixed momentum fractions in k^-, this implies that the k^+ momentum fractions must also be strongly ordered. This is the Bjorken aligned jet [122] configuration. Thus it is more convenient to work in momentum space to make the connection to the Wüsthoff result. To make this more specific, we choose the transverse momenta for the three-particle $q\bar{q}g$ final state (after the shockwave) as

$$\mathbf{P}_i \coloneqq \mathbf{p}_i - z_i\mathbf{q}, \tag{109}$$

$$\hat{\mathbf{P}}_{q\bar{q}} \coloneqq \frac{z_0\mathbf{P}_1 - z_1\mathbf{P}_0}{z_0 + z_1}, \tag{110}$$

$$\hat{\mathbf{K}}_{g\tilde{g}} \coloneqq (z_0 + z_1)\mathbf{P}_2 - z_2(\mathbf{P}_0 + \mathbf{P}_1), \tag{111}$$

$$\hat{\mathbf{\Delta}} \coloneqq -\mathbf{P}_0 - \mathbf{P}_1 - \mathbf{P}_2, \tag{112}$$

where $\mathbf{p}_i$ are the transverse momenta of the final state partons, and the variable $\hat{\mathbf{P}}_{q\bar{q}}$ can be interpreted as the relative momentum between the quark and antiquark, whereas $\hat{\mathbf{K}}_{g\tilde{g}}$ is the momentum of the gluon with respect to the quark-antiquark system, i.e., the "effective gluon." The total transverse momentum transfer in the scattering process is $\hat{\mathbf{\Delta}}$. The corresponding conjugate variables in transverse coordinate space can be determined by writing

$$\mathbf{P}_0 \cdot \mathbf{x}_0 + \mathbf{P}_1 \cdot \mathbf{x}_1 + \mathbf{P}_2 \cdot \mathbf{x}_2 \equiv \hat{\mathbf{P}}_{q\bar{q}} \cdot \mathbf{u} + \hat{\mathbf{K}}_{g\tilde{g}} \cdot \mathbf{r} + \hat{\mathbf{\Delta}} \cdot \mathbf{b}, \tag{113}$$

from which one obtains

$$\mathbf{u} = \mathbf{x}_1 - \mathbf{x}_0, \tag{114}$$

$$\mathbf{r} = \mathbf{x}_2 - \frac{z_0\mathbf{x}_0 + z_1\mathbf{x}_1}{z_0 + z_1}, \tag{115}$$

$$\mathbf{b} = z_0\mathbf{x}_0 + z_1\mathbf{x}_1 + z_2\mathbf{x}_2. \tag{116}$$

Here $\mathbf{u}$ is the size of the $q\bar{q}$ dipole, $\mathbf{r}$ is the distance from the gluon to the center of mass of the $q\bar{q}$ system, i.e., the size of the effective gluonic dipole, and $\mathbf{b}$ is the center of mass.

In the derivation of the large Q^2 limit of our result for F_T^D, Eq. (88), our starting point is the expression (82) for the diffractive cross section. The calculation then proceeds

roughly as follows. First we Fourier transform the LFWFs back into momentum space where we can apply the kinematic approximations related to the leading $\log Q^2$ limit. Then we write the virtual photon LFWF in momentum space using natural relative momenta from the perspective of the physical picture that emerges at the large-Q^2 limit discussed above. This enables us to apply the aligned jet limit (AJL) approximations at the LFWF level, after which we proceed to integrate over surplus degrees of freedom at the cross section level. In the aligned jet kinematics we have the strong transverse momentum ordering $Q^2 \gg \mathbf{P}_{q\bar{q}}^2 \gg \mathbf{K}_{g\tilde{g}}^2 \gg \boldsymbol{\Delta}^2$. The corresponding k^+ momentum fraction ordering is $z_2 \ll z_1 \ll z_0$ or, symmetrically, $z_2 \ll z_0 \ll z_1$. Since these two limits give the same contribution, we will just concentrate on the first one and in the end multiply the result by 2.

The transverse Fourier transform of the LFWF into momentum space is defined as [32]

$$\tilde{\psi}_{\mathrm{NLO}}^{\gamma_{\mathrm{T}}^*\to q\bar{q}g} = \int \frac{\mathrm{d}^2\mathbf{k}_0}{(2\pi)^2}\int \frac{\mathrm{d}^2\mathbf{k}_1}{(2\pi)^2}\int \frac{\mathrm{d}^2\mathbf{k}_2}{(2\pi)^2}(2\pi)^2\delta^{(2)}(\mathbf{k}_0+\mathbf{k}_1+\mathbf{k}_2-\mathbf{q})e^{i(\mathbf{k}_0\cdot\mathbf{x}_0+\mathbf{k}_1\cdot\mathbf{x}_1+\mathbf{k}_2\cdot\mathbf{x}_2)}\psi_{\mathrm{NLO}}^{\gamma_{\mathrm{T}}^*\to q\bar{q}g}. \tag{117}$$

Performing the Fourier transforms and rewriting the Fourier momenta in terms of the natural momenta using

$$\mathbf{k}_1 = -\frac{z_1}{z_0+z_1}\mathbf{K}_{g\tilde{g}} + \mathbf{P}_{q\bar{q}}, \tag{118}$$

$$\mathbf{k}_0 = -\frac{z_0}{z_0+z_1}\mathbf{K}_{g\tilde{g}} - \mathbf{P}_{q\bar{q}}, \tag{119}$$

the diffractive cross section from Eq. (82) becomes

$$\begin{aligned}\frac{\mathrm{d}\sigma_{\lambda,q\bar{q}g}^{\mathrm{D}}}{\mathrm{d}M_X^2} = {}& \frac{N_{\mathrm{c}}C_{\mathrm{F}}}{4(2\pi)^2}\int_{-\infty}^{0}\mathrm{d}t\int_0^1\frac{\mathrm{d}z_0}{z_0}\int_0^1\frac{\mathrm{d}z_1}{z_1}\int_0^1\frac{\mathrm{d}z_2}{z_2}\delta(z_0+z_1+z_2-1)\int\frac{\mathrm{d}^2\hat{\mathbf{P}}_{q\bar{q}}}{(2\pi)^2}\int\frac{\mathrm{d}^2\hat{\mathbf{K}}_{g\tilde{g}}}{(2\pi)^2}\int\frac{\mathrm{d}^2\hat{\boldsymbol{\Delta}}}{(2\pi)^2}\\ &\times\delta(\hat{\boldsymbol{\Delta}}^2-|t|)\delta\left(\frac{\hat{\mathbf{K}}_{g\tilde{g}}^2}{z_2(z_0+z_1)}+\frac{z_0+z_1}{z_0z_1}\hat{\mathbf{P}}_{q\bar{q}}^2-M_X^2\right)\\ &\times\int_{\mathbf{u}}\int_{\mathbf{r}}\int_{\mathbf{b}}\int_{\bar{\mathbf{u}}}\int_{\bar{\mathbf{r}}}\int_{\bar{\mathbf{b}}}(2\pi)^6 e^{i(\bar{\mathbf{u}}-\mathbf{u})\cdot\hat{\mathbf{P}}_{q\bar{q}}}e^{i(\bar{\mathbf{r}}-\mathbf{r})\cdot\hat{\mathbf{K}}_{g\tilde{g}}}e^{i(\bar{\mathbf{b}}-\mathbf{b})\cdot\hat{\boldsymbol{\Delta}}}\\ &\times\sum_{h_0,h_1,\lambda_2}(\tilde{\psi}_{\gamma_\lambda^*\to q_0\bar{q}_1g_2})^\dagger(\tilde{\psi}_{\gamma_\lambda^*\to q_0\bar{q}_1g_2})[S_{\bar{\mathbf{u}}\bar{\mathbf{r}}\bar{\mathbf{b}}}^{(3)\dagger}-1][S_{\mathbf{urb}}^{(3)}-1],\end{aligned} \tag{120}$$

where the hatted quantities are the final state momenta. The squared virtual photon amplitude Fourier transformed to coordinate space reads

$$\begin{aligned}(\tilde{\psi}_{\gamma_\lambda^*\to q_0\bar{q}_1g_2})^\dagger(\tilde{\psi}_{\gamma_\lambda^*\to q_0\bar{q}_1g_2}) = {}& \int\frac{\mathrm{d}^2\bar{\mathbf{P}}_{q\bar{q}}}{(2\pi)^2}\int\frac{\mathrm{d}^2\bar{\mathbf{K}}_{g\tilde{g}}}{(2\pi)^2}\int\frac{\mathrm{d}^2\bar{\boldsymbol{\Delta}}}{(2\pi)^2}(\psi_{\gamma_\lambda^*\to q_0\bar{q}_1g_2})^\dagger e^{-i\bar{\mathbf{u}}\cdot\bar{\mathbf{P}}_{q\bar{q}}-i\bar{\mathbf{r}}\cdot\bar{\mathbf{K}}_{g\tilde{g}}-i\bar{\mathbf{b}}\cdot\bar{\boldsymbol{\Delta}}}(2\pi)^2\delta^{(2)}(\bar{\boldsymbol{\Delta}})\\ &\times\int\frac{\mathrm{d}^2\mathbf{P}_{q\bar{q}}}{(2\pi)^2}\int\frac{\mathrm{d}^2\mathbf{K}_{g\tilde{g}}}{(2\pi)^2}\int\frac{\mathrm{d}^2\boldsymbol{\Delta}}{(2\pi)^2}(\psi_{\gamma_\lambda^*\to q_0\bar{q}_1g_2})e^{i\mathbf{u}\cdot\mathbf{P}_{q\bar{q}}+i\mathbf{r}\cdot\mathbf{K}_{g\tilde{g}}+i\mathbf{b}\cdot\boldsymbol{\Delta}}(2\pi)^2\delta^{(2)}(\boldsymbol{\Delta}).\end{aligned} \tag{121}$$

Now we begin the application of the aligned jet limit approximations. First, we assume that the relative transverse momentum of the parent dipole $\|\mathbf{P}_{q\bar{q}}\|$ is much larger than that of either of the daughter dipoles, i.e., $\|\mathbf{P}_{q\bar{q}}\| \gg \|\mathbf{K}_{g\tilde{g}}\|$. In position space this corresponds to a configuration where the gluon is emitted far away from the parent dipole, i.e., $\|\mathbf{u}\| \ll \|\mathbf{r}\|$. This makes the dipole amplitude independent of the parent dipole size, and enables us to separate and perform the $\mathbf{u}$-integrations in Eq. (120), which yields two delta functions

$$\int\frac{\mathrm{d}^2\mathbf{u}}{(2\pi)^2}\int\frac{\mathrm{d}^2\bar{\mathbf{u}}}{(2\pi)^2}e^{i\bar{\mathbf{u}}\cdot(\hat{\mathbf{P}}_{q\bar{q}}-\bar{\mathbf{P}}_{q\bar{q}})}e^{-i\mathbf{u}\cdot(\hat{\mathbf{P}}_{q\bar{q}}-\mathbf{P}_{q\bar{q}})} = \delta^{(2)}(\hat{\mathbf{P}}_{q\bar{q}}-\bar{\mathbf{P}}_{q\bar{q}})\delta^{(2)}(\hat{\mathbf{P}}_{q\bar{q}}-\mathbf{P}_{q\bar{q}}). \tag{122}$$

This enables us to perform the $\mathbf{P}_{q\bar{q}}$ and $\bar{\mathbf{P}}_{q\bar{q}}$ integrations in Eq. (120).

