

# Proof of sum rules for double parton distributions in QCD

M. Diehl<sup>1</sup>, P. Plößl<sup>2</sup> and A. Schäfer<sup>2</sup>

<sup>1</sup> Deutsches Elektronen-Synchrotron DESY, 22603 Hamburg, Germany

<sup>2</sup> Institut für Theoretische Physik, Universität Regensburg, 93040 Regensburg, Germany

**Abstract:** Double hard scattering can play an important role for producing multi-particle final states in hadron-hadron collisions. The associated cross sections depend on double parton distributions, which at present are only weakly constrained by theory or measurements. A set of sum rules for these distributions has been proposed by Gaunt and Stirling some time ago. We give a proof for these sum rules at all orders in perturbation theory, including a detailed analysis of the renormalisation of ultraviolet divergences. As a by-product of our study, we obtain the form of the inhomogeneous evolution equation for double parton distributions at arbitrary perturbative order.

# Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Definitions and sum rules</b>	<b>3</b>
<b>3</b>	<b>Analysis of low-order graphs and its limitations</b>	<b>5</b>
3.1	Sum rules with a gluon PDF	5
3.2	Sum rules with a quark PDF	9
<b>4</b>	<b>Light-cone perturbation theory</b>	<b>11</b>
<b>5</b>	<b>All-order proof for bare distributions</b>	<b>13</b>
5.1	Representation of PDFs and DPDs in LCPT	13
5.2	Equality between PDF and DPD graphs	15
5.3	Number sum rule	17
5.4	Momentum sum rule	18
<b>6</b>	<b>Renormalisation</b>	<b>18</b>
6.1	Convolution integrals	18
6.2	Renormalisation of DPDs	20
6.3	Number sum rule	23
6.4	Momentum sum rule	24
<b>7</b>	<b>DPD evolution and its consequences</b>	<b>25</b>
<b>8</b>	<b>Conclusion</b>	<b>28</b>
	<b>References</b>	<b>28</b>

---

## 1 Introduction

In high-energy hadron hadron collisions, two or more partons in each incoming hadron may simultaneously take part in hard-scattering subprocesses. The importance of such multiparton interactions increases with the collision energy, because the density of partons in a hadron increases as their momentum fraction becomes smaller. Theoretical predictions with highest possible accuracy are required for a wide range of final states and kinematic regions in order to use the full potential of the LHC and of possible future hadron colliders. It is therefore highly desirable to develop a theory and phenomenology of multiparton interactions based on first principles in QCD. The most important contribution typically comes from double parton scattering (DPS), which is the subject of this work. To compute DPS cross sections one needs double parton distributions, which give the probability density for finding two specified partons inside a hadron. Our knowledge of these distributions is still very poor, given their dependence on several variables and the difficulty to separate the DPS contribution to a physical process from the contribution due to single hard scattering. Often one makes the simplest assumption that the two partons are entirely uncorrelated. This can at best be a first approximation. In the region of relatively large momentum fractions many studies in dynamical models find that correlations are actually strong (see [1] for a recent review).

In such a situation, theoretical constraints on double parton distributions (DPDs) are valuable. One type of constraint comes from the mechanism in which the two partons originate

from the short-distance splitting of a single parton. This mechanism, which we will call  $1 \rightarrow 2$  splitting, dominates DPDs at small distance between the two partons and can be expressed in terms of perturbative splitting functions and the usual single parton distributions (PDFs). The impact of  $1 \rightarrow 2$  splitting on DPDs and DPS cross sections has been investigated from several points of view, see e.g. [2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13]. A different constraint is provided by the DPD sum rules proposed by Gaunt and Stirling [14], which express the conservation of quark flavour and of momentum and nicely fit together with the interpretation of DPDs as probability densities. Attempts to construct DPDs satisfying these sum rules were made in [14] and [15]. A study of  $W^+W^+$  and  $W^-W^-$  production with the DPDs proposed in [14] can be found in [16].

DPDs depend on a renormalisation scale, just like ordinary PDFs. This dependence is described by a generalisation of the DGLAP evolution equations. Due to the  $1 \rightarrow 2$  splitting mechanism, these equations contain an inhomogeneous term whose form at leading order (LO) has been extensively discussed in the literature [17, 18, 19, 20]. Gaunt and Stirling noted that if the sum rules they postulated are valid at some scale, then their validity is preserved by LO evolution to any other scale [14]. More detailed analyses of how the momentum sum rule remains valid in the presence of parton splitting can be found in [8] and [21]. Still lacking is however a full proof that the sum rules hold at some starting scale. A derivation using the light-cone wave function representation is given in appendix C of [22]. This framework formalises the physical picture of the parton model and in particular allows one to keep track of kinematic and combinatorial factors. An explicit analysis of ultraviolet (UV) divergences and of the associated scale dependence is however missing in [22]. The same holds for rapidity divergences that are present in light-cone wave functions (but cancel in the DPDs appearing in the sum rules).

The aim of the present paper is to provide an explicit proof of the DPD sum rules in QCD. After defining DPDs and stating the sum rules in section 2, we show in section 3 how the sum rules arise at the first nontrivial order in a simple perturbative toy model. We recall the essentials of light-cone perturbation theory in section 4 and then use this formalism in section 5 to give an all-order proof of the sum rules for bare (i.e. unrenormalised) DPDs. In section 6 we show how DPDs are renormalised and establish that the sum rules remain valid if UV divergences are subtracted in a suitable scheme. In Section 7 we explore the consequences of our analysis on the evolution of DPDs: we obtain the general form of the inhomogeneous term beyond LO (confirming the NLO result given in [20]), we derive sum rules for the associated evolution kernels, and we cross check that the sum rules are preserved by evolution at any order in  $\alpha_s$ . Our findings are briefly summarised in section 8.

## 2 Definitions and sum rules

In this section, we recall the definition of PDFs and DPDs and then state the sum rules to be proven in the remainder of this work. We restrict ourselves to unpolarised partons and transverse-momentum integrated distributions throughout. Bare PDFs and DPDs are defined as

$$\begin{aligned}
 f_B^{j_1}(x; \mu) &= (x_1 p^+)^{-n_1} \int \frac{dz_1^-}{2\pi} e^{ix_1 z_1^- p^+} \langle p | \mathcal{O}_{j_1}(0, z_1) | p \rangle \Big|_{z_1^+ = 0, \mathbf{z}_1 = \mathbf{0}}, \\
 F_B^{j_1 j_2}(x_1, x_2, \mathbf{y}) &= (x_1 p^+)^{-n_1} (x_2 p^+)^{-n_2} 2p^+ \int dy^- \frac{dz_1^-}{2\pi} \frac{dz_2^-}{2\pi} e^{i(x_1 z_1^- + x_2 z_2^-) p^+} \\
 &\quad \times \langle p | \mathcal{O}_{j_1}(y, z_1) \mathcal{O}_{j_2}(0, z_2) | p \rangle \Big|_{z_1^+ = z_2^+ = y^+ = 0, \mathbf{z}_1 = \mathbf{z}_2 = \mathbf{0}}, \tag{1}
 \end{aligned}$$

where  $n_j = 0$  if parton  $j$  is a fermion,  $n_j = 1$  if it is a gluon, and  $n_j = -1$  if one considers scalar partons. We use light-cone coordinates  $v^\pm = (v^0 \pm v^3)/\sqrt{2}$  for any four-vector  $v^\mu$  and write its transverse part in boldface,  $\mathbf{v} = (v^1, v^2)$ . The twist-two operators read

$$\begin{aligned}\mathcal{O}_q(y, z) &= \frac{1}{2} \bar{q}\left(y - \frac{z}{2}\right) \gamma^+ q\left(y + \frac{z}{2}\right), & \mathcal{O}_{\bar{q}}(y, z) &= -\frac{1}{2} \bar{q}\left(y + \frac{z}{2}\right) \gamma^+ q\left(y - \frac{z}{2}\right), \\ \mathcal{O}_g(y, z) &= G^{+i}\left(y - \frac{z}{2}\right) G^{+i}\left(y + \frac{z}{2}\right)\end{aligned}\quad (2)$$

for unpolarised quarks, antiquarks and gluons, respectively and are constructed from bare (i.e. unrenormalised) fields. The transverse index  $i$  in (2) is to be summed over. We only consider colour-singlet DPDs here, so that the colour indices of the quark or gluon fields in (2) are implicit and to be summed over. For scalar partons we have

$$\mathcal{O}_\phi(y, z) = \phi\left(y - \frac{z}{2}\right) \phi\left(y + \frac{z}{2}\right). \quad (3)$$

Throughout this work, we use light-cone gauge  $A^+ = 0$ , so that Wilson lines do not appear in the above operators. Note that, just like ordinary PDFs, colour-singlet DPDs are free of rapidity divergences. Modifications of the standard light-like Wilson lines in Feynman gauge (or their equivalents in other gauges) can be employed to regulate such divergences as discussed for DPS in [23, 24], but this is of no concern for our present work.

The DPD in (1) depends on the transverse distance  $\mathbf{y}$  between the two partons. The discussion of sum rules requires introducing their Fourier transform to transverse momentum space,

$$F_B^{j_1 j_2}(x_1, x_2, \mathbf{\Delta}) = \int d^{D-2} \mathbf{y} e^{i\mathbf{y}\mathbf{\Delta}} F_B^{j_1 j_2}(x_1, x_2, \mathbf{y}), \quad (4)$$

where  $D = 4 - 2\varepsilon$  is the number of space-time dimensions in dimensional regularisation.

For the discussion of graphs in the next sections, it is useful to introduce the momentum space Green functions

$$\begin{aligned}\mathcal{G}^{j_1}(k_1) &= \int d^D z_1 e^{ik_1 z_1} \langle p | \mathcal{O}_{j_1}(z_1) | p \rangle, \\ \mathcal{G}^{j_1 j_2}(k_1, k_2, \Delta) &= \int d^D z_1 e^{ik_1 z_1} \int d^D z_2 e^{ik_2 z_2} \int d^D y e^{-iy\Delta} \langle p | \mathcal{O}_{j_1}(y, z_1) \mathcal{O}_{j_2}(0, z_2) | p \rangle,\end{aligned}\quad (5)$$

in terms of which we have

$$f_B^{j_1}(x_1) = (x_1 p^+)^{-n_1} \int \frac{dk_1^- d^{D-2} \mathbf{k}_1}{(2\pi)^D} \mathcal{G}^{j_1}(k_1), \quad (6)$$

$$F_B^{j_1 j_2}(x_1, x_2, \mathbf{\Delta}) = \left[ \prod_{i=1}^2 (x_i p^+)^{-n_i} \int \frac{dk_i^- d^{D-2} \mathbf{k}_i}{(2\pi)^D} \right] 2p^+ \int \frac{d\Delta^-}{2\pi} \mathcal{G}^{j_1 j_2}(k_1, k_2, \Delta). \quad (7)$$

The above distributions contain ultraviolet divergences and require renormalisation. Renormalised distributions will be denoted without the subscript  $B$  and have an additional dependence on the renormalisation scale  $\mu$ . We postpone a detailed discussion to section 6 but already note here that the integral over  $\mathbf{y}$  in (4) diverges at  $\mathbf{y} = \mathbf{0}$  in  $D = 4$  dimensions and hence requires additional renormalisation. This divergence is due to the  $1 \rightarrow 2$  splitting mechanism mentioned in the introduction and leads to the inhomogeneous term in the evolution equations for DPDs. It only appears for DPDs depending on the momentum  $\mathbf{\Delta}$ , but not for their counterparts depending on the distance  $\mathbf{y}$ , as was already noted in [2].

The DPD sum rules hold for the momentum space distributions at zero transverse momentum, which we abbreviate as

$$F_{j_1 j_2}(x_i; \mu) = F_{j_1 j_2}(x_i, \mathbf{\Delta} = \mathbf{0}; \mu). \quad (8)$$

Up to the additional renormalisation just mentioned, these distributions correspond to the integral of  $F_{j_1 j_2}(x_i, \mathbf{y}; \mu)$  over all  $\mathbf{y}$ , as can be seen in (4). The momentum sum rule for DPDs reads

$$\sum_{j_2} \int_0^{1-x_1} dx_2 x_2 F^{j_1 j_2}(x_1, x_2; \mu) = (1-x_1) f^{j_1}(x_1; \mu), \quad (9)$$

whereas the conservation of quark flavour in QCD corresponds to the number sum rule

$$\int_0^{1-x_1} dx_2 F^{j_1 j_2, v}(x_1, x_2; \mu) = (N_{j_2, v} + \delta_{j_1, \bar{j}_2} - \delta_{j_1, j_2}) f^{j_1}(x_1; \mu), \quad (10)$$

where  $j_2$  denotes a quark or an antiquark. The parton label  $j_v$  indicates the difference of parton and antiparton distributions, i.e.  $F_{j_1 j_2, v} = F_{j_1 j_2} - F_{j_1 \bar{j}_2}$ . The number of valence partons  $j$  in a hadron is denoted by  $N_{j_v}$ , so that e.g.  $N_{u_v} = \int dx f^{u_v}(x) = 2$  in a proton. Using the convention that  $\bar{j}$  denotes a quark if  $j$  is an antiquark, we have  $N_{\bar{u}_v} = 2$  in an antiproton.