To perform the t integration of (120), we separate the closely related transverse momentum transfer integrals

$$\int_{-\infty}^{0} \mathrm{d}t \int \frac{\mathrm{d}^2\bar{\mathbf{b}}}{2\pi} \int \frac{\mathrm{d}^2\mathbf{b}}{2\pi} \int \frac{\mathrm{d}^2\hat{\mathbf{\Delta}}}{(2\pi)^2} \int \frac{\mathrm{d}^2\bar{\mathbf{\Delta}}}{(2\pi)^2} \int \frac{\mathrm{d}^2\mathbf{\Delta}}{(2\pi)^2} \delta(\hat{\Delta}^2 - |t|)(2\pi)^4 \delta^{(2)}(\bar{\mathbf{\Delta}})\delta^{(2)}(\mathbf{\Delta}) e^{i\bar{\mathbf{b}}\cdot(\hat{\mathbf{\Delta}}-\bar{\mathbf{\Delta}})} e^{-i\mathbf{b}\cdot(\hat{\mathbf{\Delta}}-\mathbf{\Delta})}$$
$$= \int_{-\infty}^{0} \mathrm{d}t \int \frac{\mathrm{d}^2\bar{\mathbf{b}}}{2\pi} \int \frac{\mathrm{d}^2\mathbf{b}}{2\pi} \int \frac{\mathrm{d}^2\hat{\mathbf{\Delta}}}{(2\pi)^2} e^{i\hat{\mathbf{\Delta}}\cdot(\bar{\mathbf{b}}-\mathbf{b})} \delta(\hat{\Delta}^2 - |t|) = \int \frac{\mathrm{d}^2\bar{\mathbf{b}}}{2\pi} \int \frac{\mathrm{d}^2\mathbf{b}}{2\pi} \delta^{(2)}(\bar{\mathbf{b}} - \mathbf{b}). \tag{123}$$

An essential step that we need to take before we proceed is to consider how to connect our results written in terms of the plus-momentum fractions to the minus momentum fractions in the Wüsthoff result (108). The essential idea here is to think in terms of invariant masses of different multiparton states. To connect the frames of reference, we need to approximate the invariant masses in the aligned jet limit as

$$M^2_{q\bar{q}} = \frac{1-\xi}{\xi} Q^2 = (1-z_2)^2 \frac{\mathbf{P}^2_{q\bar{q}}}{z_0 z_1} \approx \frac{\mathbf{P}^2_{q\bar{q}}}{z_1}, \tag{124}$$

$$M^2_{q\bar{q}g} + Q^2 = Q^2 + \frac{z_0+z_1}{z_0 z_1}\mathbf{P}^2_{q\bar{q}} + \frac{\mathbf{K}^2_{g\tilde{g}}}{z_2(z_0+z_1)}$$
$$\approx Q^2 + \frac{\mathbf{P}^2_{q\bar{q}}}{z_1} + \frac{\mathbf{K}^2_{g\tilde{g}}}{z_2} = \frac{Q^2}{\xi} + \frac{\mathbf{K}^2_{g\tilde{g}}}{z_2}, \tag{125}$$

where $zx_{\mathbb{P}} = x_{\mathrm{Bj}}/\xi$ is the fraction of the target momentum carried by the $q\bar{q}$ system. Given the above relations, we can reparametrize the often occurring combination

$$\mathbf{P}^2_{q\bar{q}} + z_1 Q^2 = z_1(M^2_{q\bar{q}} + Q^2) = z_1 Q^2/\xi. \tag{126}$$

It is useful to write the $q\bar{q}g$ energy denominator in terms of the diffractive state mass

$$M^2_X = \frac{1-\beta}{\beta} Q^2 \approx \frac{\hat{\mathbf{P}}^2_{q\bar{q}}}{z_1} + \frac{\hat{\mathbf{K}}^2_{g\tilde{g}}}{z_2}, \tag{127}$$

where $\hat{\mathbf{K}}_{g\tilde{g}}$ is the gluon relative momentum after the shockwave. Recall that $\mathbf{P}_{q\bar{q}} \equiv \hat{\mathbf{P}}_{q\bar{q}}$ is conserved in the shockwave which enables us to leverage final state information about M^2_X to simplify the γ^* LFWF before the shockwave. We now want to eliminate z_2 using these variables, so that

$$z_2 = \frac{\beta \hat{\mathbf{K}}^2_{g\tilde{g}}}{Q^2(1-z)}. \tag{128}$$

With this relation the three particle state invariant mass before the shockwave becomes

$$M^2_{q\bar{q}g} = \frac{Q^2}{\beta \hat{\mathbf{K}}^2_{g\tilde{g}}}[(z-\beta)\hat{\mathbf{K}}^2_{g\tilde{g}} + (1-z)\mathbf{K}^2_{g\tilde{g}}], \tag{129}$$

and the "outer" (i.e., $q\bar{q}g$ state) LC energy denominator in the NLO virtual photon LFWF will be

$$M^2_{q\bar{q}g} + Q^2 = \frac{Q^2}{\beta \hat{\mathbf{K}}^2_{g\tilde{g}}}[z\hat{\mathbf{K}}^2_{g\tilde{g}} + (1-z)\mathbf{K}^2_{g\tilde{g}}]. \tag{130}$$

Note the distinction that this is for the $q\bar{q}g$ state before the shockwave. The momentum $\hat{\mathbf{K}}^2_{g\tilde{g}}$ after the shockwave is fixed by the final state kinematics. It, or the $q\bar{q}g$ invariant mass, is not the same before ($M^2_{q\bar{q}g}$) and after (M^2_X) the shockwave. The momentum argument of the wave function before the shockwave $\mathbf{K}_{g\tilde{g}}$ will need to be Fourier transformed into coordinate space, see Eq. (121), in order to include the interaction with the target shockwave. The momenta $\mathbf{K}_{g\tilde{g}}$ and $\bar{\mathbf{K}}_{g\tilde{g}}$ before the shockwave are separate in the DA and CCA, which have to be Fourier transformed separately.

Next we move on to manipulate the virtual photon splitting light-front wave function in the aligned jet limit. The $\gamma^*_\lambda \to q\bar{q}g$ LFWF in momentum space in the convention of Refs. [31,32] is

$$\psi^{\gamma^*_{\mathrm{T}}\to q\bar{q}g}_{\mathrm{NLO}} = 4ee_f(gt^a_{\alpha\beta})\sqrt{z_0 z_1}\Bigg\{-\bar{\Sigma}^{ijkl}_{(b)} \frac{(-\mathbf{k}_1)^i \mathbf{m}^k \boldsymbol{\varepsilon}^j_\lambda \boldsymbol{\varepsilon}^{*l}_\sigma}{[\mathbf{k}_1^2 + \bar{Q}^2_{(b)}][\mathbf{m}^2 + \omega_{(b)}(\mathbf{k}_1^2 + \bar{Q}^2_{(b)})]}$$
$$- \bar{\Sigma}^{ijkl}_{(c)} \frac{\mathbf{k}_0^i \mathbf{l}^k \boldsymbol{\varepsilon}^j_\lambda \boldsymbol{\varepsilon}^{*l}_\sigma}{[\mathbf{k}_0^2 + \bar{Q}^2_{(c)}][\mathbf{l}^2 + \omega_{(c)}(\mathbf{k}_0^2 + \bar{Q}^2_{(c)})]}$$
$$- \bar{\Sigma}^{ij}_{(d)} \frac{\boldsymbol{\varepsilon}^{*i}_\sigma \boldsymbol{\varepsilon}^j_\lambda}{[\mathbf{m}^2 + \omega_{(d)}(\mathbf{k}_1^2 + \bar{Q}^2_{(d)})]} + \bar{\Sigma}^{ij}_{(e)} \frac{\boldsymbol{\varepsilon}^{*i}_\sigma \boldsymbol{\varepsilon}^j_\lambda}{[\mathbf{l}^2 + \omega_{(e)}(\mathbf{k}_0^2 + \bar{Q}^2_{(e)})]}\Bigg\}, \tag{131}$$

where the momenta are defined as

$$\mathbf{l} \coloneqq \frac{z_1}{(z_0+z_1)(z_1+z_2)}\mathbf{K}_{g\tilde{g}} - \frac{z_2}{z_1+z_2}\mathbf{P}_{q\bar{q}}, \tag{132}$$

$$\mathbf{m} \coloneqq \frac{z_0}{(z_0+z_1)(z_0+z_2)}\mathbf{K}_{g\tilde{g}} + \frac{z_2}{z_0+z_2}\mathbf{P}_{q\bar{q}}, \tag{133}$$

and the vertex factors are

$$\bar{\Sigma}^{ijkl}_{(b)} = \frac{1}{4}\frac{1}{z_0+z_2}[(2z_0+z_2)\delta^{kl} - i(2h_0)z_2\epsilon^{kl}][(2z_1-1)\delta^{ij} - i(2h_0)\epsilon^{ij}], \tag{134}$$

$$\bar{\Sigma}^{ijkl}_{(c)} = \frac{1}{4}\frac{1}{z_1+z_2}[(2z_1+z_2)\delta^{kl} + i(2h_0)z_2\epsilon^{kl}][(2z_0-1)\delta^{ij} + i(2h_0)\epsilon^{ij}], \tag{135}$$

$$\bar{\Sigma}^{ij}_{(d)} = \frac{1}{4}\frac{z_0 z_2}{(z_0+z_2)^2}[\delta^{ij} - i(2h_0)\epsilon^{ij}], \tag{136}$$

$$\bar{\Sigma}^{ij}_{(e)} = \frac{1}{4}\frac{z_1 z_2}{(z_1+z_2)^2}[\delta^{ij} + i(2h_0)\epsilon^{ij}], \tag{137}$$

with

$$\bar{Q}^2_{(b)} = z_1(z_0+z_2)Q^2, \qquad \omega_{(b)} = \frac{z_0 z_2}{z_1(z_0+z_2)^2}, \tag{138}$$

$$\bar{Q}^2_{(c)} = z_0(z_2+z_1)Q^2, \qquad \omega_{(c)} = \frac{z_2 z_1}{z_0(z_2+z_1)^2}, \tag{139}$$

$$\bar{Q}^2_{(d)} = z_1(z_0+z_2)Q^2, \qquad \omega_{(d)} = \frac{z_0 z_2}{z_1(z_0+z_2)^2}, \tag{140}$$

$$\bar{Q}^2_{(e)} = z_0(z_2+z_1)Q^2, \qquad \omega_{(e)} = \frac{z_2 z_1}{z_0(z_2+z_1)^2}. \tag{141}$$

In the aligned jet limit, these LC structures simplify significantly. Let us now take the strongly ordered limit $z_2 \ll z_1 \ll z_0$. Note that the leading term at $z_2 = 0$, $\mathbf{K}_{g\tilde{g}} \to 0$ cancels out, and thus one has to include subleading terms in the small-$\mathbf{K}_{g\tilde{g}}$ expansion to get the development right. We are left with

$$\mathbf{k}_1 \to \mathbf{P}_{q\bar{q}}, \tag{142}$$

$$\mathbf{k}_0 \to -\mathbf{K}_{g\tilde{g}} - \mathbf{P}_{q\bar{q}}, \tag{143}$$

$$\mathbf{l} \to \mathbf{K}_{g\tilde{g}} - \frac{z_2}{z_1}\mathbf{P}_{q\bar{q}}, \tag{144}$$

$$\mathbf{m} \to \mathbf{K}_{g\tilde{g}}, \tag{145}$$

and

$$\bar{\Sigma}^{ijkl}_{(b)} = -\frac{1}{2}[\delta^{ij} + i(2h_0)\epsilon^{ij}]\delta^{kl}, \tag{146}$$

$$\bar{\Sigma}^{ijkl}_{(c)} = \frac{1}{2}[\delta^{ij} + i(2h_0)\epsilon^{ij}]\delta^{kl}, \tag{147}$$

$$\bar{\Sigma}^{ij}_{(d)} = \frac{1}{4}z_2[\delta^{ij} - i(2h_0)\epsilon^{ij}], \tag{148}$$

$$\bar{\Sigma}^{ij}_{(e)} = \frac{1}{4}\frac{z_2}{z_1}[\delta^{ij} + i(2h_0)\epsilon^{ij}]. \tag{149}$$

Additionally it will be useful to simplify the momentum dependence of the $q\bar{q}g$ LFWF by using the three particle state invariant mass:

$$\mathbf{m}^2 + \omega_{(b)}(\mathbf{k}_1^2 + \bar{Q}^2_{(b)}) = \frac{z_0 z_2}{z_0+z_2}(M^2_{q\bar{q}g} + Q^2) \approx z_2(M^2_{q\bar{q}g} + Q^2), \tag{150}$$

$$\mathbf{l}^2 + \omega_{(c)}(\mathbf{k}_0^2 + \bar{Q}^2_{(c)}) = \frac{z_1 z_2}{z_1+z_2}(M^2_{q\bar{q}g} + Q^2) \approx z_2(M^2_{q\bar{q}g} + Q^2). \tag{151}$$

With these simplifications the normal emission part of the wave function (131) becomes

$$\psi^{\gamma^*_T \to q\bar{q}g}_{\mathrm{NLO}}\Big|_{\mathrm{AJL}} = 4ee_f(gt^a_{\alpha\beta})\sqrt{z_0 z_1}\frac{\boldsymbol{\epsilon}^j_\lambda \boldsymbol{\epsilon}^{*l}_\sigma \frac{1}{2}[\delta^{ij} + i(2h_0)\epsilon^{ij}]\delta^{kl}}{z_2(M^2_{q\bar{q}g} + Q^2)} \times \left[-\frac{\mathbf{P}^i_{q\bar{q}}}{\mathbf{P}^2_{q\bar{q}} + z_1 Q^2}\mathbf{K}^k_{g\tilde{g}} + \frac{(\mathbf{P}_{q\bar{q}} + \mathbf{K}_{g\tilde{g}})^i}{(\mathbf{P}_{q\bar{q}} + \mathbf{K}_{g\tilde{g}})^2 + z_1 Q^2}\left(\mathbf{K}^k_{g\tilde{g}} - \frac{z_2}{z_1}\mathbf{P}^k_{q\bar{q}}\right)\right]. \tag{152}$$

In the kinematic regime $\mathbf{P}_{q\bar{q}}^2 \gg \mathbf{K}_{g\tilde{g}}^2$ and we can first simplify the term in the square brackets as

$$[\ldots] = \frac{-\mathbf{P}_{q\bar{q}}^i((\mathbf{P}_{q\bar{q}}+\mathbf{K}_{g\tilde{g}})^2+z_1Q^2)\mathbf{K}_{g\tilde{g}}^k + (\mathbf{P}_{q\bar{q}}+\mathbf{K}_{g\tilde{g}})^i(\mathbf{P}_{q\bar{q}}^2+z_1Q^2)(\mathbf{K}_{g\tilde{g}}^k - \frac{z_2}{z_1}\mathbf{P}_{q\bar{q}}^k)}{((\mathbf{P}_{q\bar{q}}+\mathbf{K}_{g\tilde{g}})^2+z_1Q^2)(\mathbf{P}_{q\bar{q}}^2+z_1Q^2)}$$
$$\approx \frac{-2\mathbf{P}_{q\bar{q}}\cdot\mathbf{K}_{g\tilde{g}}\mathbf{P}_{q\bar{q}}^i + (\mathbf{P}_{q\bar{q}}^2+z_1Q^2)\mathbf{K}_{g\tilde{g}}^i}{(\mathbf{P}_{q\bar{q}}^2+z_1Q^2)^2}\mathbf{K}_{g\tilde{g}}^k - \frac{z_2}{z_1}\frac{\mathbf{P}_{q\bar{q}}^i\mathbf{P}_{q\bar{q}}^k}{\mathbf{P}_{q\bar{q}}^2+z_1Q^2} + \mathcal{O}(\mathbf{K}_{g\tilde{g}}^3), \tag{153}$$

where the last term is kept since in this strong ordering limit we have $z_2/z_1 \sim \mathbf{K}_{g\tilde{g}}^2/\mathbf{P}_{q\bar{q}}^2$, whereas the associated term proportional to $(z_2/z_1)\mathbf{K}_{g\tilde{g}}^i\mathbf{P}_{q\bar{q}}^k$ is discarded as a higher order term. Out of the instantaneous emission terms, the contribution from diagram in Fig. 4(e) is enhanced by a factor of $1/z_1$ with respect to diagram in Fig. 4(d), which means that in the aligned jet limit, we can neglect the contribution from the latter, and only the former contributes. Using the identity

$$\frac{1}{2}[\delta^{ij}+i(2h_0)\epsilon^{ij}]\boldsymbol{\varepsilon}_\lambda^j = \delta_{h_0,\lambda}\boldsymbol{\varepsilon}_\lambda^i \tag{154}$$

we see that the quark (and, consequently, antiquark) helicity is completely fixed by the photon polarization. This is a common feature of LCPT vertices: the particle carrying all the longitudinal momentum in a splitting inherits the light front helicity of the parent [80]. We can now combine the leading instantaneous contribution with the regular emissions

$$\psi_{\mathrm{NLO}}^{\gamma_{\mathrm{T}}^*\to q\bar{q}g}\Big|_{\mathrm{AJL}} = 4ee_f(gt_{\alpha\beta}^a)\sqrt{z_0z_1}\frac{\boldsymbol{\varepsilon}_\lambda^i\boldsymbol{\varepsilon}_\sigma^{*j}\delta_{h_0,\lambda}}{z_2(M_{q\bar{q}g}^2+Q^2)}$$
$$\times\left[\frac{-2\mathbf{P}_{q\bar{q}}\cdot\mathbf{K}_{g\tilde{g}}\mathbf{P}_{q\bar{q}}^i + (\mathbf{P}_{q\bar{q}}^2+z_1Q^2)\mathbf{K}_{g\tilde{g}}^i}{(\mathbf{P}_{q\bar{q}}^2+z_1Q^2)^2}\mathbf{K}_{g\tilde{g}}^j - \frac{z_2}{z_1}\frac{\mathbf{P}_{q\bar{q}}^i\mathbf{P}_{q\bar{q}}^j}{\mathbf{P}_{q\bar{q}}^2+z_1Q^2} + \frac{1}{2}\frac{z_2}{z_1}\delta^{ij}\right]$$
$$= 4ee_f(gt_{\alpha\beta}^a)\sqrt{z_0z_1}\frac{\boldsymbol{\varepsilon}_\lambda^i\boldsymbol{\varepsilon}_\sigma^{*j}\delta_{h_0,\lambda}}{z_2(M_{q\bar{q}g}^2+Q^2)}\frac{\xi}{z_1Q^2}\left[-2\frac{\xi}{z_1Q^2}(\mathbf{P}_{q\bar{q}}\cdot\mathbf{K}_{g\tilde{g}})\mathbf{P}_{q\bar{q}}^i\mathbf{K}_{g\tilde{g}}^j + \mathbf{K}_{g\tilde{g}}^i\mathbf{K}_{g\tilde{g}}^j - \xi\frac{z}{1-z}\frac{\hat{\mathbf{K}}_{g\tilde{g}}^2}{z_1Q^2}\mathbf{P}_{q\bar{q}}^i\mathbf{P}_{q\bar{q}}^j\right.$$
$$\left.+\frac{1}{2}\frac{z\hat{\mathbf{K}}_{g\tilde{g}}^2}{1-z}\delta^{ij}\right]. \tag{155}$$