### 3 Analysis of low-order graphs and its limitations

In this section, we show for a simple example how the DPD sum rules for bare distributions can be obtained from Feynman graphs in covariant perturbation theory. We consider all graphs at lowest order in  $\alpha_s$ . In a second example, we exhibit the limitations of this covariant approach. This leads us to use light-cone perturbation theory in the following sections, where we formulate a proof that is valid at all orders in  $\alpha_s$ . It is clear that neither covariant nor light-cone perturbation theory are suitable for actually computing parton distributions, which are non-perturbative objects. We must thus assume that general properties of Green functions – in our case the sum rules – remain valid beyond perturbation theory. This is similar to the spirit of perturbative proofs of factorisation in QCD [25, 26].

We use a toy model with scalar “quarks” of two flavours,  $u$  and  $d$ , which we take to be mass degenerate. The coupling between these quarks and the gluons is as required by gauge invariance. We consider a scalar “hadron” that has a pointlike coupling to  $u$  and  $\bar{d}$  and compute parton distributions at lowest order in perturbation theory.

#### 3.1 Sum rules with a gluon PDF

Let us first consider the case in which parton 1 in the sum rules is a gluon. At lowest order, the gluon PDF appearing on the r.h.s. of the sum rules is given by four graphs, two of which are depicted in figure 1. The other two are obtained by reversing the arrows on the quark lines (so that in graph a the gluon couples to the  $\bar{d}$  instead of the  $u$ ). They lead to identical expressions due to the symmetries of our model, viz. charge conjugation and the identical

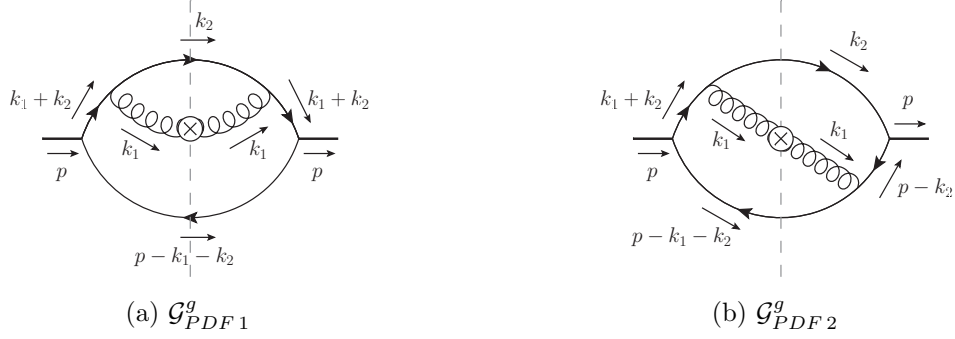


Figure 1: Two graphs contributing to the gluon PDF at order  $\mathcal{O}(\alpha_s)$  in our model. Two more graphs are obtained by reversing the arrows that indicate the flow of quark number.

masses of the two quark flavours. The contributions of the graphs to the gluon PDF are

$$f_1^g(x_1) = 4 \int d\Gamma_{PDF} (x_1 \mathbf{k}_2 - x_2 \mathbf{k}_1)^2 \times \left[ ((k_1 + k_2)^2 - m^2 + i\epsilon) (k_1^2 + i\epsilon) (k_1^2 - i\epsilon) ((k_1 + k_2)^2 - m^2 - i\epsilon) \right]^{-1}, \quad (11)$$

$$f_2^g(x_1) = 4 \int d\Gamma_{PDF} (x_1 \mathbf{k}_2 - x_2 \mathbf{k}_1) (x_2 \mathbf{k}_1 - x_1 \mathbf{k}_2 - \mathbf{k}_1) \times \left[ ((k_1 + k_2)^2 - m^2 + i\epsilon) (k_1^2 + i\epsilon) (k_1^2 - i\epsilon) ((p - k_2)^2 - m^2 - i\epsilon) \right]^{-1}, \quad (12)$$

where the integration element  $d\Gamma_{PDF}$  is given by

$$d\Gamma_{PDF} = \frac{g^2 \mu^{D-4} C_F p^+}{x_1} \frac{dk_1^- d^{D-2} \mathbf{k}_1}{(2\pi)^D} \frac{d^D k_2}{(2\pi)^D} \times 2\pi \delta(k_2^2 - m^2) 2\pi \delta((p - k_1 - k_2)^2 - m^2). \quad (13)$$

Here  $g$  denotes the strong coupling and  $m$  the quark mass. For simplicity we set the coupling between the hadron and the quarks to 1. A factor of two for the diagrams with reversed arrows on the quark lines is included in these expressions. For brevity, we omit the subscript  $B$  for bare distributions throughout this section.

To obtain the graphs for the DPDs appearing in the sum rules, one can start from the PDF graphs and insert the operator for parton 2 on one of the lines that go across the final state cut in the PDF, as illustrated in figure 2. For our specific case, the result of this procedure is

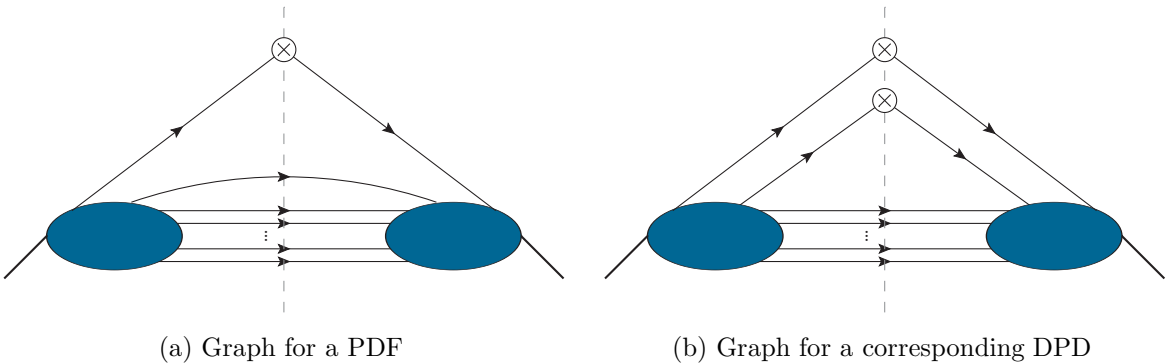


Figure 2: Transition from a given PDF graph to a corresponding DPD graph.

shown in figure 3. The operator (3) for scalar quarks or antiquarks simply provides a factor 1 in the graphs, and we obtain

$$\begin{aligned}
F_{1.1}^{gu}(x_1, x_2; \mu) &= 4 \int d\Gamma_{DPD,1} x_2 (x_1 \mathbf{k}_2 - x_2 \mathbf{k}_1)^2 \\
&\times \left[ ((k_1 + k_2)^2 - m^2 + i\epsilon) ((k_2 + \Delta/2)^2 - m^2 + i\epsilon) ((k_1 - \Delta/2)^2 + i\epsilon) \right]^{-1} \\
&\times \left[ ((k_1 + \Delta/2)^2 - i\epsilon) ((k_2 - \Delta/2)^2 - m^2 - i\epsilon) ((k_1 + k_2)^2 - m^2 - i\epsilon) \right]^{-1}, \quad (14)
\end{aligned}$$

$$\begin{aligned}
F_{1.2}^{g\bar{d}}(x_1, 1 - x_1 - x_2) &= 4 \int d\Gamma_{DPD,2} (1 - x_1 - x_2) (x_1 \mathbf{k}_2 - x_2 \mathbf{k}_1)^2 \\
&\times \left[ ((k_1 + k_2)^2 - m^2 + i\epsilon) ((p - k_1 - k_2 + \Delta/2)^2 - m^2 + i\epsilon) ((k_1 - \Delta/2)^2 + i\epsilon) \right]^{-1} \\
&\times \left[ ((k_1 + \Delta/2)^2 - i\epsilon) ((p - k_1 - k_2 - \Delta/2)^2 - m^2 - i\epsilon) ((k_1 + k_2)^2 - m^2 - i\epsilon) \right]^{-1}, \quad (15)
\end{aligned}$$

$$\begin{aligned}
F_{2.1}^{gu}(x_1, x_2) &= 4 \int d\Gamma_{DPD,1} x_2 (x_1 \mathbf{k}_2 - x_2 \mathbf{k}_1) (x_2 \mathbf{k}_1 - x_1 \mathbf{k}_2 - \mathbf{k}_1) \\
&\times \left[ ((k_1 + k_2)^2 - m^2 + i\epsilon) ((k_2 + \Delta/2)^2 - m^2 + i\epsilon) ((k_1 - \Delta/2)^2 + i\epsilon) \right]^{-1} \\
&\times \left[ ((k_1 + \Delta/2)^2 - i\epsilon) ((k_2 - \Delta/2)^2 - m^2 - i\epsilon) ((p - k_2)^2 - m^2 - i\epsilon) \right]^{-1}, \quad (16)
\end{aligned}$$

$$\begin{aligned}
F_{2.2}^{g\bar{d}}(x_1, 1 - x_1 - x_2) &= 4 \int d\Gamma_{DPD,2} (1 - x_1 - x_2) (x_1 \mathbf{k}_2 - x_2 \mathbf{k}_1) (x_2 \mathbf{k}_1 - x_1 \mathbf{k}_2 - \mathbf{k}_1) \\
&\times \left[ ((k_1 + k_2)^2 - m^2 + i\epsilon) ((p - k_1 - k_2 + \Delta/2)^2 - m^2 + i\epsilon) ((k_1 - \Delta/2)^2 + i\epsilon) \right]^{-1} \\
&\times \left[ ((k_1 + \Delta/2)^2 - i\epsilon) ((p - k_1 - k_2 - \Delta/2)^2 - m^2 - i\epsilon) ((p - k_2)^2 - m^2 - i\epsilon) \right]^{-1}, \quad (17)
\end{aligned}$$

where the integration elements  $d\Gamma_{DPD,i}$  ( $i = 1, 2$ ) are given by

$$d\Gamma_{DPD,i} = \frac{2g^2 \mu^{D-4} C_F (p^+)^3}{x_1} \frac{dk_1^- d^{D-2} \mathbf{k}_1}{(2\pi)^D} \frac{dk_2^- d^{D-2} \mathbf{k}_2}{(2\pi)^D} \frac{d\Delta^-}{2\pi} 2\pi \delta(\ell_i^2 - m^2) \quad (18)$$

with  $\ell_1 = p - k_1 - k_2$  and  $\ell_2 = k_2$ . Notice the particular form of the second momentum fraction arguments in  $F^{g\bar{d}}$ , which will turn out to be useful for showing the sum rules. Reversing the arrows on the quark lines in figure 3, one obtains the remaining one-loop contributions to the DPDs with one gluon. Given the symmetries of our model, we have  $F_{1.1}^{gu}(x_1, x_2; \mu) = F_{1.1}^{g\bar{d}}(x_1, x_2; \mu)$  and likewise for the other three graphs.

We notice that the expressions from PDF and DPD graphs that correspond to each other are already quite similar. While the momentum dependent part of the numerators agrees exactly, the main difference are the two additional denominator factors in the DPD, which are for the line that runs across the final state cut in the PDF. In order to get rid of these, we perform the integrations over their minus momentum components using the theorem of residues. After suitable variable substitutions, we can close the integration contour in such a way that we only pick up the poles of the two lines corresponding to parton 2. This removes the additional denominator factors in the DPD and sets the corresponding momenta to their on-shell value, just like it is the case in the PDF. The relevant substitutions read

$$k_1^- - \Delta^-/2 = k^-, \quad k_1^- + \Delta^-/2 = k'^-, \quad k_1^- + k_2^- = K^- \quad (19)$$

for  $F_{1.1}^{gu}$  and  $F_{2.1}^{gu}$ , and

$$k_1^- - \Delta^-/2 = k^-, \quad k_1^- + \Delta^-/2 = k'^- \quad (20)$$

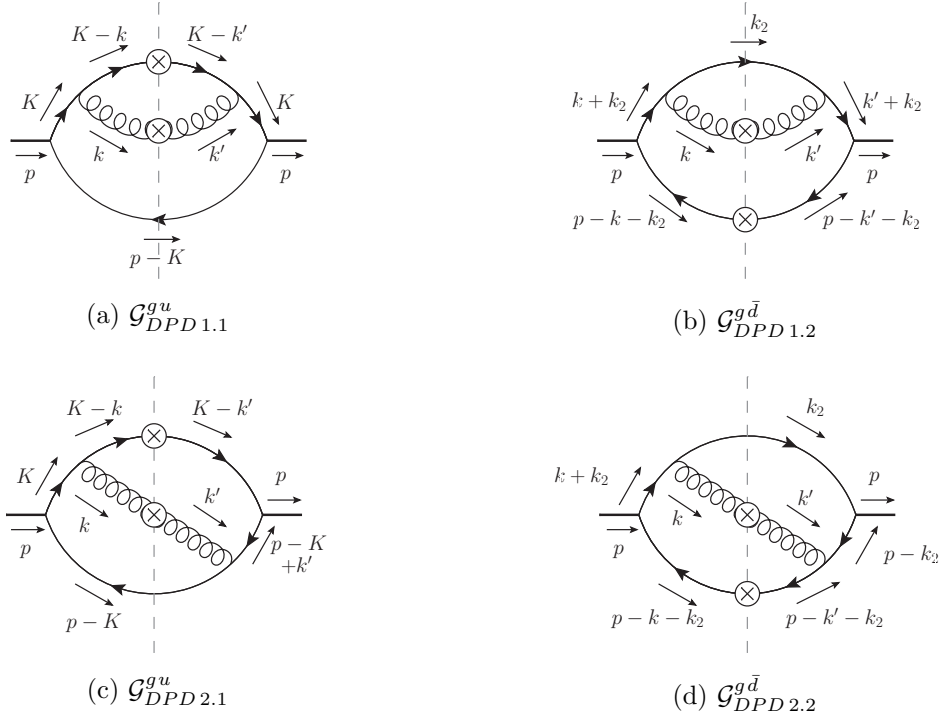


Figure 3: Graphs for gluon-quark or gluon-antiquark DPDs corresponding to the gluon PDF graphs in figure 1. The loop momenta are defined in (19) and (20), and each graph is for the momentum fractions specified in (15) to (17). Four more graphs are obtained by reversing the direction of arrows on the quark lines.

for  $F_{1.2}^{g\bar{d}}$  and  $F_{2.2}^{g\bar{d}}$ , where in the last case  $k_2^-$  is kept as an integration variable. The integration over  $k^-$  sets the line corresponding to parton 2 on the l.h.s. of the cut to its on-shell value, while the same is achieved on the r.h.s. of the cut by integrating over  $k'^-$ . After performing the integrations over all minus components, we obtain for the expressions in (11) to (17)

$$f_1^g(x_1) = \int_0^{1-x_1} dx_2 x_2 (1-x_1-x_2)^3 \int d\Gamma \frac{(x_1 \mathbf{k}_2 - x_2 \mathbf{k}_1)^2}{((\mathbf{k}_1 + \mathbf{k}_2)^2 + m^2)^2 \mathcal{D}^2}, \quad (21)$$

$$f_2^g(x_1) = \frac{1}{2} \int_0^{1-x_1} dx_2 x_2^2 (1-x_1-x_2)^2 \int d\Gamma \frac{(x_1 \mathbf{k}_2 - x_2 \mathbf{k}_1)^2 - x_2 \mathbf{k}_1^2 + x_1 \mathbf{k}_1 \mathbf{k}_2}{(\mathbf{k}_2^2 + m^2) ((\mathbf{k}_1 + \mathbf{k}_2)^2 + m^2) \mathcal{D}^2}, \quad (22)$$

and

$$\begin{aligned} F_{1.1}^{gu}(x_1, x_2) &= F_{1.2}^{g\bar{d}}(x_1, 1-x_1-x_2) \\ &= \frac{1}{2} \int_0^{1-x_1} dx_2 x_2 (1-x_1-x_2)^3 \int d\Gamma \frac{(x_1 \mathbf{k}_2 - x_2 \mathbf{k}_1)^2}{((\mathbf{k}_1 + \mathbf{k}_2)^2 + m^2)^2 \mathcal{D}^2}, \end{aligned} \quad (23)$$

$$\begin{aligned} F_{2.1}^{gu}(x_1, x_2) &= F_{2.2}^{g\bar{d}}(x_1, 1-x_1-x_2) \\ &= \frac{1}{4} x_2^2 (1-x_1-x_2)^2 \int d\Gamma \frac{(x_1 \mathbf{k}_2 - x_2 \mathbf{k}_1)^2 - x_2 \mathbf{k}_1^2 + x_1 \mathbf{k}_1 \mathbf{k}_2}{(\mathbf{k}_2^2 + m^2) ((\mathbf{k}_1 + \mathbf{k}_2)^2 + m^2) \mathcal{D}^2}, \end{aligned} \quad (24)$$

where we have abbreviated

$$\mathcal{D} = x_1 \mathbf{k}_2^2 + x_2 \mathbf{k}_1^2 - (x_1 \mathbf{k}_2 - x_2 \mathbf{k}_1)^2 + x_1 (1-x_1) m^2 \quad (25)$$



and the measure for the remaining integrations is

$$d\Gamma = \frac{g^2 \mu^{D-4} C_F}{x_1} \frac{d^{D-2} \mathbf{k}_1}{(2\pi)^{D-1}} \frac{d^{D-2} \mathbf{k}_2}{(2\pi)^{D-1}}. \quad (26)$$

The similarity between PDF and DPD expressions has become very close. To show the sum rules, we note that at order  $\alpha_s$ , the only nonzero DPDs involving a gluon are those we have already discussed. In particular,  $F^{g\bar{u}}$ ,  $F^{gd}$  and  $F^{gg}$  only appear at order  $\alpha_s^2$ . For the number sum rule for  $u$  quarks, we first combine all contributions from graphs 1.1 and 1.2:

$$\begin{aligned} \int_0^{1-x_1} dx_2 \left[ F_{1.1}^{gu}((x_1, x_2)) + F_{1.2}^{gu}(x_1, x_2) \right] &= \int_0^{1-x_1} dx_2 \left[ F_{1.1}^{gu}((x_1, x_2)) + F_{1.2}^{g\bar{d}}(x_1, x_2) \right] \\ &= \int_0^{1-x_1} dx_2 \left[ F_{1.1}^{gu}((x_1, x_2)) + F_{1.2}^{g\bar{d}}(x_1, 1-x_1-x_2) \right] = f_1^g(x_1). \end{aligned} \quad (27)$$

In the first step we used the symmetry between  $u$  and  $\bar{d}$  in our model, and in the second step we performed a change of variables in  $F^{g\bar{d}}$ . The last equality is easily seen from the explicit expressions in (21) and (23). One readily derives the analogue of (27) for the contributions from graphs 2.1 and 2.2 and hence for the sum over all graphs. Since  $N_{uv} = 1$  in our model, this shows that the number sum rule for  $u$  quarks is fulfilled. In the same manner, one can show the number sum rule for  $\bar{d}$  quarks. For the momentum sum rule, we start again with the contributions from graphs 1.1 and 1.2:

$$\begin{aligned} \int_0^{1-x_1} dx_2 x_2 \left[ F_{1.1}^{gu}((x_1, x_2)) + F_{1.2}^{gu}(x_1, x_2) + F_{1.1}^{g\bar{d}}((x_1, x_2)) + F_{1.2}^{g\bar{d}}(x_1, x_2) \right] \\ = 2 \int_0^{1-x_1} dx_2 \left[ x_2 F_{1.1}^{gu}((x_1, x_2)) + x_2 F_{1.2}^{g\bar{d}}(x_1, x_2) \right] \\ = 2 \int_0^{1-x_1} dx_2 \left[ x_2 F_{1.1}^{gu}((x_1, x_2)) + (1-x_1-x_2) F_{1.2}^{g\bar{d}}(x_1, 1-x_1-x_2) \right] \\ = 2 \int_0^{1-x_1} dx_2 \left[ x_2 F_{1.1}^{gu}((x_1, x_2)) + (1-x_1-x_2) F_{1.1}^{gu}(x_1, x_2) \right] = (1-x_1) f_1^g(x_1), \end{aligned} \quad (28)$$

The first two steps are justified just like their analogues in (27), and the last two steps make again use of (21) and (23). An analogous relation can be derived for graphs 2.1 and 2.2 and thus for the sum over all graphs, which confirms the momentum sum rule.

We have seen that both sum rules hold individually for each PDF graph and the set of corresponding DPD graphs. This is already suggested by figure 2 and will remain true in the all-order proof in section 5. However, a crucial step in the preceding derivation was that for each DPD graph we could perform the integrations over minus momenta in such a way that, after applying the theorem of residues, the momentum of parton 2 to the left and to the right of the final state cut was set on shell. This is not readily possible for other graphs, as we shall now see.

### 3.2 Sum rules with a quark PDF

We consider now the case in which parton 1 in the sum rules is a  $u$  quark (the expressions for  $\bar{d}$  quarks are identical for symmetry reasons). At order  $\alpha_s$  the number of graphs contributing

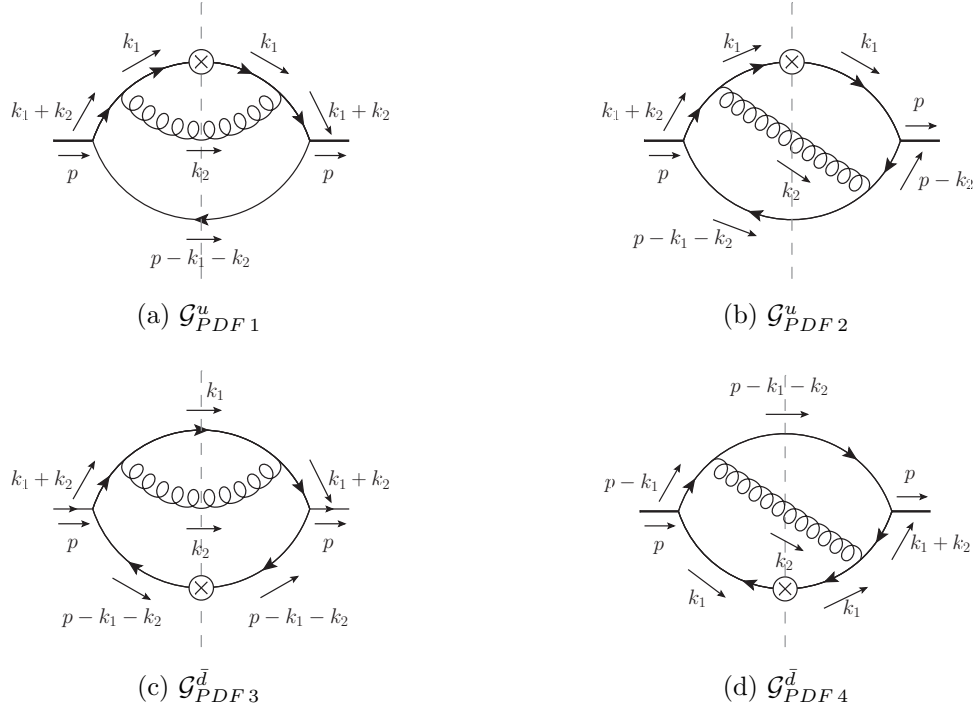


Figure 4: Real emission graphs contributing to quark or antiquark distributions. Four more graphs are obtained by reversing the arrow on the quark line.

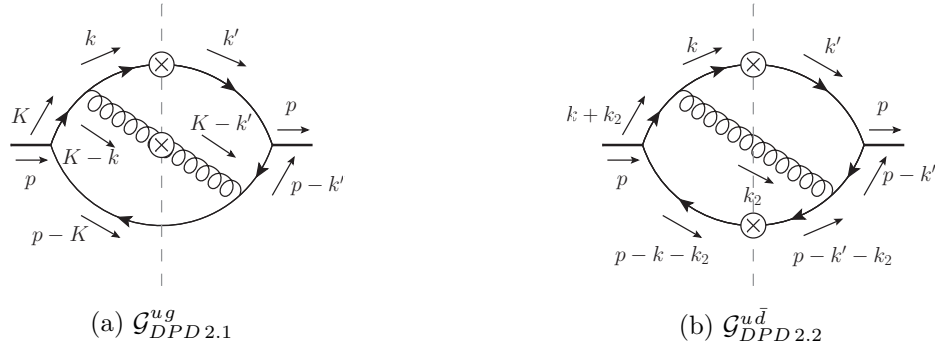


Figure 5: DPD graphs corresponding to the PDF graph in figure 4b.

to each PDF is much higher than in the previous section. There are the real emission graphs in figure 4, as well as graphs with a cut quark loop and a vertex or propagator correction to the left or to the right of the cut. For the graph in figure 4a and the corresponding graphs for DPDs, one can establish the validity of the sum rules exactly as in the previous section. This situation is however different for the remaining graphs of figure 4.