Next we move on to calculate the squared amplitude. Here it is important to recall that the momenta $\mathbf{K}_{g\tilde{g}}$ and $\bar{\mathbf{K}}_{g\tilde{g}}$ in the DA and CCA, respectively, are separate, whereas $\mathbf{P}_{q\bar{q}} \equiv \bar{\mathbf{P}}_{q\bar{q}} \equiv \hat{\mathbf{P}}_{q\bar{q}}$ due to Eq. (122). We need the following algebra:

$$\left[-\xi\frac{\hat{\mathbf{P}}_{q\bar{q}}^i\hat{\mathbf{P}}_{q\bar{q}}^j}{z_1Q^2}+\frac{1}{2}\delta^{ij}\right]\left[-\xi\frac{\hat{\mathbf{P}}_{q\bar{q}}^i\hat{\mathbf{P}}_{q\bar{q}}^j}{z_1Q^2}+\frac{1}{2}\delta^{ij}\right] = \xi^2\frac{\hat{\mathbf{P}}_{q\bar{q}}^4}{(z_1Q^2)^2}+\xi\frac{\hat{\mathbf{P}}_{q\bar{q}}^2}{z_1Q^22}+\frac{1}{2} = \frac{1}{2}(\xi^2+(1-\xi)^2) \tag{156}$$

and

$$\left[-2\frac{\xi}{z_1Q^2}(\hat{\mathbf{P}}_{q\bar{q}}\cdot\mathbf{K}_{g\tilde{g}})\hat{\mathbf{P}}_{q\bar{q}}^i\mathbf{K}_{g\tilde{g}}^j+\mathbf{K}_{g\tilde{g}}^i\mathbf{K}_{g\tilde{g}}^j\right]\left[-\xi\frac{\hat{\mathbf{P}}_{q\bar{q}}^i\hat{\mathbf{P}}_{q\bar{q}}^j}{z_1Q^2}+\frac{1}{2}\delta^{ij}\right] = 2\xi^2\frac{\hat{\mathbf{P}}_{q\bar{q}}^2(\hat{\mathbf{P}}_{q\bar{q}}\cdot\mathbf{K}_{g\tilde{g}})^2}{(z_1Q^2)^2} - 2\xi\frac{(\hat{\mathbf{P}}_{q\bar{q}}\cdot\mathbf{K}_{g\tilde{g}})^2}{z_1Q^2}+\frac{1}{2}\mathbf{K}_{g\tilde{g}}^2. \tag{157}$$

Now we remember that at the cross section level we are integrating over the phase space $\mathrm{d}^2\hat{\mathbf{P}}_{q\bar{q}}$ with $\hat{\mathbf{P}}_{q\bar{q}}^2 = z_1(1-\xi)Q^2/\xi$ fixed, i.e., over the angle of $\hat{\mathbf{P}}_{q\bar{q}}$. This means that we can replace

$$\hat{\mathbf{P}}_{q\bar{q}}^i\hat{\mathbf{P}}_{q\bar{q}}^j \to \frac{1}{2}\hat{\mathbf{P}}_{q\bar{q}}^2\delta_{ij} = \frac{1}{2}\frac{1-\xi}{\xi}z_1Q^2\delta_{ij}. \tag{158}$$

Thus we get

$$2\xi^2\frac{\hat{\mathbf{P}}_{q\bar{q}}^2(\hat{\mathbf{P}}_{q\bar{q}}\cdot\mathbf{K}_{g\tilde{g}})^2}{(z_1Q^2)^2}-2\xi\frac{(\hat{\mathbf{P}}_{q\bar{q}}\cdot\mathbf{K}_{g\tilde{g}})^2}{z_1Q^2}+\frac{1}{2}\mathbf{K}_{g\tilde{g}}^2\to\frac{1}{2}(\xi^2+(1-\xi)^2)\mathbf{K}_{g\tilde{g}}^2. \tag{159}$$

Analogously the other cross term yields $\frac{1}{2}(\xi^2+(1-\xi)^2)\bar{\mathbf{K}}_{g\tilde{g}}^2$. Finally, the last term gives

$$\left[-2\frac{\xi}{z_1Q^2}(\hat{\mathbf{P}}_{q\bar{q}}\cdot\mathbf{K}_{g\tilde{g}})\hat{\mathbf{P}}_{q\bar{q}}^i\mathbf{K}_{g\tilde{g}}^j+\mathbf{K}_{g\tilde{g}}^i\mathbf{K}_{g\tilde{g}}^j\right]\left[-2\frac{\xi}{z_1Q^2}(\hat{\mathbf{P}}_{q\bar{q}}\cdot\bar{\mathbf{K}}_{g\tilde{g}})\hat{\mathbf{P}}_{q\bar{q}}^i\bar{\mathbf{K}}_{g\tilde{g}}^j+\bar{\mathbf{K}}_{g\tilde{g}}^i\bar{\mathbf{K}}_{g\tilde{g}}^j\right]=(\xi^2+(1-\xi)^2)(\mathbf{K}_{g\tilde{g}}\cdot\bar{\mathbf{K}}_{g\tilde{g}})^2. \tag{160}$$

With these the full squared amplitude can be written as

$$\begin{aligned}\frac{1}{2}\sum_{h,\lambda,\sigma}(\psi_{\mathrm{NLO}}^{\gamma_{\mathrm{T}}^*\to q\bar{q}g}|_{\mathrm{AJL}})^\dagger\psi_{\mathrm{NLO}}^{\gamma_{\mathrm{T}}^*\to q\bar{q}g}|_{\mathrm{AJL}}&=8e^2e_f^2g^2\frac{z_0z_1}{z_2^2}\frac{(\frac{\xi}{z_1Q^2})^2(\xi^2+(1-\xi)^2)}{(M_{q\bar{q}g}^2(\mathbf{K}_{g\tilde{g}})+Q^2)(M_{q\bar{q}g}^2(\bar{\mathbf{K}}_{g\tilde{g}})+Q^2)}\\&\times\left[\mathbf{K}_{g\tilde{g}}^i\mathbf{K}_{g\tilde{g}}^j+\frac{1}{2}\frac{z\hat{\mathbf{K}}_{g\tilde{g}}^2}{1-z}\delta^{ij}\right]\left[\bar{\mathbf{K}}_{g\tilde{g}}^i\bar{\mathbf{K}}_{g\tilde{g}}^j+\frac{1}{2}\frac{z\hat{\mathbf{K}}_{g\tilde{g}}^2}{1-z}\delta^{ij}\right],\end{aligned} \tag{161}$$

where we note that after the aligned jet configuration approximations the squared amplitude took a form that factorizes into a DA-like and a CCA-like parts similar to what has been seen in literature [58,60,123,124], even though this was not possible at earlier stages. To elaborate, we were only able to find the factorized form of the cross section with the "effective $q\bar{q}g$ wave function" after squaring the amplitude and integrating over the angle of the relative momentum of the $q\bar{q}$ pair. The same traceless rank-2 tensor structure is found in Ref. [123], also through a procedure of first squaring the amplitude, summing over internal degrees of freedom and then refactorizing the result. Appendix C of Ref. [123] gives an interesting interpretation of the tensor structure of the $\gamma\to g\tilde{g}$ effective wave function in terms of polarization vectors in the projectile and target light cone gauges. Note, however, that the recent paper [125] that appeared shortly after the first version of this manuscript, succeeds in extracting the traceless structure for the wave function already at the amplitude level, without squaring it first and without averaging over the angle of the hard momentum $\mathbf{P}$. We also note that the DGLAP splitting function $(\xi+(1-\xi)^2)$ manifests at this stage of the calculation, i.e., it is an underlying feature of the $\gamma_{\mathrm{T}}^*\to q\bar{q}g$ splitting function.