To illustrate where the problem is, let us consider the graph in figure 4b and the corresponding DPD graphs in figure 5. Before integrating over minus momenta, we make the same change of variables as previously, using (19) for  $F_{2.1}^{ug}$  and (20) for  $F_{2.1}^{u\bar{d}}$ . At this point, one finds that in both graphs it is impossible to close integration contours such that one picks up only the propagator poles of parton 2, given that on the r.h.s. of the cut their pole in  $k'^-$  is on the same side of the real axis as the propagator pole of the  $\bar{d}$  that directly couples to the

hadron. Performing the integrations over the minus momenta, we find

$$f_2^u(x_1) = x_1^2 \int_0^{1-x_1} dx_2 (1-x_1-x_2)^3 \int d\Gamma \frac{(x_1 \mathbf{k}_2 - x_2 \mathbf{k}_1)^2 - x_1 \mathbf{k}_2^2 + x_2 \mathbf{k}_1 \mathbf{k}_2}{((\mathbf{k}_1 + \mathbf{k}_2)^2 + m^2) \mathcal{D}_1^2 \mathcal{D}_2} \quad (29)$$

and

$$\begin{aligned} F_{2.1}^{ug}(x_1, x_2) &= F_{2.2}^{u\bar{d}}(x_1, 1-x_1-x_2) \\ &= x_1^2 (1-x_1-x_2)^3 \int d\Gamma \frac{(x_1 \mathbf{k}_2 - x_2 \mathbf{k}_1)^2 - x_1 \mathbf{k}_2^2 + x_2 \mathbf{k}_1 \mathbf{k}_2}{((\mathbf{k}_1 + \mathbf{k}_2)^2 + m^2) \mathcal{D}_1^2 \mathcal{D}_2} \\ &\quad - \frac{x_1^2 (1-x_1-x_2)^2 (1-x_1)}{x_2} \int d\Gamma \frac{(x_1 \mathbf{k}_2 - x_2 \mathbf{k}_1)^2 - x_1 \mathbf{k}_2^2 + x_2 \mathbf{k}_1 \mathbf{k}_2}{((\mathbf{k}_1 + \mathbf{k}_2)^2 + m^2) (\mathbf{k}_1^2 + m^2) \mathcal{D}_1 \mathcal{D}_2}, \end{aligned} \quad (30)$$

where

$$\begin{aligned} \mathcal{D}_1 &= x_1 \mathbf{k}_2^2 + x_2 \mathbf{k}_1^2 - (x_1 \mathbf{k}_2 - x_2 \mathbf{k}_1)^2 + x_2 (1-x_2) m^2, \\ \mathcal{D}_2 &= (x_1 \mathbf{k}_2 - x_2 \mathbf{k}_1)^2 + (1-2x_1) \mathbf{k}_2^2 + 2x_2 \mathbf{k}_1 \mathbf{k}_2 + x_2^2 m^2 \end{aligned} \quad (31)$$

and  $d\Gamma$  is defined as before in (26). Whereas the first term in (30) has the desired structure for establishing the sum rules, this is not the case for the second term, which originates from the “unwanted” propagator pole when the minus momentum integrals are performed using residues. We have explicitly checked that this extra term disappears when one sums over all contributing graphs, and the sum rules remain valid also in this example. However, the simple proof we used in the previous section no longer works.

This situation bears some similarity to the proof of cancellation of Glauber gluons in single [25, 26] or double hard scattering [27]. For simple cases, this cancellation can be established in covariant perturbation theory, using the theorem of residues for integrations over minus momenta in a similar way as here, but for more complicated graphs, that method turns out to be cumbersome [27]. It is not clear whether a general proof could be given at all in that framework. In the next sections we will see that light-cone perturbation theory provides a powerful tool for proving the sum rules at all orders in the strong coupling, as it is for establishing Glauber gluon cancellation [25, 26, 27].

## 4 Light-cone perturbation theory

Light-cone perturbation theory (LCPT, also called light-front perturbation theory) is quite similar to old-fashioned time ordered perturbation theory, with the difference that the vertices of a graph are ordered in “light-cone time”  $x^+ = (x^0 + x^3)/\sqrt{2}$  instead of ordinary time  $x^0$ . The rules of LCPT can be derived from regular covariant perturbation theory by performing the integrations over all internal minus momenta, thus setting all internal lines on-shell. This is for instance shown in chapter 7.2.3 in [26], whose normalisation conventions we adopt in the following. Further discussion of LCPT can be found in [28, 29, 30, 31, 32, 33, 34, 35].

Because in LCPT all internal lines are treated as on-shell, PDF and DPD graphs with identical  $x^+$  ordering of vertices automatically have the same denominator structure, which we will use to prove the validity of the DPD sum rules to all orders for bare distributions. For brevity we give only the basic rules of LCPT here and discuss details and subtleties when we encounter them. As in the rest of this paper, we use light-cone gauge  $A^+ = 0$  for the gluon.

1. Starting from a given Feynman graph, one assigns a light-cone time  $x_j^+$  to each vertex and considers all possible orderings of the  $x_j^+$ . When drawing LCPT graphs, we follow the convention that  $x^+$  increases from left to right on the l.h.s. of the final state cut, while on the r.h.s. it increases from right to left.
2. Coupling constants and vertex factors are the ones known from covariant perturbation theory. An exception are momentum dependent vertices, which are discussed below.
3. Plus and transverse momentum components,  $k_l^+$  and  $\mathbf{k}_l$ , of a line  $l$  are conserved at the vertices.
4. For each propagating line  $l$  in a graph, one has a factor  $1/(2k_l^+)$  together with a Heaviside step function  $\Theta(k_l^+)$  if the routing of  $k_l$  is from smaller to larger values of  $x^+$ .
5. Loop momenta  $\ell$  are integrated over their plus and transverse components with measure

$$\int \frac{d\ell^+ d^{D-2}\boldsymbol{\ell}}{(2\pi)^{D-1}}. \quad (32)$$

6. For each state  $i$  between two vertices at consecutive light-cone times  $x_i^+$  and  $x_{i+1}^+$  there is a factor

$$\frac{1}{P_i^- - \sum_{l \in i} k_{l, \text{os}}^- + i\epsilon}, \quad (33)$$

where  $P_i^-$  is the minus component of the sum of all external momenta entering the graph before  $x_i^+$ . The sum is over the on-shell values of the minus components

$$k_{l, \text{os}}^- = \frac{\mathbf{k}_l^2 + m_l^2}{2k_l^+} \quad (34)$$

of all lines  $l$  between  $x_i^+$  and  $x_{i+1}^+$ .

For particles with spin, one furthermore has to consider that the dependence of the propagator numerators on the particles minus momenta leads to a separation into propagating and instantaneous contributions. Decomposing the covariant four-momentum as

$$k = k_{\text{os}} + (k - k_{\text{os}}), \quad (35)$$

where by definition  $k$  and  $k_{\text{os}}$  only differ in their minus components, one can rewrite the covariant fermion propagator as

$$G_f(k) = \frac{\Theta(k^+)}{2k^+} \frac{i(\not{k}_{\text{os}}) + m}{k^- - \frac{\mathbf{k}^2 + m^2}{2k^+} + i\epsilon} + \frac{\Theta(-k^+)}{-2k^+} \frac{i(\not{k}_{\text{os}}) + m}{-k^- - \frac{\mathbf{k}^2 + m^2}{-2k^+} + i\epsilon} + \frac{i\gamma^+}{2k^+}. \quad (36)$$

The first term describes the propagation of a fermion and the second term the propagation of an antifermion, both with positive plus momentum as stated in point 4 above. The third term is independent of  $k^-$  and hence instantaneous. In LCPT graphs it gives rise to a vertical fermion line whose ends are associated with the same light-cone time  $x^+$ . For the gluon propagator in light-cone gauge, we have

$$G_g^{\mu\nu}(k) = \frac{i}{k^2 + i\epsilon} \left( -g^{\mu\nu} + \frac{n^\mu k_{\text{os}}^\nu + k_{\text{os}}^\nu n^\mu}{k^+} \right) + \frac{in^\mu n^\nu}{(k^+)^2}, \quad (37)$$

where  $n$  is the light-like vector projecting on plus components,  $kn = k^+$ . In analogy with (36), the first term in (37) can be further decomposed into parts with  $\Theta(k^+)$  or  $\Theta(-k^+)$ . The last term in (37) is again instantaneous.

The decomposition (35) has to be made for all Feynman rules in covariant perturbation theory that have a dependence on minus momenta in the numerator, in particular for the three-gluon vertex. We will however only be concerned with gluon lines that are either internal to a graph or associated with the twist-two operator (2) for an observed parton. Any gluon vertex in a graph then has all its Lorentz indices contracted with a gluon propagator. The difference  $(k - k_{\text{os}})^\mu \propto n^\mu$  between a covariant and an on-shell momentum gives zero when contracted with  $G_g^{\mu\nu}$ , so that in the three-gluon vertex we can simply replace  $k$  with  $k_{\text{os}}$  in the numerator factor.

## 5 All-order proof for bare distributions

In this section we derive the DPD sum rules for bare, i.e. unrenormalised, distributions. We consider graphs in LCPT at arbitrary fixed order in  $\alpha_s$ , using perturbation theory in the same spirit as discussed at the beginning of section 3. Having established the sum rules for any fixed order in  $\alpha_s$ , one immediately obtains their validity for the sum over all perturbative orders.

### 5.1 Representation of PDFs and DPDs in LCPT

To begin with, we adapt the representations (6) and (7) of PDFs and DPDs in terms of Green functions to the LCPT formalism. For the bare PDF we find

$$f_B^{j_1}(x_1) = \sum_g (x_1 p^+)^{-n_1} (p^+)^{N(g)-2} \int \frac{dk_1^- d^{D-2}\mathbf{k}_1}{(2\pi)^D} \left( \prod_{i=2}^{N(g)} \frac{dx_i d^{D-2}\mathbf{k}_i}{(2\pi)^{D-1}} \right) \times \mathcal{G}_g^{j_1}(\{x\}, \{\mathbf{k}\}) 2\pi \delta \left( p^- - k_1^- - \sum_{i=2}^{M(g)} k_{i,\text{os}}^- \right) \delta \left( 1 - \sum_{i=1}^{M(g)} x_i \right), \quad (38)$$

where

$$x_i = k_i^+ / p^+. \quad (39)$$

The label  $g$  now specifies a cut LCPT graph, which has a definite light-cone time ordering of vertices. We label the independent momenta in a graph by  $k_i$ , always starting with  $k_1$  for the observed parton. The total number of independent momenta is  $N(g) - 1$ , and  $M(g) - 1$  of these go across the final state cut. We collectively write  $\{x\}$  and  $\{\mathbf{k}\}$  for the set of light-cone momentum fraction and transverse momentum arguments of a graph.

As already discussed in section 3, we can obtain DPDs from the PDF for parton  $j_1$  by inserting the operator for a second parton on one of the final state parton lines. The result is

$$F_B^{j_1 j}(x_1, z) = \sum_g \sum_l \delta_{j,f(l)} (x_1 p^+)^{-n_1} (x_l p^+)^{-n_l} 2 (p^+)^{N(g)-2} \times \int dx_l \delta(x_l - z) \int \frac{dK^- d^{D-2}\mathbf{k}_1 d^{D-2}\mathbf{k}_l}{(2\pi)^{2D-1}} \left( \prod_{\substack{i=2 \\ i \neq l}}^{N(g)} \frac{dx_i d^{D-2}\mathbf{k}_i}{(2\pi)^{D-1}} \right) \times \mathcal{G}_{g,l}^{j_1 f(l)}(\{x\}, \{\mathbf{k}\}) 2\pi \delta \left( p^- - K^- - \sum_{\substack{i=2 \\ i \neq l}}^{M(g)} k_{i,\text{os}}^- \right) \delta \left( 1 - \sum_{i=1}^{M(g)} x_i \right), \quad (40)$$

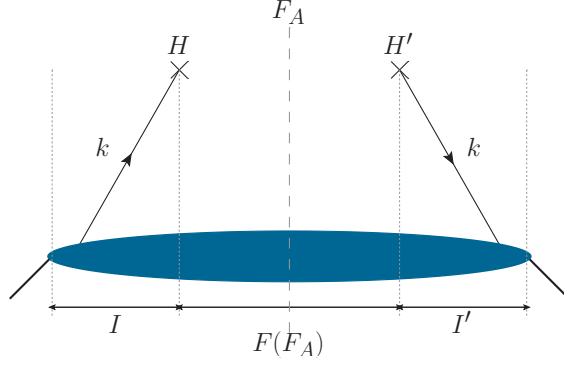


Figure 6: Schematic illustration of an LCPT graph for a PDF.  $H$  ( $H'$ ) denotes the light-cone time of the vertex for the insertion of the twist-two operator on the left (right) of the final-state cut. Since the order of vertices matters for evaluating LCPT graphs, one cannot place that operator insertion on the line indicating the final state cut, as was done for the Feynman graphs in section 3.

where we have a sum over all parton lines  $l$  on which the operator for parton  $j$  can be inserted, with  $\delta_{j,f(l)}$  selecting the parton type and  $\delta(x_l - z)$  the plus momentum fraction. Notice that both the plus momentum and the transverse momentum components of the two observed partons are equal on both sides of the final state cut. The former is always the case for a DPD, whilst the latter holds because for the sum rules we are considering the case  $\mathbf{\Delta} = \mathbf{0}$ .

Starting from (7) we have arrived at (40) by making the variable substitutions

$$K^- = (k_1 + k_l)^- \quad k^- = (k_l - \Delta)^- / 2, \quad k'^- = (k_l + \Delta)^- / 2 \quad (41)$$

and performing the integrations over  $k^-$  and  $k'^-$ . As described in Appendix B of [27], the result of these integrations is that on each side of the final state cut the two vertices corresponding to the operator insertions for the observed partons  $j_1$  and  $j$  are associated with the *same* light-cone time  $x^+$ . This is not surprising, because the two operators  $\mathcal{O}_{j_1}$  and  $\mathcal{O}_{j_2}$  in the definition (1) of a DPD are taken at the same light-cone time.

We now perform the integration over  $k_1^-$  in the expression (38) of the PDF. The following is a simplified version of the argument in chapter 14.4.3 of [26]. A general LCPT graph for the PDF, illustrated in figure 6, can be written in the form

$$\mathcal{G}_g^{j_1} = I F(F_A) I' \times \{\text{numerator}\} \quad (42)$$

with

$$I = \prod_{\substack{\text{states } \xi \\ \xi < H}} \frac{1}{p^- - \sum_{l \in \xi} k_{l, \text{os}}^- + i\epsilon}, \quad I' = \prod_{\substack{\text{states } \xi \\ \xi < H'}} \frac{1}{p^- - \sum_{l \in \xi} k_{l, \text{os}}^- - i\epsilon} \quad (43)$$

and

$$F(F_A) = \prod_{\substack{\text{states } \xi \\ H < \xi < F_A}} \frac{1}{p^- - k_1^- - \sum_{l \in \xi} k_{l, \text{os}}^- + i\epsilon} 2\pi\delta \left( p^- - k_1^- - \sum_{l \in F_A} k_{l, \text{os}}^- \right) \\ \times \prod_{\substack{\text{states } \xi \\ H' < \xi < F_A}} \frac{1}{p^- - k_1^- - \sum_{l \in \xi} k_{l, \text{os}}^- - i\epsilon}. \quad (44)$$

The ‘‘numerator’’ in (42) includes vertex factors, propagator numerators and factors of  $\pm i$ , and its detailed form is not relevant for the present argument. The graph is cut across the final state  $F_A$ , and the sums in  $I, I'$  and  $F$  are over intermediate states  $\xi$  either before or after the light-cone time of the operator insertion. Let us now take the sum over all graphs  $g$  that differ only by the state  $F_A$  where the cut is made but are otherwise identical. Numbering the states in  $F(F_A)$  from 1 to  $N$ , we can write

$$\sum_{F_A} F(F_A) = \sum_{c=1}^N \left[ \prod_{f=1}^{c-1} \frac{1}{p^- - k_1^- - D_f + i\epsilon} 2\pi\delta(p^- - k_1^- - D_c) \prod_{f=c+1}^N \frac{1}{p^- - k_1^- - D_f - i\epsilon} \right], \quad (45)$$

where we abbreviated  $D_f = \sum_{l \in f} k_{l,os}^-$ . Rewriting the  $\delta$  function in this expression as

$$2\pi\delta(x) = i \left[ \frac{1}{x + i\epsilon} - \frac{1}{x - i\epsilon} \right], \quad (46)$$

we find

$$\sum_{F_A} F(F_A) = i \left[ \prod_{f=1}^N \frac{1}{p^- - k_1^- - D_f + i\epsilon} - \prod_{f=1}^N \frac{1}{p^- - k_1^- - D_f - i\epsilon} \right]. \quad (47)$$

Using the theorem of residues, one can easily see that this expression vanishes for  $N \geq 2$  when one integrates over  $k_1^-$ . For  $N = 1$  on the other hand, the initial  $\delta$  function in (40) is reproduced. Thus one can conclude that for a PDF only needs to consider those  $x^+$  orderings of the vertices for which there are no states ‘‘later’’, i.e. with greater  $x^+$ , than the operator insertion vertices  $H$  and  $H'$  on each side of the final state cut.

The preceding argument can be repeated for the  $K^-$  integration in the expression (40) for a DPD, and one finds again that only time orderings with no intermediate state after the operator insertions contribute. As a result, we have

$$f_B^{j_1}(x_1) = \sum_g (x_1 p^+)^{-n_1} (p^+)^{N(g)-2} \int \frac{d^{D-2}\mathbf{k}_1}{(2\pi)^{D-1}} \times \left( \prod_{i=2}^{N(g)} \frac{dx_i d^{D-2}\mathbf{k}_i}{(2\pi)^{D-1}} \right) \mathcal{G}_g^{j_1}(\{x\}, \{\mathbf{k}\}) \delta \left( 1 - \sum_{i=1}^{M(g)} x_i \right), \quad (48)$$

$$\int_0^{1-x_1} dz z^m F_B^{j_1 j}(x_1, z) = \sum_g \sum_l \delta_{j,f(l)} (x_1 p^+)^{-n_1} (p^+)^{N(g)-2} \int \frac{d^{D-2}\mathbf{k}_1}{(2\pi)^{D-1}} \times \left( \prod_{i=2}^{N(g)} \frac{dx_i d^{D-2}\mathbf{k}_i}{(2\pi)^{D-1}} \right) 2x_l^m (x_l p^+)^{-n_l} \mathcal{G}_{g,l}^{j_1 f(l)}(\{x\}, \{\mathbf{k}\}) \delta \left( 1 - \sum_{i=1}^{M(g)} x_i \right), \quad (49)$$

where the sum over all graphs  $g$  can be restricted to the time orderings just discussed. We integrated the DPD over its second momentum fraction  $z$  with weight  $z^m$ , as is required for the sum rules (where  $m = 0$  or  $m = 1$ ).

## 5.2 Equality between PDF and DPD graphs

The next step in the proof is to establish the equality

$$2(x_l p^+)^{-n_l} \mathcal{G}_{g,l}^{j_1 f(l)} \stackrel{!}{=} \mathcal{G}_g^{j_1} \quad (50)$$

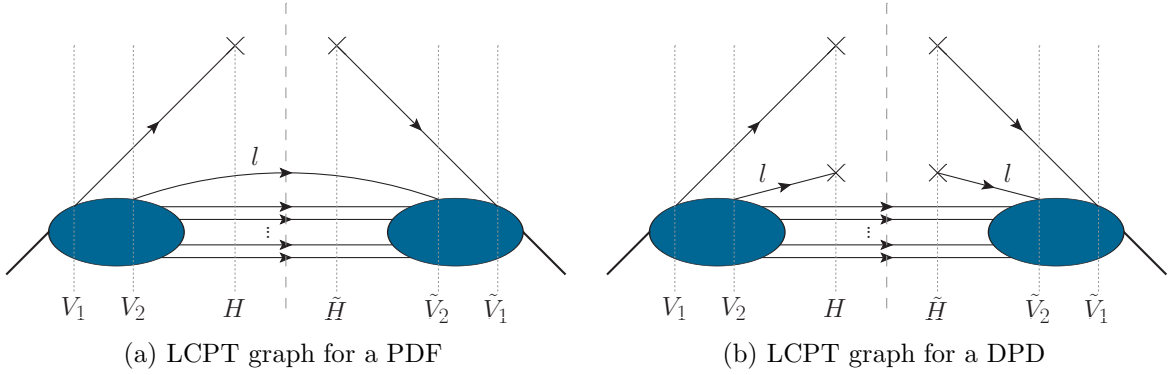


Figure 7: LCPT graphs for a PDF or a DPD. It is understood that the subgraphs denoted by blobs are identical in panels (a) and (b). One may interchange the time ordering between the vertices  $V_1$  and  $V_2$ , and independently the time ordering between the vertices  $\tilde{V}_1$  and  $\tilde{V}_2$ .

for all graphs  $g$  and all partons  $l$  that contribute in (48) and (49). This includes the statement that there is a unique correspondence between the LCPT graphs  $g$  that contribute to  $f^{j_1}$  and those that contribute to  $F^{j_1 j}$ , as indicated in figure 7.

We already showed in the last section that in a PDF graph the vertex  $H$  of the twist-two operator insertion must be later in light-cone time than the vertex  $V_2$  where the final state line  $l$  leaves the graph. Let us now show that there can be no instantaneous propagator attached to a twist-two operator insertion. The vertex  $V_1$  where parton 1 leaves the graph must then come before  $H$  as well, and in a DPD graph both  $V_1$  and  $V_2$  must come before  $H$ . In both PDF and DPD graphs,  $V_1$  may come before or after  $V_2$ . Corresponding statements hold for  $\tilde{V}_1$ ,  $\tilde{V}_2$  and  $\tilde{H}$  to the right of the cut.

To show this, we identify the vertex rules (in momentum space) associated with the unpolarised twist-two operators (2). For quarks we get a factor  $\gamma^+/2$  that connects the Dirac indices of the parton to the left and the right of the cut, and for antiquark we get a factor  $-\gamma^+/2$ . This gives zero when multiplied with the instantaneous part of the fermion propagator (36) because  $(\gamma^+)^2 = 0$ . In light-cone gauge  $A^+ = 0$ , the twist-two operator for gluons gives a factor  $(k^+)^2 \delta^{ii'}$ . Here  $k^+$  denotes the gluon plus momentum, which is equal on both sides of the cut. The index  $i$  ( $i'$ ) denotes the gluon polarisation index to the left (right) of the cut and is restricted to be transverse. This gives zero when contracted with the instantaneous part of the gluon propagator (37).

We thus find that LCPT graphs contributing to a PDF are related to those contributing to a DPD by inserting the operator for the second parton on a final state line in the PDF graph. This is analogous to the statement we used for Feynman graphs in section 3 but now includes the statement about the relative time orderings between vertices. With PDF and DPD graphs having the same time ordering of vertices, they have identical light-cone energy denominators (33). We now compare the numerator factor associated with the line  $l$  selected by the operator for parton 2 in the DPD and the factor associated with the corresponding final state line in the PDF. In the DPD, the momenta  $k$  and  $k'$  carried by  $l$  to the left and the right of the cut have the same plus and transverse components, as already noted after (40). Their on-shell values are hence identical as well, i.e.

$$k_{\text{os}} = k'_{\text{os}}. \quad (51)$$

Furthermore,  $k$  is equal to the momentum of the corresponding final state line in the PDF. If  $l$  is a quark, we get

$$2 \frac{k_{\text{os}} + m}{2k^+} \frac{\gamma^+}{2} \frac{k_{\text{os}} + m}{2k^+} = \frac{k_{\text{os}} + m}{2k^+} \quad (52)$$



in the DPD graph. Here a factor  $(k_{\text{os}}^+ + m)/(2k^+)$  appears for each propagating line (see (36)), the factor  $\gamma^+/2$  comes from the twist-two quark operator, and the factor 2 on the l.h.s. is taken from the l.h.s. of (50). On the r.h.s. of (52) we recognise the factor for the final state line  $l$  in the PDF graph, which proves (50) for quarks. The same argument is easily repeated for antiquarks. If  $l$  is a gluon, we get

$$\begin{aligned} & \frac{2}{k^+} \frac{1}{2k^+} \left( -g^{\mu i} + \frac{n^\mu k_{\text{os}}^i + n^i k_{\text{os}}^\mu}{k^+} \right) (k^+)^2 \delta^{i i'} \frac{1}{2k^+} \left( -g^{\mu' i'} + \frac{n^{\mu'} k_{\text{os}}^{i'} + n^{i'} k_{\text{os}}^{\mu'}}{k^+} \right) \\ &= \frac{1}{2k^+} \left( -g^{\mu \mu'} + \frac{n^\mu k_{\text{os}}^{\mu'} + n^{\mu'} k_{\text{os}}^\mu}{k^+} \right) \end{aligned} \quad (53)$$

in the DPD graph, where the factor  $2/k^+ = 2(x_l p^+)^{-1}$  on the l.h.s. is again taken from the l.h.s. of (50). The r.h.s. of (53) is the factor for a final state gluon in the PDF graph, which completes the proof of (50).

### 5.3 Number sum rule

We can now insert the relation (50) into the expression (49) for the integral of the DPD over its second momentum argument and set the power  $m = 0$ . The validity of the number sum rule for a bare DPD then requires that

$$\begin{aligned} & \sum_g \sum_l (\delta_{j_2, f(l)} - \delta_{\overline{j_2}, f(l)}) (x_1 p^+)^{-n_1} (p^+)^{N(g)-2} \int \frac{d^{D-2} \mathbf{k}_1}{(2\pi)^{D-1}} \left( \prod_{i=2}^{N(g)} \frac{dx_i d^{D-2} \mathbf{k}_i}{(2\pi)^{D-1}} \right) \\ & \quad \times \mathcal{G}_g^{j_1}(\{x\}, \{\mathbf{k}\}) \delta \left( 1 - \sum_{i=1}^{M(g)} x_i \right) \\ & \stackrel{!}{=} (N_{j_2, v} + \delta_{j_1, \overline{j_2}} - \delta_{j_1, j_2}) \sum_g (x_1 p^+)^{-n_1} (p^+)^{N(g)-2} \int \frac{d^{D-2} \mathbf{k}_1}{(2\pi)^{D-1}} \left( \prod_{i=2}^{N(g)} \frac{dx_i d^{D-2} \mathbf{k}_i}{(2\pi)^{D-1}} \right) \\ & \quad \times \mathcal{G}_g^{j_1}(\{x\}, \{\mathbf{k}\}) \delta \left( 1 - \sum_{i=1}^{M(g)} x_i \right). \end{aligned} \quad (54)$$

It thus remains to show that

$$\delta_{j_1, j_2} - \delta_{j_1, \overline{j_2}} + \sum_l (\delta_{j_2, f(l)} - \delta_{\overline{j_2}, f(l)}) \stackrel{!}{=} N_{j_2, v}, \quad (55)$$

where  $j_2$  denotes either a quark or an antiquark. The sum over  $l$  on the l.h.s. gives the number of partons with flavour  $j_2$  crossing the final state cut in the PDF graph, minus the corresponding number of partons with flavour  $\overline{j_2}$ . If the observed parton  $j_1$  in the PDF has flavour  $j_2$  ( $\overline{j_2}$ ), that number is increased (decreased) by 1. The result is obviously equal to the difference of partons with flavour  $j_2$  and those with flavour  $\overline{j_2}$  in the hadron, which is indeed  $N_{j_2, v}$ . The number sum rule is thus verified.