After this refactorization, the Fourier transforms over $\mathbf{K}_{g\tilde{g}}$ and $\bar{\mathbf{K}}_{g\tilde{g}}$ now separate, and can be evaluated as

$$\int\frac{\mathrm{d}^2\mathbf{K}}{(2\pi)^2}e^{i\mathbf{K}\cdot\mathbf{r}}\frac{\mathbf{K}^i\mathbf{K}^j+\frac{1}{2}\frac{z\hat{\mathbf{K}}_{g\tilde{g}}^2}{1-z}\delta^{ij}}{M_{q\bar{q}g}^2(\mathbf{K})+Q^2}=\frac{\beta\hat{\mathbf{K}}_{g\tilde{g}}^2}{(1-z)Q^2}\int\frac{\mathrm{d}^2\mathbf{K}}{(2\pi)^2}e^{i\mathbf{K}\cdot\mathbf{r}}\frac{\mathbf{K}^i\mathbf{K}^j+\frac{1}{2}\frac{z\hat{\mathbf{K}}_{g\tilde{g}}^2}{1-z}\delta^{ij}}{\mathbf{K}^2+\frac{z\hat{\mathbf{K}}_{g\tilde{g}}^2}{1-z}} \tag{162}$$

$$\begin{aligned}&=\frac{\beta\hat{\mathbf{K}}_{g\tilde{g}}^2}{(1-z)Q^2}\left[\left(\frac{\mathbf{r}^i\mathbf{r}^j}{\mathbf{r}^2}-\frac{1}{2}\delta^{ij}\right)\left(\delta^{(2)}(\mathbf{r})-\frac{1}{2\pi}\frac{z\hat{\mathbf{K}}_{g\tilde{g}}^2}{1-z}\mathrm{K}_2\left(\sqrt{\frac{z\hat{\mathbf{K}}_{g\tilde{g}}^2}{1-z}\mathbf{r}^2}\right)\right)+\frac{1}{2}\delta^{ij}\delta^{(2)}(\mathbf{r})\right]\\&\cong\frac{1}{2\pi}\frac{\beta\hat{\mathbf{K}}_{g\tilde{g}}^2}{(1-z)Q^2}\frac{z\hat{\mathbf{K}}_{g\tilde{g}}^2}{1-z}\left(\frac{1}{2}\delta^{ij}-\frac{\mathbf{r}^i\mathbf{r}^j}{\mathbf{r}^2}\right)\mathrm{K}_2\left(\sqrt{\frac{z\hat{\mathbf{K}}_{g\tilde{g}}^2}{1-z}\mathbf{r}^2}\right),\end{aligned} \tag{163}$$

where $\cong$ is used to imply effective equivalence since the contributions with $\mathbf{r}\equiv 0$ vanish identically as the dipole amplitude vanishes at $\mathbf{r}\equiv 0$. Thus after the transverse Fourier transforms we find for the squared virtual photon amplitude

$$\begin{aligned}&\int\frac{\mathrm{d}^2\mathbf{K}_{g\tilde{g}}}{(2\pi)^2}e^{i\mathbf{K}_{g\tilde{g}}\cdot\mathbf{r}}\int\frac{\mathrm{d}^2\bar{\mathbf{K}}_{g\tilde{g}}}{(2\pi)^2}e^{-i\bar{\mathbf{K}}_{g\tilde{g}}\cdot\mathbf{r}}\frac{1}{2}\sum_{h,\lambda,\sigma}(\psi_{\mathrm{NLO}}^{\gamma_{\mathrm{T}}^*\to q\bar{q}g}|_{\mathrm{AJL}})^\dagger\psi_{\mathrm{NLO}}^{\gamma_{\mathrm{T}}^*\to q\bar{q}g}|_{\mathrm{AJL}}\\&=2\frac{(4\pi)^2}{(2\pi)^2}\alpha_{\mathrm{em}}\alpha_{\mathrm{s}}e_f^2\frac{z_0}{z_1}\frac{\beta^4\hat{\mathbf{K}}_{g\tilde{g}}^8}{(1-z)^4Q^8}[\xi^2+(1-\xi)^2]\left(\delta^{ij}-2\frac{\mathbf{r}^i\mathbf{r}^j}{\mathbf{r}^2}\right)\left(\delta^{ij}-2\frac{\bar{\mathbf{r}}^i\bar{\mathbf{r}}^j}{\bar{\mathbf{r}}^2}\right)\mathrm{K}_2\left(\sqrt{\frac{z\hat{\mathbf{K}}_{g\tilde{g}}^2}{1-z}\mathbf{r}^2}\right)\mathrm{K}_2\left(\sqrt{\frac{z\hat{\mathbf{K}}_{g\tilde{g}}^2}{1-z}\bar{\mathbf{r}}^2}\right).\end{aligned} \tag{164}$$

For the next stage we need the following identities which enforce the physical invariant masses at the cross section level:

$$\int \frac{\mathrm{d}^2\hat{\mathbf{P}}_{q\bar{q}}}{(2\pi)^2}\delta\left(\frac{\hat{\mathbf{P}}^2_{q\bar{q}}}{z_1}+\frac{\hat{\mathbf{K}}^2_{g\tilde{g}}}{z_2}-M_X^2\right)=\frac{z_1}{4\pi}, \tag{165}$$

and

$$\int_0^1 \frac{\mathrm{d}z_2}{z_2^3}\delta\left(z-\left(1-\frac{\beta\hat{\mathbf{K}}^2_{g\tilde{g}}}{Q^2 z_2}\right)\right)=\frac{(1-z)Q^4}{\beta^2\hat{\mathbf{K}}^4_{g\tilde{g}}}. \tag{166}$$

We need to perform the change of variables $(z_2,\hat{\mathbf{K}}^2_{g\tilde{g}})\mapsto(z,k^2)$, where $k^2\coloneqq\hat{\mathbf{K}}^2_{g\tilde{g}}/(1-z)$ is the mean virtuality of the exchanged t-channel gluon in the two-gluon exchange model [58,60]. This change is enacted by inserting the relations between these quantities as δ-function integrals:

$$\begin{aligned}\int \frac{\mathrm{d}^2\hat{\mathbf{K}}_{g\tilde{g}}}{(2\pi)^2}\int\frac{\mathrm{d}z_2}{z_2^3}&=\int\frac{\mathrm{d}^2\hat{\mathbf{K}}_{g\tilde{g}}}{(2\pi)^2}\int\frac{\mathrm{d}z_2}{z_2^3}\int\mathrm{d}k^2\int\mathrm{d}z\,\delta\left(z-\left(1-\frac{\beta\hat{\mathbf{K}}^2_{g\tilde{g}}}{Q^2z_2}\right)\right)\delta\left(k^2-\frac{\hat{\mathbf{K}}^2_{g\tilde{g}}}{1-z}\right)\\&=\frac{1}{4\pi}\int\mathrm{d}k^2\int\mathrm{d}z\frac{Q^4}{\beta^2k^4}.\end{aligned} \tag{167}$$

After this the remaining longitudinal momentum fraction integrals separate and may be performed. Keeping in mind the assumption that $1\gtrsim z_0\gg z_1\gg z_2$, we find,

$$\begin{aligned}\int_0^1\mathrm{d}z_0\int_{z_{1,\min}}^{z_{1,\max}}\frac{\mathrm{d}z_1}{z_1}\delta(z_0+z_1+z_2-1)&=\int_{1-z_2-z_{1,\max}}^{1-z_2-z_{1,\min}}\frac{\mathrm{d}z_0}{1-z_0-z_2}\\&=\log\left(\frac{z_{1,\max}}{z_{1,\min}}\right)\equiv\log\left(\frac{Q^2}{\beta k^2}\right)\approx\log\left(\frac{Q^2}{k^2}\right).\end{aligned} \tag{168}$$

Here we have taken $z_{1,\max}=1$ and $z_{1,\min}\equiv z_2$, reflecting the kinematics of the aligned jet limit. We are also taking β to be of order one, and calculating in the leading large logarithmic limit in Q^2, thus the constant under the log is not under control at this point of the calculation. We will additionally need the relation

$$\int\frac{\mathrm{d}^2\mathbf{r}}{2\pi}\int\frac{\mathrm{d}^2\bar{\mathbf{r}}}{2\pi}e^{i\mathbf{k}_t\cdot(\mathbf{r}-\bar{\mathbf{r}})}\left(\delta^{ij}-2\frac{\mathbf{r}^i\mathbf{r}^j}{\mathbf{r}^2}\right)\left(\delta^{ij}-2\frac{\bar{\mathbf{r}}^i\bar{\mathbf{r}}^j}{\bar{\mathbf{r}}^2}\right)=2\int\mathrm{d}r\,r\mathrm{J}_2(k_tr)\int\mathrm{d}\bar{r}\,\bar{r}\mathrm{J}_2(k_t\bar{r}) \tag{169}$$

to simplify the tensor structure by computing the angular integrals.

Recalling that $\xi\coloneqq\frac{\beta}{z}$, we can finally collect our results and write the final result for the $q\bar{q}g$ contribution to the transverse diffractive structure function using Eq. (2) at the large-Q^2 limit

$$\begin{aligned}x_{\mathbb{P}}F_{T,q\bar{q}g}^{\mathrm{D}(3),\mathrm{LL}(Q^2)}&=2N_cC_F\frac{\alpha_s\beta}{16\pi^4}\sum_f e_f^2\int\mathrm{d}^2\mathbf{b}\int_\beta^1\mathrm{d}z\left[\left(\frac{\beta}{z}\right)^2+\left(1-\frac{\beta}{z}\right)^2\right]\int_0^{Q^2}\mathrm{d}k^2\,k^4\log\left(\frac{z_{1,\max}}{\beta}\frac{Q^2}{k^2}\right)\\&\times\int\mathrm{d}r\int\mathrm{d}\bar{r}\,r\mathrm{J}_2(\sqrt{1-z}kr)\bar{r}\mathrm{J}_2(\sqrt{1-z}k\bar{r})\mathrm{K}_2(\sqrt{z}kr)\mathrm{K}_2(\sqrt{z}k\bar{r})[1-S_{\bar{\mathbf{r}}\mathbf{b}}]^\dagger[1-S_{\mathbf{r}\mathbf{b}}]\\&=\frac{\alpha_s\beta}{8\pi^4}\sum_f e_f^2\int\mathrm{d}^2\mathbf{b}\int_\beta^1\mathrm{d}z\left[\left(\frac{\beta}{z}\right)^2+\left(1-\frac{\beta}{z}\right)^2\right]\int_0^{Q^2}\mathrm{d}k^2\,k^4\log\left(\frac{Q^2}{k^2}\right)\\&\times\left\{\int\mathrm{d}r\,r\mathrm{J}_2(\sqrt{1-z}kr)\mathrm{K}_2(\sqrt{z}kr)(2[1-S_{\mathbf{r}\mathbf{b}}])\right\}^2,\end{aligned} \tag{170}$$

where the leading factor of 2 accounts for the other momentum ordering where one has $z_1\gg z_0\gg z_2$. Explicitly writing out the color factor and taking into account the normalization of the dipole amplitude, we find exact agreement with the result of Ref. [62] shown in Eq. (108). To reiterate, we were able to find exact agreement with the large-Q^2 limiting $q\bar{q}g$ contribution to F_T^D with a calculation beginning from the corresponding full $\mathcal{O}(\alpha_s)$-accuracy

LCPT result. This precise agreement is very reassuring, given that the Wüsthoff result was derived in a very different formalism more in the spirit of a perturbative two-gluon exchange [58,60]. Results similar to the Wüsthoff result can be found in Refs. [123,124], where a semiclassical color field picture of the diffractive scattering is utilized.