## 5.4 Momentum sum rule

Inserting the relation (50) into (49) with  $m = 1$ , we find that the momentum sum rule holds if

$$\begin{aligned}
& \sum_g \sum_l (x_1 p^+)^{-n_1} (p^+)^{N(g)-2} \int \frac{d^{D-2} \mathbf{k}_1}{(2\pi)^{D-1}} \left( \prod_{i=2}^{N(g)} \frac{dx_i d^{D-2} \mathbf{k}_i}{(2\pi)^{D-1}} \right) \\
& \quad \times x_l \mathcal{G}_g^{j_1}(\{x\}, \{\mathbf{k}\}) \delta \left( 1 - \sum_{i=1}^{M(g)} x_i \right) \\
& \stackrel{!}{=} (1 - x_1) \sum_g (x_1 p^+)^{-n_1} (p^+)^{N(g)-2} \int \frac{d^{D-2} \mathbf{k}_1}{(2\pi)^{D-1}} \left( \prod_{i=2}^{N(g)} \frac{dx_i d^{D-2} \mathbf{k}_i}{(2\pi)^{D-1}} \right) \\
& \quad \times \mathcal{G}_g^{j_1}(\{x\}, \{\mathbf{k}\}) \delta \left( 1 - \sum_{i=1}^{M(g)} x_i \right). \tag{56}
\end{aligned}$$

This just means that the sum over the momentum fractions  $x_l$  of all partons crossing the final state cut in a PDF graph must be equal to  $1 - x_1$ , which is a direct consequence of momentum conservation. This concludes our proof of the sum rules for bare DPDs.

## 6 Renormalisation

Up to now, we have essentially shown that the parton model interpretation of the DPD sum rules is reflected in the graphs that represent single or double parton distributions in QCD, where quark number and parton momentum are conserved quantities. However, these graphs have short-distance singularities that must be renormalised. It is known that the literal interpretation of PDFs as probability densities can be invalidated by renormalisation. This is most obvious for the positivity of the distributions, because one has to *subtract* terms that become infinite if the UV regulator is removed. It is hence important to establish whether the sum rules retain their validity after renormalisation in a specified scheme. We will show that this is indeed the case for the  $\overline{\text{MS}}$  scheme. As in the previous section, our arguments are valid at arbitrary order in  $\alpha_s$ .

### 6.1 Convolution integrals

The calculations in the following sections make heavy use of multiple convolution integrals involving functions of one or two momentum fractions. To keep expressions readable, we introduce a shorthand notation that avoids giving the explicit arguments of the integrands, and we derive some useful rules of computation. In the following let  $D$  be a function of two momentum fractions, while  $A, B, C$  are functions of one momentum fraction only. We define

$$A \otimes_1 D = \int \frac{dz}{z} A\left(\frac{x_1}{z}\right) D(z, x_2), \tag{57}$$

where the integration boundaries are determined by the support properties of the functions under the integral. Specifically, when a one-variable function, say  $A$ , is a PDF we have  $A(x) = 0$  if  $x < 0$  or  $x > 1$ , and when a two-variable function is a DPD we have  $D(x_1, x_2) = 0$  if  $x_1 < 0$  or  $x_2 < 0$  or  $x_1 + x_2 > 1$ . The same properties hold also for the associated evolution kernels and renormalisation factors (which may include delta and plus distributions at the

endpoints of their support). The convolution with respect to the second argument of  $D$  is introduced in analogy to (57) and denoted by  $\otimes_2$ . We then have a combined convolution

$$A \otimes_1 B \otimes_2 D = A \otimes_1 \left[ B \otimes_2 D \right]. \quad (58)$$

A second type of convolution integral we will encounter is

$$D \otimes_{12} A = \int \frac{dz}{z^2} D \left( \frac{x_1}{z}, \frac{x_2}{z} \right) A(z). \quad (59)$$

Interchanging the order of integrations, it is straightforward to show that

$$\begin{aligned} \left[ A \otimes_1 D \right] \otimes_{12} B &= A \otimes_1 \left[ D \otimes_{12} B \right], \\ \left[ A \otimes_1 B \otimes_2 D \right] \otimes_{12} C &= A \otimes_1 B \otimes_2 \left[ D \otimes_{12} C \right], \end{aligned} \quad (60)$$

so that we can write these convolutions without the brackets. Using the support properties of the functions stated above and making the integration boundaries explicit, one also has

$$\begin{aligned} \left[ D \otimes_{12} A \right] \otimes_{12} B &= \int_{x_1+x_2}^1 \frac{dy}{y^2} \int_{(x_1+x_2)/y}^1 \frac{dz}{z^2} D \left( \frac{x_1}{yz}, \frac{x_2}{yz} \right) A(z) B(y) \\ &= \int_{x_1+x_2}^1 \frac{dw}{w^2} D \left( \frac{x_1}{yz}, \frac{x_2}{yz} \right) \int_w^1 \frac{dy}{y} A \left( \frac{w}{y} \right) B(y) \\ &= D \otimes_{12} [A \otimes B], \end{aligned} \quad (61)$$

where  $\otimes$  denotes the usual Mellin convolution for functions of a single momentum fraction.

As a shorthand for integrals over momentum fractions, we write

$$\int A = \int dx A(x), \quad \int_2 D = \int dx_2 D(x_1, x_2) \quad (62)$$

for functions of one or two momentum fractions, respectively. We further introduce operators  $X^n$  and  $X_2^n$  that act on a function by multiplying with a power of the appropriate momentum fraction,

$$(X^n A)(x) = x^n A(x), \quad (X_2^n D)(x_1, x_2) = x_2^n D(x_1, x_2), \quad (63)$$

such that in convolution integrals  $X^n A$  and  $X_2^n D$  can be used without explicitly giving their momentum arguments. For functions of one argument, one has the well-known rules

$$X^n (A \otimes B) = (X^n A) \otimes (X^n B), \quad \int X^n (A \otimes B) = \int X^n A \int X^n B, \quad (64)$$

and it is easy to see that for functions of two momentum fractions

$$\int_2 X_2 (A \otimes_2 D) = \int X A \int_2 X_2 D. \quad (65)$$

We will also make use of the following identity:

$$\begin{aligned}
\int_2 X_2^n (D \otimes_{12} A) &= \int dx_2 x_2^n \int \frac{dz}{z^2} D\left(\frac{x_1}{z}, \frac{x_2}{z}\right) A(z) \\
&= \int \frac{du_1}{u_1} \int du_2 u_2^n D(u_1, u_2) \left(\frac{x_1}{u_1}\right)^n A\left(\frac{x_1}{u_1}\right) \\
&= \left(\int_2 X_2^n D\right) \otimes_1 (X^n A), \tag{66}
\end{aligned}$$

where the subscript on the convolution symbol in the last line indicates that the result depends on the momentum fraction  $x_1$ .

## 6.2 Renormalisation of DPDs

We will now analyse how the DPDs in the sum rules, i.e. momentum space DPDs evaluated at  $\Delta = 0$ , have to be renormalised. Let us first consider position space DPDs  $F_B(y)$ . These have short-distance singularities for each of the twist-two operators in their definition (1). For a single-parton distribution these singularities are renormalised with factors  $Z_{i,j}(x; \mu, \epsilon)$  as follows:

$$f^i(\mu) = \sum_j Z_{i,j}(\mu) \otimes f_B^j. \tag{67}$$

Correspondingly a bare DPD is renormalised with a  $Z$  factor for each parton, i.e.

$$F^{i_1 i_2}(y; \mu) = \sum_{j_1, j_2} Z_{i_1, j_1}(\mu) \otimes_1 Z_{i_2, j_2}(\mu) \otimes_2 F_B^{j_1 j_2}(y). \tag{68}$$

The Fourier transform (4) from position to momentum space involves an integration over all  $y$  and produces an additional short-distance singularity, which is due to the  $1 \rightarrow 2$  splitting mechanism. As discussed in [36, 13], this mechanism dominates the DPD at small  $y$  and gives

$$F_B^{i_1 i_2}(y; \mu) \Big|_{y \rightarrow 0} = \frac{\Gamma(1 - \epsilon)}{(\pi y^2)^{1 - \epsilon}} \sum_j V_B^{i_1 i_2, j}(y; \mu) \otimes_{12} f_B^j \tag{69}$$

with perturbative coefficient functions  $V_B(x_1, x_2, y; \mu)$  that depend on  $y$  via powers of  $(y\mu)^{2\epsilon} \alpha_s(\mu)$ . The Fourier transform of (69) w.r.t. the transverse distance  $y$  has a logarithmic singularity in  $D = 4$  dimensions and gives a simple pole  $1/\epsilon$  for  $D = 4 - 2\epsilon$ . Notice the difference between this and the UV divergences in the twist-two operators, which lead to higher powers of  $1/\epsilon$  with increasing powers of  $\alpha_s$ . The splitting singularity in  $F_B(\Delta)$  is renormalised additively with a renormalisation factor  $Z_{i_1 i_2, j}(x_1, x_2; \mu, \epsilon)$  depending on two momentum fractions:

$$F^{i_1 i_2}(\Delta; \mu) = \sum_{j_1, j_2} Z_{i_1, j_1}(\mu) \otimes_1 Z_{i_2, j_2}(\mu) \otimes_2 F_B^{j_1 j_2}(\Delta) + \sum_j Z_{i_1 i_2, j}(\mu) \otimes_{12} f_B^j. \tag{70}$$

Introducing single and double Mellin moments by the integrals

$$A(m) = \int dx x^{m-1} A(x), \quad D(m_1, m_2) = \int dx_1 x_1^{m_1-1} \int dx_2 x_2^{m_2-1} D(x_1, x_2), \tag{71}$$

one easily sees that (70) turns into

$$F^{i_1 i_2}(m_1, m_2, \Delta; \mu) = \sum_{j_1, j_2} Z_{i_1, j_1}(m_1; \mu) Z_{i_2, j_2}(m_2; \mu) F_B^{i_1 i_2}(m_1, m_2, \Delta) + \sum_j Z_{i_1 i_2, j}(m_1, m_2; \mu) f_B^j(m_1 + m_2 - 1), \quad (72)$$

in agreement with the leading-order analyses in [17, 18].

The renormalisation factor for PDFs has a perturbative expansion

$$Z_{i,j}(x; \mu, \epsilon) = \delta_{i,j} \delta(1-x) + \sum_{n=1}^{\infty} \alpha_s^n(\mu) Z_{i,j}^{(n)}(x; \epsilon), \quad (73)$$

whereas its analogue for the splitting singularity has no tree-level term:

$$Z_{i_1 i_2, j}(x_1, x_2; \mu, \epsilon) = \sum_{n=1}^{\infty} \alpha_s^n(\mu) Z_{i_1 i_2, j}^{(n)}(x_1, x_2; \epsilon). \quad (74)$$

For later use, we define the inverse  $Z^{-1}$  of the PDF renormalisation factor by

$$\sum_j [Z_{i,j} \otimes Z_{j,k}^{-1}](x) = \delta_{i,k} \delta(1-x). \quad (75)$$

The expansion (73) implies a corresponding expansion

$$Z_{i,j}^{-1}(x; \mu, \epsilon) = \delta_{i,j} \delta(1-x) + \sum_{n=1}^{\infty} \alpha_s^n(\mu) Z_{i,j}^{-1(n)}(x; \epsilon), \quad (76)$$

whose coefficients  $Z_{i,j}^{-1(n)}$  can easily be expressed in terms of  $Z_{i,j}^{(n)}$  by solving (75) order by order in  $\alpha_s$ . At first order, one simply has  $Z_{i,j}^{-1(1)} = -Z_{i,j}^{(1)}$ .

**Implementation of the  $\overline{\text{MS}}$  scheme.** The derivations in the remainder of this paper will be significantly simplified by using a particular implementation of the  $\overline{\text{MS}}$  renormalisation scheme. We start with the definition of this scheme given in section 3.2.6 of [26], where the bare and the renormalised couplings are related by

$$\alpha_0 = \alpha_s \mu^{2\epsilon} \left[ 1 + \sum_{n=1}^{\infty} \alpha_s^n S_\epsilon^n \sum_{m=1}^n \frac{B_{nm}}{\epsilon^m} \right], \quad (77)$$

and a generic renormalisation factor reads

$$Z = Z^{(0)} + \sum_{n=1}^{\infty} \alpha_s^n S_\epsilon^n \sum_{m=1}^{M(n)} \frac{Z_{nm}}{\epsilon^m}, \quad (78)$$

where the order  $M(n)$  of the highest pole depends on the quantity being renormalised. The tree-level value  $Z^{(0)}$  is not important here. The coefficients  $B_{nm}$  and  $Z_{nm}$  are independent of  $\epsilon$ , but  $Z$  and thus  $Z^{(0)}$  and  $Z_{nm}$  may depend on additional variables like momentum fractions. The standard choice for the factor  $S_\epsilon$  is  $S_\epsilon = (4\pi e^{-\gamma})^\epsilon$ , where  $\gamma$  is the Euler–Mascheroni constant. The alternative  $S_\epsilon = (4\pi)^\epsilon / \Gamma(1-\epsilon)$  was proposed in [26], and the following arguments are valid in both cases. The counterterms in (77) and (78) contain finite parts that result from multiplying powers of  $1/\epsilon$  with powers of  $S_\epsilon$ .

We now define a second renormalisation scheme by

$$\alpha_0 = \alpha'_s \frac{\mu^{2\varepsilon}}{S_\varepsilon} \left[ 1 + \sum_{n=1}^{\infty} \alpha_s'^n \sum_{m=1}^n \frac{B'_{nm}}{\varepsilon^m} \right], \quad (79)$$

and

$$Z = Z^{(0)} + \sum_{n=1}^{\infty} \alpha_s'^n \sum_{m=1}^{M(n)} \frac{Z'_{nm}}{\varepsilon^m}, \quad (80)$$

where  $B'_{nm}$  and  $Z'_{nm}$  are again independent of  $\varepsilon$ . The counterterms in this scheme are pure poles in  $\varepsilon$ . This will be essential for the arguments in the following sections.

Let us show that the two schemes defined by (77), (78) and by (79), (80) give the same renormalised quantities at  $\varepsilon = 0$ . To this end we introduce a rescaled strong coupling

$$\bar{\alpha}_s(\varepsilon) = \alpha'_s / S_\varepsilon, \quad (81)$$

so that (79) and (80) become

$$\alpha_0 = \bar{\alpha}_s(\varepsilon) \mu^{2\varepsilon} \left[ 1 + \sum_{n=1}^{\infty} \bar{\alpha}_s^n(\varepsilon) S_\varepsilon^n \sum_{m=1}^n \frac{B'_{nm}}{\varepsilon^m} \right], \quad (82)$$

and

$$Z = Z^{(0)} + \sum_{n=1}^{\infty} \bar{\alpha}_s^n(\varepsilon) S_\varepsilon^n \sum_{m=1}^{M(n)} \frac{Z'_{nm}}{\varepsilon^m}. \quad (83)$$

Consider now a renormalised quantity  $R(\alpha_s, \varepsilon, B_{nm}, Z_{nm})$  in the first scheme and its counterpart  $R'(\bar{\alpha}_s(\varepsilon), \varepsilon, B'_{nm}, Z'_{nm})$  in the second scheme. Because (77) and (78) have the same functional form as (82) and (83), the renormalised quantities in both schemes also have the same functional form, i.e.

$$R'(\bar{\alpha}_s, \varepsilon, B'_{nm}, Z'_{nm}) = R(\bar{\alpha}_s, \varepsilon, B'_{nm}, Z'_{nm}). \quad (84)$$

The coefficients  $B_{nm}$  and  $Z_{nm}$  ( $B'_{nm}$  and  $Z'_{nm}$ ) are uniquely fixed by the requirement that renormalised quantities have no poles in  $\varepsilon$  when expanded in  $\alpha_s$  and  $\varepsilon$  ( $\alpha'_s$  and  $\varepsilon$ ). Since  $\bar{\alpha}_s(\varepsilon)$  only differs from  $\alpha'_s$  by terms of order  $\varepsilon$ , one must also obtain an expression without poles in  $\varepsilon$  when expanding  $R'$  in  $\bar{\alpha}_s$  and  $\varepsilon$ . As a consequence, the renormalisation coefficients in the two schemes are identical,  $B_{nm} = B'_{nm}$  and  $Z_{nm} = Z'_{nm}$ , so that (84) implies

$$R'(\bar{\alpha}_s, \varepsilon, B'_{nm}, Z'_{nm}) = R(\bar{\alpha}_s, \varepsilon, B_{nm}, Z_{nm}). \quad (85)$$

With  $\bar{\alpha}_s(0) = \alpha'_s$  we thus find that in the primed scheme the value of  $R'$  at the physical point is

$$\lim_{\varepsilon \rightarrow 0} R'(\bar{\alpha}_s, \varepsilon, B'_{nm}, Z'_{nm}) = R(\alpha'_s, 0, B_{nm}, Z_{nm}). \quad (86)$$

In the original scheme, the value of  $R$  is

$$\lim_{\varepsilon \rightarrow 0} R(\alpha_s, \varepsilon, B_{nm}, Z_{nm}) = R(\alpha_s, 0, B_{nm}, Z_{nm}) \quad (87)$$

at the physical point. Consider now a case in which the renormalised quantity is an observable (e.g. the hadronic vacuum polarisation). Then it must have the same value in the two schemes,

and we can conclude that  $\alpha'_s = \alpha_s$ . It follows that for any other quantity, including quantities that are not observables (such as renormalised PDFs or DPDs), the two schemes give the same result at the physical point  $\varepsilon = 0$ .

For the standard choice  $S_\varepsilon = (4\pi e^{-\gamma})^\varepsilon$ , the relation (79) takes the form of a minimal subtraction scheme with  $\bar{\mu}^2 = \mu^2/(S_\varepsilon)^{1/\varepsilon} = \mu^2 e^\gamma/(4\pi)$ . This way of implementing  $\overline{\text{MS}}$  subtraction is in fact well known in the literature. Our above argument shows that one can also use the implementation of (79) and (80) for different choices of  $S_\varepsilon$ . In the remainder of this work, we will use this implementation, omitting the primes on  $\alpha_s$ ,  $B_{nm}$  and  $Z_{nm}$ .

### 6.3 Number sum rule

We are now ready to prove the number sum rule for renormalised DPDs. Using the notation introduced in section 6.1, we need to show that the difference

$$\Delta^{i_1 i_2} = \int_2 F^{i_1 i_2, v} - (N_{i_2, v} + \delta_{i_1, \bar{i}_2} - \delta_{i_1, i_2}) f^{i_1} \quad (88)$$

is zero. Using the expression (70) for the renormalised DPD and the rules of computation from section 6.1, we get

$$\begin{aligned} \int_2 F^{i_1 i_2, v} &= \sum_{j_1, j_2} Z_{i_1, j_1} \otimes_1 \int_2 (Z_{i_2, v, j_2} \otimes_2 F_B^{j_1 j_2}) + \sum_j \int_2 (Z_{i_1 i_2, v, j} \otimes_{12} f_B^j) \\ &= \sum_{j_1, j_2} Z_{i_1, j_1} \otimes_1 \int_2 Z_{i_2, v, j_2} \int_2 F_B^{j_1 j_2} + \sum_j \left( \int_2 Z_{i_1 i_2, v, j} \right) \otimes_1 f_B^j. \end{aligned} \quad (89)$$

The number sum rule for renormalised PDFs implies a sum rule for their renormalisation factors, which reads

$$\int Z_{i, v, j} = \int (Z_{i, j} - Z_{\bar{i}, j}) = \delta_{i, j} - \delta_{i, \bar{j}}. \quad (90)$$

A proof can be found in section 8.6 of [26]. We recall that in the case where  $i$  is a gluon we define  $\bar{i} = i$ , so that the above relation is valid for all parton labels. For the first term in (89) we thus have

$$\begin{aligned} \sum_{j_1, j_2} Z_{i_1, j_1} \otimes_1 \int_2 Z_{i_2, v, j_2} \int_2 F_B^{j_1 j_2} &= \sum_{j_1} Z_{i_1, j_1} \otimes_1 \int_2 F_B^{j_1 i_2, v} \\ &= \sum_{j_1} Z_{i_1, j_1} \otimes_1 (N_{i_2, v} + \delta_{j_1, \bar{i}_2} - \delta_{j_1, i_2}) f_B^{j_1} \\ &= N_{i_2, v} f^{i_1} + \sum_j (\delta_{j, \bar{i}_2} - \delta_{j, i_2}) Z_{i_1, j} \otimes_1 f_B^j. \end{aligned} \quad (91)$$

Here we have used the number sum rule for unrenormalised DPDs, as well as the relation between bare and renormalised PDF in the term proportional to  $N_{i_2, v}$ . Putting all terms together, we obtain

$$\begin{aligned} \Delta^{i_1 i_2} &= \sum_j \left( \int_2 Z_{i_1 i_2, v, j} \right) \otimes_1 f_B^j + \sum_j (\delta_{j, \bar{i}_2} - \delta_{j, i_2}) Z_{i_1, j} \otimes_1 f_B^j - (\delta_{i_1, \bar{i}_2} - \delta_{i_1, i_2}) f^{i_1} \\ &= \sum_k R_k^{i_1 i_2} \otimes_1 f^k \end{aligned} \quad (92)$$

with

$$R_k^{i_1 i_2} = \sum_j \left( \int_2 Z_{i_1 i_2, v, j} \right) \otimes_1 Z_{j, k}^{-1} - \sum_j (\delta_{i_1, \bar{i}_2} - \delta_{i_1, i_2} - \delta_{j, \bar{i}_2} + \delta_{j, i_2}) Z_{i_1, j} \otimes_1 Z_{j, k}^{-1}. \quad (93)$$

In the last step of (92) we expressed the bare PDF in terms of the renormalised one, using the inverse renormalisation factor introduced earlier. In the  $\overline{\text{MS}}$  scheme, the perturbative expansion coefficients  $Z_{i_1 i_2, j}^{(n)}$  and  $Z_{i, j}^{(n)}$  involve only pole terms in  $\epsilon$ , and it is easy to see from (75) that the same is true for  $Z_{i, j}^{-1(n)}$ . The tree-level part of  $Z_{i_1, j}$  in (93) is proportional to  $\delta_{i_1, j}$  and hence vanishes when multiplied with the combination of Kronecker symbols in parentheses.  $R_k^{i_1 i_2}$  is hence a sum of pure pole terms in  $\epsilon$ . Since  $\Delta^{i_1 i_2}$  is finite at  $\epsilon = 0$  according to its definition (88), we can conclude that  $R_k^{i_1 i_2} = 0$  and hence  $\Delta^{i_1 i_2} = 0$ . This proves the number sum rule for renormalised DPDs.

The previous argument implies that  $\sum_k R_k^{i_1 i_2} \otimes Z_{k, j} = 0$ , which together with (93) yields a number sum rule

$$\int_2 Z_{i_1 i_2, v, j} = (\delta_{i_1, \bar{i}_2} - \delta_{i_1, i_2} - \delta_{j, \bar{i}_2} + \delta_{j, i_2}) Z_{i_1, j} \quad (94)$$

for the renormalisation factor of the splitting singularity.