In Eq. (170) we can recognize the origins of some key features. The wave function of a gluon splitting to the effective adjoint dipole is a Bessel K_2 function and the wave function overlap of the effective dipole and a final state with invariant mass M_X is a J_2. This particular structure, in stead of the usual K_1, J_1 for a transverse photon, originates from the transverse tensor structure of the $\gamma \to g\tilde{g}$ wave function at the aligned jet limit. Furthermore, we see the DGLAP $g \to q\bar{q}$ splitting function which emerged from the squared virtual photon wave function in momentum space at the aligned jet limit, corresponding to the first step in the DGLAP evolution of a diffractive quark parton distribution function. It is interesting that similar Bessel K_2 functions appear in a forward (small transverse momentum) limit also in the calculation of Ref. [44], although it is not completely obvious whether they have the same origin as here.

The transverse structure function, Eq. (170), is proportional to $\log Q^2$ and as such dominates at large Q^2. This logarithm originates from the aligned jet configuration part of the phase space when integrating over the kinematics of the $q\bar{q}$ pair. This is explicitly visible in Eq. (168) where the $\log Q^2/k^2$ contribution is obtained when integrating over the quark momentum fraction in the aligned jet limit. As the longitudinal photon wave function suppresses the aligned jet configurations where $z_0 \ll 1$ or $z_1 \ll 1$, the same log would not be present in the high-Q^2 limit of the longitudinal structure function.

IX. DISCUSSION AND CONCLUSIONS

In conclusion, we have here taken a significant step towards calculating diffractive structure functions at NLO accuracy in the color dipole picture applicable to the saturation regime of QCD. Our calculation includes the "radiative" part of the NLO correction, i.e., the $q\bar{q}g$ component in the terminology used in earlier works. For the diffractive structure function one calculates the cross section for a fixed invariant mass, a much more inclusive final state than for jet production. Compared to dijet production, this observable thus does not require a jet definition. The subset of the NLO contribution calculated here should be finite by itself.

We have also checked that we can independently reproduce two earlier results for the diffractive structure functions appearing in the literature. In the limit of a soft gluon (i.e., large mass diffractive state), we reproduce an earlier result derived by many authors (including, e.g., the derivation by Munier and Shoshi in a framework very similar to ours). More nontrivially, in the limit of a fixed $\beta = Q^2/(M_X^2 + Q^2)$ and large Q^2, we recover the earlier result used by Golec-Biernat and Wüsthoff (GBW) and in many other phenomenological studies. The original derivation of this result is perhaps not that clearly documented in the available literature. Certainly it is performed in a collinear factorization-type framework very different from ours. We have thus provided a completely independent rederivation of the large-Q^2 result in the dipole picture, fully agreeing with it.[8] Specifically, our calculation shows how to obtain the ingredients of the large-Q^2 result: a DGLAP-type logarithm in Q^2, a splitting function $P_{g\to q\bar{q}}$ and also the somewhat curious traceless rank-2 tensor "photon to effective gluon dipole wave function" [see Eq. (161)], from a dipole picture calculation. Thus we believe that the method of this calculation can be helpful in more general for matching the dipole picture with the collinear factorization limit.

Experimentally, the diffractive structure function is a key part of the program in high energy DIS experiments, both at HERA and at the future EIC. Compared to, e.g., diffractive dijets, it is a clean and well-defined observable without requiring high transverse momentum or heavy quarks in the final state. This makes it possible to access smaller values of x_{Bj} at a finite collision energy than for dijet observables, and thus to achieve a better sensitivity to gluon saturation. The nuclear modification of diffractive structure functions has already been identified as a key observable for gluon saturation at the EIC [1].

Our results are presented in a form that can directly be applied to phenomenology. They generalize the large-Q^2 and large-M_X results used in earlier phenomenological studies to a more precise kinematics. Depending on assumptions on the impact parameter dependence of the dipole amplitude, various different simplifications are possible. However, our main result is completely general in this regard and can be applied to any impact parameter dependence. It will be interesting to evaluate the diffractive structure function numerically, both in order to compare to earlier limiting results, and to test dipole amplitude parametrizations against a new set of experimental data. On the theory side, we have in this paper also outlined the necessary steps to complete the NLO calculation of the diffractive structure function. Here there are many recent results that can be taken advantage of. The loop corrections to the $\gamma \to q\bar{q}g$ light cone wavefunction are known [31–33]. So is the procedure to factorize the large logarithms of $x_{\mathbb{P}}$ into the BK/JIMWLK evolution of the target [65,126], once the corresponding diagrams (with a gluon crossing the shockwave but not the cut) are calculated. The treatment of final state gluon exchanges poses interesting conceptual questions that are new in the context of LCPT.

[8] Apart from the treatment of the color factor of the adjoint dipole, which was noted already in Refs. [61,62].

While in many ways straightforward, these further calculations are sizable enough projects that they are best left for future publications.

ACKNOWLEDGMENTS

We thank K. Golec-Biernat, A. van Hameren, C. Marquet, R. Paatelainen, and J. Penttala for discussions. H. H., T. L., and H. M. are supported by the Academy of Finland, the Centre of Excellence in Quark Matter (project 346324) and Projects No. 338263, No. 346567, and No. 321840. G. B. is supported in part by the National Science Centre (Poland) under the research Grant No. 2020/38/E/ST2/00122 (SONATA BIS 10). Y. M. acknowledges financial support from Xunta de Galicia (Centro singular de investigación de Galicia accreditation 2019–2022); the "María de Maeztu" Units of Excellence program MDM2016-0692 and the Spanish Research State Agency under Project No. PID2020–119632 GB-I00; European Union ERDF. G. B. and Y. M. acknowledge financial support from MSCA RISE 823947 "Heavy ion collisions: collectivity and precision in saturation physics" (HIEIC). This work was also supported under the European Union's Horizon 2020 research and innovation programme by the European Research Council (ERC, Grant Agreement No. ERC-2018-ADG-835105 YoctoLHC) and by the STRONG-2020 project (Grant Agreement No. 824093). The content of this article does not reflect the official opinion of the European Union and responsibility for the information and views expressed therein lies entirely with the authors.

[1] A. Accardi *et al.*, Electron ion collider: The next QCD frontier: Understanding the glue that binds us all, Eur. Phys. J. A **52**, 268 (2016).
[2] E. C. Aschenauer, S. Fazio, J. H. Lee, H. Mäntysaari, B. S. Page, B. Schenke, T. Ullrich, R. Venugopalan, and P. Zurita, The electron–ion collider: Assessing the energy dependence of key measurements, Rep. Prog. Phys. **82**, 024301 (2019).
[3] R. Abdul Khalek *et al.*, Science requirements and detector concepts for the electron-ion collider: EIC yellow report, Nucl. Phys. **A1026**, 122447 (2022).
[4] J. L. Abelleira Fernandez *et al.* (LHeC Study Group), A large hadron electron collider at CERN: Report on the physics and design concepts for machine and detector, J. Phys. G **39**, 075001 (2012).
[5] P. Agostini *et al.* (LHeC and FCC-he Study Groups), The large hadron-electron collider at the HL-LHC, J. Phys. G **48**, 110501 (2021).
[6] E. Iancu and R. Venugopalan, The color glass condensate and high-energy scattering in QCD, Quark Gluon Plasma **3**, 249 (2003).
[7] H. Weigert, Evolution at small x_{Bj}: The color glass condensate, Prog. Part. Nucl. Phys. **55**, 461 (2005).
[8] F. Gelis, E. Iancu, J. Jalilian-Marian, and R. Venugopalan, The color glass condensate, Annu. Rev. Nucl. Part. Sci. **60**, 463 (2010).
[9] N. N. Nikolaev and B. G. Zakharov, Color transparency and scaling properties of nuclear shadowing in deep inelastic scattering, Z. Phys. C **49**, 607 (1991).
[10] N. Nikolaev and B. G. Zakharov, Pomeron structure function and diffraction dissociation of virtual photons in perturbative QCD, Z. Phys. C **53**, 331 (1992).
[11] A. H. Mueller, Soft gluons in the infinite momentum wave function and the BFKL pomeron, Nucl. Phys. **B415**, 373 (1994).
[12] A. H. Mueller and B. Patel, Single and double BFKL pomeron exchange and a dipole picture of high-energy hard processes, Nucl. Phys. **B425**, 471 (1994).
[13] A. H. Mueller, Unitarity and the BFKL pomeron, Nucl. Phys. **B437**, 107 (1995).
[14] I. Balitsky, Operator expansion for high-energy scattering, Nucl. Phys. **B463**, 99 (1996).
[15] Y. V. Kovchegov, Unitarization of the BFKL pomeron on a nucleus, Phys. Rev. D **61**, 074018 (2000).
[16] Y. V. Kovchegov, Small-x F_2 structure function of a nucleus including multiple pomeron exchanges, Phys. Rev. D **60**, 034008 (1999).
[17] I. Balitsky and G. A. Chirilli, Next-to-leading order evolution of color dipoles, Phys. Rev. D **77**, 014019 (2008).
[18] I. Balitsky and G. A. Chirilli, Rapidity evolution of Wilson lines at the next-to-leading order, Phys. Rev. D **88**, 111501 (2013).
[19] A. Kovner, M. Lublinsky, and Y. Mulian, Jalilian-Marian, Iancu, McLerran, Weigert, Leonidov, Kovner evolution at next to leading order, Phys. Rev. D **89**, 061704 (2014).
[20] I. Balitsky and A. V. Grabovsky, NLO evolution of 3-quark Wilson loop operator, J. High Energy Phys. 01 (2015) 009.
[21] G. Beuf, Improving the kinematics for low-x QCD evolution equations in coordinate space, Phys. Rev. D **89**, 074039 (2014).
[22] T. Lappi and H. Mäntysaari, Direct numerical solution of the coordinate space Balitsky-Kovchegov equation at next to leading order, Phys. Rev. D **91**, 074016 (2015).
[23] E. Iancu, J. D. Madrigal, A. H. Mueller, G. Soyez, and D. N. Triantafyllopoulos, Resumming double logarithms in the QCD evolution of color dipoles, Phys. Lett. B **744**, 293 (2015).
[24] E. Iancu, J. D. Madrigal, A. H. Mueller, G. Soyez, and D. N. Triantafyllopoulos, Collinearly-improved BK evolution meets the HERA data, Phys. Lett. B **750**, 643 (2015).
[25] J. L. Albacete, Resummation of double collinear logs in BK evolution versus HERA data, Nucl. Phys. **A957**, 71 (2017).

[26] T. Lappi and H. Mäntysaari, Next-to-leading order Balitsky-Kovchegov equation with resummation, Phys. Rev. D **93**, 094004 (2016).
[27] M. Lublinsky and Y. Mulian, High energy QCD at NLO: From light-cone wave function to JIMWLK evolution, J. High Energy Phys. 05 (2017) 097.
[28] I. Balitsky and G. A. Chirilli, Photon impact factor in the next-to-leading order, Phys. Rev. D **83**, 031502 (2011).
[29] I. Balitsky and G. A. Chirilli, Photon impact factor and k_T-factorization for DIS in the next-to-leading order, Phys. Rev. D **87**, 014013 (2013).
[30] G. Beuf, NLO corrections for the dipole factorization of DIS structure functions at low x, Phys. Rev. D **85**, 034039 (2012).
[31] G. Beuf, Dipole factorization for DIS at NLO: Loop correction to the $\gamma^*_{T,L} \to q\bar{q}$ light-front wave functions, Phys. Rev. D **94**, 054016 (2016).
[32] G. Beuf, Dipole factorization for DIS at NLO: Combining the $q\bar{q}$ and $q\bar{q}g$ contributions, Phys. Rev. D **96**, 074033 (2017).
[33] H. Hänninen, T. Lappi, and R. Paatelainen, One-loop corrections to light cone wave functions: The dipole picture DIS cross section, Ann. Phys. (Amsterdam) **393**, 358 (2018).
[34] G. Beuf, H. Hänninen, T. Lappi, and H. Mäntysaari, Color glass condensate at next-to-leading order meets HERA data, Phys. Rev. D **102**, 074028 (2020).
[35] T. Lappi, H. Mäntysaari, and A. Ramnath, Next-to-leading order Balitsky-Kovchegov equation beyond large N_c, Phys. Rev. D **102**, 074027 (2020).
[36] P. Caucal, F. Salazar, and R. Venugopalan, Dijet impact factor in DIS at next-to-leading order in the color glass condensate, J. High Energy Phys. 11 (2021) 222.
[37] P. Taels, T. Altinoluk, G. Beuf, and C. Marquet, Dijet photoproduction at low x at next-to-leading order and its back-to-back limit, J. High Energy Phys. 10 (2022) 184.
[38] G. Beuf, T. Lappi, and R. Paatelainen, Massive quarks in NLO dipole factorization for DIS: Longitudinal photon, Phys. Rev. D **104**, 056032 (2021).
[39] G. Beuf, T. Lappi, and R. Paatelainen, Massive Quarks at One Loop in the Dipole Picture of Deep Inelastic Scattering, Phys. Rev. Lett. **129**, 072001 (2022).
[40] G. Beuf, T. Lappi, and R. Paatelainen, Massive quarks in NLO dipole factorization for DIS: Transverse photon, Phys. Rev. D **106**, 034013 (2022).
[41] K. J. Golec-Biernat and M. Wusthoff, Saturation in diffractive deep inelastic scattering, Phys. Rev. D **60**, 114023 (1999).
[42] H. Kowalski, T. Lappi, and R. Venugopalan, Nuclear Enhancement of Universal Dynamics of High Parton Densities, Phys. Rev. Lett. **100**, 022303 (2008).
[43] N. Armesto, P. R. Newman, W. Słomiński, and A. M. Staśto, Inclusive diffraction in future electron-proton and electron-ion colliders, Phys. Rev. D **100**, 074022 (2019).
[44] R. Boussarie, A. V. Grabovsky, L. Szymanowski, and S. Wallon, Impact factor for high-energy two and three jets diffractive production, J. High Energy Phys. 09 (2014) 026.
[45] R. Boussarie, A. V. Grabovsky, L. Szymanowski, and S. Wallon, On the one loop $\gamma^{(*)} \to q\bar{q}$ impact factor and the exclusive diffractive cross sections for the production of two or three jets, J. High Energy Phys. 11 (2016) 149.
[46] R. Boussarie, A. V. Grabovsky, D. Y. Ivanov, L. Szymanowski, and S. Wallon, Next-to-Leading Order Computation of Exclusive Diffractive Light Vector Meson Production in a Saturation Framework, Phys. Rev. Lett. **119**, 072002 (2017).
[47] M. A. Escobedo and T. Lappi, Dipole picture and the non-relativistic expansion, Phys. Rev. D **101**, 034030 (2020).
[48] T. Lappi, H. Mäntysaari, and J. Penttala, Relativistic corrections to the vector meson light front wave function, Phys. Rev. D **102**, 054020 (2020).
[49] H. Mäntysaari and J. Penttala, Exclusive heavy vector meson production at next-to-leading order in the dipole picture, Phys. Lett. B **823**, 136723 (2021).
[50] H. Mäntysaari and J. Penttala, Exclusive production of light vector mesons at next-to-leading order in the dipole picture, Phys. Rev. D **105**, 114038 (2022).
[51] H. Mäntysaari and J. Penttala, Complete calculation of exclusive heavy vector meson production at next-to-leading order in the dipole picture, J. High Energy Phys. 08 (2022) 247.
[52] E. Iancu, A. H. Mueller, and D. N. Triantafyllopoulos, Probing Parton Saturation and the Gluon Dipole via Diffractive Jet Production at the Electron-Ion Collider, Phys. Rev. Lett. **128**, 202001 (2022).
[53] Y. Hatta, B.-W. Xiao, and F. Yuan, Semi-inclusive diffractive deep inelastic scattering at small-x, arXiv:2205.08060.
[54] G. P. Lepage and S. J. Brodsky, Exclusive processes in perturbative quantum chromodynamics, Phys. Rev. D **22**, 2157 (1980).
[55] F. Hautmann, Z. Kunszt, and D. E. Soper, Diffractive Deeply Inelastic Scattering of Hadronic States with Small Transverse Size, Phys. Rev. Lett. **81**, 3333 (1998).
[56] F. Hautmann, Z. Kunszt, and D. E. Soper, Hard scattering factorization and light cone Hamiltonian approach to diffractive processes, Nucl. Phys. **B563**, 153 (1999).
[57] F. Hautmann and D. E. Soper, Color transparency in deeply inelastic diffraction, Phys. Rev. D **63**, 011501 (2001).
[58] M. Wusthoff, Large rapidity gap events in deep inelastic scattering, Phys. Rev. D **56**, 4311 (1997).
[59] K. J. Golec-Biernat and M. Wusthoff, Saturation in diffractive deep inelastic scattering, Phys. Rev. D **60**, 114023 (1999).
[60] K. J. Golec-Biernat and M. Wusthoff, Saturation in diffractive deep inelastic scattering, Phys. Rev. D **60**, 114023 (1999).
[61] C. Marquet, A unified description of diffractive deep inelastic scattering with saturation, Phys. Rev. D **76**, 094017 (2007).
[62] H. Kowalski, T. Lappi, C. Marquet, and R. Venugopalan, Nuclear enhancement and suppression of diffractive structure functions at high energies, Phys. Rev. C **78**, 045201 (2008).
[63] M. S. Kugeratski, V. P. Goncalves, and F. S. Navarra, Saturation in diffractive deep inelastic eA scattering, Eur. Phys. J. C **46**, 413 (2006).
[64] D. Bendova, J. Cepila, J. G. Contreras, t. V. P. Gonçalves, and M. Matas, Diffractive deeply inelastic scattering in future electron-ion colliders, Eur. Phys. J. C **81**, 211 (2021).

[65] H. Hänninen, Deep inelastic scattering in the dipole picture at next-to-leading order, Ph.D. thesis, University of Jyväskylä, 2021.

[66] S. Munier and A. Shoshi, Diffractive photon dissociation in the saturation regime from the Good and Walker picture, Phys. Rev. D **69**, 074022 (2004).

[67] M. L. Good and W. D. Walker, Diffraction disssociation of beam particles, Phys. Rev. **120**, 1857 (1960).

[68] H. I. Miettinen and J. Pumplin, Diffraction scattering and the parton structure of hadrons, Phys. Rev. D **18**, 1696 (1978).

[69] A. Caldwell and H. Kowalski, Investigating the gluonic structure of nuclei via J/ψ scattering, Phys. Rev. C **81**, 025203 (2010).

[70] H. Mäntysaari and B. Schenke, Evidence of Strong Proton Shape Fluctuations from Incoherent Diffraction, Phys. Rev. Lett. **117**, 052301 (2016).

[71] H. Mäntysaari, Review of proton and nuclear shape fluctuations at high energy, Rep. Prog. Phys. **83**, 082201 (2020).

[72] A. AKTAS *et al.* (H1 Collaboration), Measurement and QCD analysis of the diffractive deep-inelastic scattering cross-section at HERA, Eur. Phys. J. C **48**, 715 (2006).

[73] A. Aktas *et al.* (H1 Collaboration), Diffractive deep-inelastic scattering with a leading proton at HERA, Eur. Phys. J. C **48**, 749 (2006).

[74] A. Aktas *et al.* (H1 Collaboration), Diffractive open charm production in deep-inelastic scattering and photoproduction at HERA, Eur. Phys. J. C **50**, 1 (2007).

[75] F. D. Aaron *et al.* (H1 and ZEUS Collaborations), Combined inclusive diffractive cross sections measured with forward proton spectrometers in deep inelastic ep scattering at HERA, Eur. Phys. J. C **72**, 2175 (2012).

[76] J. D. Bjorken, J. B. Kogut, and D. E. Soper, Quantum electrodynamics at infinite momentum: Scattering from an external field, Phys. Rev. D **3**, 1382 (1971).

[77] J. B. Kogut and D. E. Soper, Quantum electrodynamics in the infinite momentum frame, Phys. Rev. D **1**, 2901 (1970).

[78] E. Iancu and Y. Mulian, Forward dijets in proton-nucleus collisions at next-to-leading order: The real corrections, J. High Energy Phys. 03 (2021) 005.

[79] G. A. Chirilli, B.-W. Xiao, and F. Yuan, Inclusive hadron productions in pA collisions, Phys. Rev. D **86**, 054005 (2012).

[80] T. Lappi and R. Paatelainen, The one loop gluon emission light cone wave function, Ann. Phys. (Amsterdam) **379**, 34 (2017).

[81] Y. V. Kovchegov and E. Levin, *Quantum Chromodynamics at High Energy* (Cambridge University Press, Cambridge, England, 2012), Vol. 33.

[82] J. Bartels, S. Gieseke, and C. F. Qiao, The $\gamma^* \to q\bar{q}$ reggeon vertex in next-to-leading order QCD, Phys. Rev. D **63**, 056014 (2001); Erratum, Phys. Rev. D **65**, 079902 (2002).

[83] J. Bartels, S. Gieseke, and A. Kyrieleis, The process $\gamma_L^* + q \to q\bar{q}g + q$: Real corrections to the virtual photon impact factor, Phys. Rev. D **65**, 014006 (2002).

[84] J. Bartels, D. Colferai, S. Gieseke, and A. Kyrieleis, NLO corrections to the photon impact factor: Combining real and virtual corrections, Phys. Rev. D **66**, 094017 (2002).

[85] J. Bartels and A. Kyrieleis, NLO corrections to the γ^* impact factor: First numerical results for the real corrections to γ_L^*, Phys. Rev. D **70**, 114003 (2004).

[86] T. Altinoluk, N. Armesto, G. Beuf, and A. H. Rezaeian, Diffractive dijet production in deep inelastic scattering and photon-hadron collisions in the color glass condensate, Phys. Lett. B **758**, 373 (2016).

[87] F. Dominguez, C. Marquet, B.-W. Xiao, and F. Yuan, Universality of unintegrated gluon distributions at small x, Phys. Rev. D **83**, 105005 (2011).

[88] J. Jalilian-Marian, A. Kovner, L. D. McLerran, and H. Weigert, The intrinsic glue distribution at very small x, Phys. Rev. D **55**, 5414 (1997).

[89] J. Jalilian-Marian, A. Kovner, A. Leonidov, and H. Weigert, The BFKL equation from the Wilson renormalization group, Nucl. Phys. **B504**, 415 (1997).

[90] J. Jalilian-Marian, A. Kovner, A. Leonidov, and H. Weigert, The Wilson renormalization group for low x physics: Towards the high density regime, Phys. Rev. D **59**, 014014 (1998).

[91] E. Iancu and L. D. McLerran, Saturation and universality in QCD at small x, Phys. Lett. B **510**, 145 (2001).

[92] E. Ferreiro, E. Iancu, A. Leonidov, and L. McLerran, Nonlinear gluon evolution in the color glass condensate. 2., Nucl. Phys. **A703**, 489 (2002).

[93] E. Iancu, A. Leonidov, and L. D. McLerran, The renormalization group equation for the color glass condensate, Phys. Lett. B **510**, 133 (2001).

[94] E. Iancu, A. Leonidov, and L. D. McLerran, Nonlinear gluon evolution in the color glass condensate. 1., Nucl. Phys. **A692**, 583 (2001).

[95] A. H. Mueller, A simple derivation of the JIMWLK equation, Phys. Lett. B **523**, 243 (2001).

[96] F. D. Aaron *et al.* (H1 and ZEUS Collaborations), Combined measurement and QCD analysis of the inclusive e+- p scattering cross sections at HERA, J. High Energy Phys. 01 (2010) 109.

[97] H. Abramowicz *et al.* (H1 and ZEUS Collaborations), Combination of measurements of inclusive deep inelastic $e^{\pm}p$ scattering cross sections and QCD analysis of HERA data, Eur. Phys. J. C **75**, 580 (2015).

[98] H. Abramowicz *et al.* (H1 and ZEUS Collaborations), Combination and QCD analysis of charm and beauty production cross-section measurements in deep inelastic ep scattering at HERA, Eur. Phys. J. C **78**, 473 (2018).

[99] H. Abramowicz *et al.* (H1 and ZEUS Collaborations), Combination and QCD analysis of charm production cross section measurements in deep-inelastic ep scattering at HERA, Eur. Phys. J. C **73**, 2311 (2013).

[100] T. Lappi and H. Mäntysaari, Single inclusive particle production at high energy from HERA data to proton-nucleus collisions, Phys. Rev. D **88**, 114020 (2013).

[101] J. L. Albacete, N. Armesto, J. G. Milhano, P. Quiroga-Arias, and C. A. Salgado, AAMQS: A non-linear QCD analysis of new HERA data at small-x including heavy quarks, Eur. Phys. J. C **71**, 1705 (2011).

[102] B. Ducloué, E. Iancu, G. Soyez, and D. N. Triantafyllopoulos, HERA data and collinearly-improved BK dynamics, Phys. Lett. B **803**, 135305 (2020).
[103] H. Mäntysaari and B. Schenke, Confronting impact parameter dependent JIMWLK evolution with HERA data, Phys. Rev. D **98**, 034013 (2018).
[104] H. Kowalski and D. Teaney, An impact parameter dipole saturation model, Phys. Rev. D **68**, 114005 (2003).
[105] H. Mäntysaari and P. Zurita, In depth analysis of the combined HERA data in the dipole models with and without saturation, Phys. Rev. D **98**, 036002 (2018).
[106] A. H. Rezaeian, M. Siddikov, M. Van de Klundert, and R. Venugopalan, Analysis of combined HERA data in the impact-parameter dependent Saturation model, Phys. Rev. D **87**, 034002 (2013).
[107] E. Iancu and Y. Mulian, Forward trijet production in proton–nucleus collisions, Nucl. Phys. **A985**, 66 (2019).
[108] K. J. Golec-Biernat and M. Wusthoff, Diffractive parton distributions from the saturation model, Eur. Phys. J. C **20**, 313 (2001).
[109] J. Bartels, K. J. Golec-Biernat, and K. Peters, On the dipole picture in the nonforward direction, Acta Phys. Pol. B **34**, 3051 (2003).
[110] H. Kowalski, L. Motyka, and G. Watt, Exclusive diffractive processes at HERA within the dipole picture, Phys. Rev. D **74**, 074016 (2006).
[111] Y. Hatta, B.-W. Xiao, and F. Yuan, Gluon tomography from deeply virtual compton scattering at small-x, Phys. Rev. D **95**, 114026 (2017).
[112] H. Mäntysaari, K. Roy, F. Salazar, and B. Schenke, Gluon imaging using azimuthal correlations in diffractive scattering at the electron-ion collider, Phys. Rev. D **103**, 094026 (2021).
[113] A. Dumitru, H. Mäntysaari, and R. Paatelainen, Color charge correlations in the proton at NLO: Beyond geometry based intuition, Phys. Lett. B **820**, 136560 (2021).
[114] H. Mäntysaari, N. Mueller, and B. Schenke, Diffractive dijet production and wigner distributions from the color glass condensate, Phys. Rev. D **99**, 074004 (2019).
[115] F. Salazar and B. Schenke, Diffractive dijet production in impact parameter dependent saturation models, Phys. Rev. D **100**, 034007 (2019).
[116] E. Iancu and A. H. Rezaeian, Elliptic flow from color-dipole orientation in pp and pA collisions, Phys. Rev. D **95**, 094003 (2017).
[117] Y. V. Kovchegov and E. Levin, Diffractive dissociation including multiple pomeron exchanges in high parton density QCD, Nucl. Phys. **B577**, 221 (2000).
[118] J. Bartels, H. Jung, and M. Wusthoff, Quark - anti-quark gluon jets in DIS diffractive dissociation, Eur. Phys. J. C **11**, 111 (1999).
[119] B. Z. Kopeliovich, A. Schafer, and A. V. Tarasov, Non-perturbative effects in gluon radiation and photoproduction of quark pairs, Phys. Rev. D **62**, 054022 (2000).
[120] Y. V. Kovchegov, Diffractive gluon production in proton nucleus collisions and in DIS, Phys. Rev. D **64**, 114016 (2001); Erratum, Phys. Rev. D **68**, 039901 (2003).
[121] K. J. Golec-Biernat and C. Marquet, Testing saturation with diffractive jet production in deep inelastic scattering, Phys. Rev. D **71**, 114005 (2005).
[122] J. D. Bjorken and J. B. Kogut, Correspondence arguments for high-energy collisions, Phys. Rev. D **8**, 1341 (1973).
[123] A. Hebecker, Diffractive parton distributions in the semi-classical approach, Nucl. Phys. **B505**, 349 (1997).
[124] W. Buchmuller, T. Gehrmann, and A. Hebecker, Inclusive and diffractive structure functions at small x, Nucl. Phys. **B537**, 477 (1999).
[125] E. Iancu, A. H. Mueller, D. N. Triantafyllopoulos, and S. Y. Wei, Gluon dipole factorisation for diffractive dijets, arXiv:2207.06268.
[126] B. Ducloué, H. Hänninen, T. Lappi, and Y. Zhu, Deep inelastic scattering in the dipole picture at next-to-leading order, Phys. Rev. D **96**, 094017 (2017).