#### 6.4 Momentum sum rule

The proof of the momentum sum rule for renormalised DPDs proceeds in close analogy to the previous subsection. We need to show that

$$\Delta^{i_1} = \sum_{i_2} \int_2 X_2 F^{i_1 i_2} - (1 - X_1) f^{i_1} \quad (95)$$

is zero. The first term of this expression can be rewritten as

$$\begin{aligned} \sum_{i_2} \int_2 X_2 F^{i_1 i_2} &= \sum_{i_2, j_1, j_2} Z_{i_1, j_1} \otimes_1 \int_2 X_2 (Z_{i_2, j_2} \otimes_2 F_B^{j_1 j_2}) + \sum_{i_2, j} \int_2 X_2 (Z_{i_1 i_2, j} \otimes_{12} f_B^j) \\ &= \sum_{i_2, j_1, j_2} Z_{i_1, j_1} \otimes_1 \int X Z_{i_2, j_2} \int_2 X_2 F_B^{j_1 j_2} \\ &\quad + \sum_{i_2, j} \left( \int_2 X_2 Z_{i_1 i_2, j} \right) \otimes_1 (X f_B^j). \end{aligned} \quad (96)$$

The momentum sum rule for renormalised single-parton distributions implies that

$$\sum_i \int X Z_{i, j} = 1 \quad (97)$$

for any  $j$ , as shown in section 8.6 of [26]. Using this and the momentum sum rule for bare DPDs, we have

$$\begin{aligned} \sum_{i_2, j_1, j_2} Z_{i_1, j_1} \otimes_1 \int X Z_{i_2, j_2} \int_2 X_2 F_B^{j_1 j_2} &= \sum_{j_1, j_2} Z_{i_1, j_1} \otimes_1 \int_2 X_2 F_B^{j_1 j_2} \\ &= \sum_{j_1} Z_{i_1, j_1} \otimes_1 [(1 - X) f_B^{j_1}] = f^{i_1} - \sum_j Z_{i_1, j} \otimes_1 (X f_B^j) \end{aligned} \quad (98)$$



and hence

$$\Delta^{i_1} = \sum_{i_2, j} \left( \int_2 X_2 Z_{i_1 i_2, j} \right) \otimes_1 (X f_B^j) - \sum_j Z_{i_1, j} \otimes_1 (X f_B^j) + X_1 f^{i_1}. \quad (99)$$

In this expression we can rewrite

$$X f_B^j = \sum_k X (Z_{j, k}^{-1} \otimes f^k) = \sum_k (X Z_{j, k}^{-1}) \otimes (X f^k) \quad (100)$$

using (64). Multiplying (75) with  $x$  and again using (64), we see that  $x Z_{i, j}^{-1}(x)$  is the inverse of  $x Z_{i, j}(x)$  w.r.t. Mellin convolution and matrix multiplication. This allows us to write

$$\Delta^{i_1} = \sum_k R_k^{i_1} \otimes_1 (X f^k) \quad (101)$$

with

$$R_k^{i_1} = \sum_j \left[ \sum_{i_2} \left( \int_2 X_2 Z_{i_1 i_2, j} \right) - Z_{i_1, j} + X Z_{i_1, j} \right] \otimes_1 (X Z_{j, k}^{-1}). \quad (102)$$

The tree-level part of  $Z_{i_1, j}(x)$  cancels in this expression because it is proportional to  $\delta(1-x)$ , so that in the  $\overline{\text{MS}}$  scheme  $R^{i_1}$  is a sum of pure pole terms in  $\epsilon$ . Since  $\Delta^{i_1}$  is finite at  $\epsilon = 0$ , we must have  $R^{i_1} = 0$  and hence  $\Delta^{i_1} = 0$ , which concludes our argument.

From  $\sum_k R_k^{i_1} \otimes (X Z_{k, j}) = 0$  and (102) we obtain a momentum sum rule

$$\sum_{i_2} \int_2 X_2 Z_{i_1 i_2, j} = (1 - X_1) Z_{i_1, j}. \quad (103)$$

## 7 DPD evolution and its consequences

The renormalisation of momentum space DPDs results in an inhomogeneous evolution equation, which at leading order in  $\alpha_s$  has been known for a long time [17, 18]. Our all-order formulation of DPD renormalisation in section 6.2 allows us to derive the form of the corresponding evolution equation at arbitrary order in  $\alpha_s$ . Differentiating (70) and using that bare distributions are independent of the scale  $\mu$ , we obtain

$$\frac{dF^{i_1 i_2}(\Delta)}{d \ln \mu^2} = \sum_{j_1, j_2} \left[ \frac{dZ_{i_1, j_1}}{d \ln \mu^2} \otimes_1 Z_{i_2, j_2} + Z_{i_1, j_1} \otimes_1 \frac{dZ_{i_2, j_2}}{d \ln \mu^2} \right] \otimes_2 F_B^{j_1 j_2}(\Delta) + \sum_j \frac{dZ_{i_1 i_2, j}}{d \ln \mu^2} \otimes_{12} f_B^j. \quad (104)$$

The DGLAP equations for the renormalised PDFs  $f^i = \sum_j Z_{i, j} f_B^j$  imply

$$\frac{dZ_{i, k}}{d \ln \mu^2} = \sum_j P_{i, j} \otimes Z_{j, k}, \quad (105)$$

where  $P_{i, j}(x; \mu)$  denotes the usual DGLAP evolution kernels. We therefore have

$$\begin{aligned} \frac{dF^{i_1 i_2}(\Delta)}{d \ln \mu^2} &= \sum_{k_1, k_2} \left[ \sum_{j_1} P_{i_1, j_1} \otimes_1 Z_{j_1, k_1} \otimes_1 Z_{i_2, k_2} + Z_{i_1, k_1} \otimes_1 \sum_{j_2} P_{i_2, j_2} \otimes_2 Z_{j_2, k_2} \right] \otimes_2 F_B^{k_1 k_2}(\Delta) \\ &\quad + \sum_j \frac{dZ_{i_1 i_2, j}}{d \ln \mu^2} \otimes_{12} f_B^j \\ &= \sum_{j_1} P_{i_1, j_1} \otimes_1 F^{j_1 i_2}(\Delta) + \sum_{j_2} P_{i_2, j_2} \otimes_2 F^{i_1 j_2}(\Delta) \\ &\quad + \sum_j \left[ \frac{dZ_{i_1 i_2, j}}{d \ln \mu^2} - \sum_{j_1} P_{i_1, j_1} \otimes_1 Z_{j_1 i_2, j} - \sum_{j_2} P_{i_2, j_2} \otimes_2 Z_{i_1 j_2, j} \right] \otimes_{12} f_B^j. \end{aligned} \quad (106)$$

We now define evolution kernels  $P_{i_1 i_2, j}(x_1, x_2; \mu)$  associated with the splitting singularity in DPDs:

$$P_{i_1 i_2, k} = \sum_j \left[ \frac{dZ_{i_1 i_2, j}}{d \ln \mu^2} - \sum_{j_1} P_{i_1, j_1} \otimes_1 Z_{j_1 i_2, j} - \sum_{j_2} P_{i_2, j_2} \otimes_2 Z_{i_1 j_2, j} \right] \otimes_{12} Z_{j, k}^{-1}, \quad (107)$$

which is equivalent to

$$\frac{dZ_{i_1 i_2, k}}{d \ln \mu^2} = \sum_{j_1} P_{i_1, j_1} \otimes_1 Z_{j_1 i_2, k} + \sum_{j_2} P_{i_2, j_2} \otimes_2 Z_{i_1 j_2, k} + \sum_j P_{i_1 i_2, j} \otimes_{12} Z_{j, k} \quad (108)$$

and gives us the form

$$\frac{dF^{i_1 i_2}(\Delta)}{d \ln \mu^2} = \sum_{j_1} P_{i_1, j_1} \otimes_1 F^{j_1 i_2}(\Delta) + \sum_{j_2} P_{i_2, j_2} \otimes_2 F^{i_1 j_2}(\Delta) + \sum_j P_{i_1 i_2, j} \otimes_{12} f^j \quad (109)$$

for the inhomogeneous double DGLAP equation at arbitrary order in perturbation theory. Our result confirms the form given for NLO evolution in equation (16) of [20]. Comparing (108) with (109) we see that the renormalisation factor for the splitting singularity satisfies the same form of evolution equation as the renormalised DPD, in analogy to the renormalisation factor for PDFs in (105). With the definitions (71) for single and double Mellin moments, we find

$$\begin{aligned} \frac{dF^{i_1 i_2}(m_1, m_2; \Delta)}{d \ln \mu^2} &= \sum_{j_1} P_{i_1, j_1}(m_1) F^{j_1 i_2}(m_1, m_2; \Delta) + \sum_{j_2} P_{i_2, j_2}(m_2) F^{i_1 j_2}(m_1, m_2; \Delta) \\ &+ \sum_j P_{i_1 i_2, j}(m_1, m_2) f^j(m_1 + m_2 - 1) \end{aligned} \quad (110)$$

in agreement with the LO formulae derived in [17, 18].

Equations (104) to (110) are valid in  $D = 4 - 2\epsilon$  dimensions, and whenever the l.h.s. of an equation is finite for  $\epsilon \rightarrow 0$ , it is understood that this limit can be taken. In the  $\overline{\text{MS}}$  scheme, renormalisation factors contain only pole terms in  $\epsilon$ , plus an  $\epsilon$  independent tree-level term in the case of  $Z_{i, j}$  (but not of  $Z_{i_1 i_2, j}$ ). An important consequence of this is that the evolution kernels  $P_{i, j}$  and  $P_{i_1 i_2, j}$  are independent of  $\epsilon$ . They must be finite for  $\epsilon \rightarrow 0$ , and any terms with a positive power of  $\epsilon$  would induce positive powers of  $\epsilon$  on the r.h.s. of (105) or (108). Such terms cannot appear, because the form

$$\frac{d}{d \ln \mu^2} = [\beta(\alpha_s(\mu)) - \epsilon \alpha_s(\mu)] \frac{\partial}{\partial \alpha_s(\mu)} \quad (111)$$

of the renormalisation scale derivative in  $D$  dimensions implies that there are only terms of order  $\epsilon^n$  with  $n \leq 0$  on the l.h.s. of (105) and (108).

Using these results, we can obtain the kernel  $P_{i_1 i_2, j}$  by isolating the term of order  $\epsilon^0$  on the r.h.s. of (107). The only renormalisation factor with a tree-level term in that equation is  $Z_{j, k}^{-1}$ , so that  $P_{i_1 i_2, j}$  is given by the term of order  $\epsilon^0$  in  $dZ_{i_1 i_2, j}/d \ln \mu^2$ . Using (111), we thus find that

$$P_{i_1 i_2, j}(x_1, x_2; \alpha_s(\mu)) = -\alpha_s(\mu) \frac{\partial}{\partial \alpha_s(\mu)} \text{Res} Z_{i_1 i_2, j}(x_1, x_2; \alpha_s(\mu), \epsilon), \quad (112)$$

where Res denotes the residue of the single pole in  $\epsilon$ . Applying the same type of reasoning to (105), we obtain

$$P_{i, j}(x; \alpha_s(\mu)) = -\alpha_s(\mu) \frac{\partial}{\partial \alpha_s(\mu)} \text{Res} Z_{i, j}(x; \alpha_s(\mu), \epsilon). \quad (113)$$

Combining these relations with the sum rules (94) and (103) for the  $1 \rightarrow 2$  renormalisation factor  $Z_{i_1 i_2, j}$ , we readily obtain corresponding sum rules

$$\int_0^{1-x_1} dx_2 P_{i_1 i_2, v, j}(x_1, x_2) = (\delta_{i_1, \bar{i}_2} - \delta_{i_1, i_2} - \delta_{j, \bar{i}_2} + \delta_{j, i_2}) P_{i_1, j}(x_1), \quad (114)$$

$$\sum_{i_2} \int_0^{1-x_1} dx_2 x_2 P_{i_1 i_2, j}(x_1, x_2) = (1-x_1) P_{i_1, j}(x_1) \quad (115)$$

for the  $1 \rightarrow 2$  evolution kernels, where we have given momentum fractions explicitly instead of using the compact notation of section 6.1. We note that the momentum sum rule (115) has the same form as a corresponding sum rule derived in [37, 38] for “two-body inclusive decay probabilities”, see equation (49) in [38]. These quantities were introduced to describe the evolution of hadronic jets. To investigate their relation with our  $1 \rightarrow 2$  splitting kernels goes beyond the scope of this paper, but we note that the two types of functions start to differ at order  $\alpha_s^2$  [39].

The derivation of the DPD sum rules in section 6 did not refer to any particular value of  $\mu$  and is hence valid independently of the renormalisation scale. As a consistency check, one can use the all-order form (109) of DPD evolution and the sum rules just stated to verify that the DPD sum rules (9) and (10) are stable under a change of scale. This means that the renormalisation scale derivative of their l.h.s. is equal to the renormalisation scale derivative of their r.h.s. Let us first show this for the number sum rule. Using our compact notation, the derivative of the l.h.s. of (10) can be written as

$$\frac{d}{d \ln \mu^2} \int_2 F^{i_1 i_2, v} = \sum_{j_1} P_{i_1, j_1} \otimes_1 \int_2 F^{j_1 i_2, v} + \sum_{j_2} \int P_{i_2, v, j_2} \int_2 F^{i_1 j_2} + \sum_j \left( \int_2 P_{i_1 i_2, v, j} \right) \otimes_1 f^j \quad (116)$$

according to (109) and the rules of computation from section 6.1. In the first term of (116) we can use the DPD sum rule at scale  $\mu$  and in the last term the sum rule (114) for the  $1 \rightarrow 2$  evolution kernel. The second term is zero thanks to the sum rule

$$\int P_{i_2, v, j} = 0 \quad (117)$$

for the PDF evolution kernels, which readily follows from (90) and (113). We thus get

$$\begin{aligned} \frac{d}{d \ln \mu^2} \int_2 F^{i_1 i_2, v} &= \sum_{j_1} P_{i_1, j_1} \otimes_1 (N_{i_2, v} + \delta_{j_1, \bar{i}_2} - \delta_{j_1, i_2}) f^{j_1} \\ &\quad + \sum_j (\delta_{i_1, \bar{i}_2} - \delta_{i_1, i_2} - \delta_{j, \bar{i}_2} + \delta_{j, i_2}) P_{i_1, j} \otimes_1 f^j \\ &= (N_{i_2, v} + \delta_{i_1, \bar{i}_2} - \delta_{i_1, i_2}) \sum_j P_{i_1, j} \otimes_1 f^j. \end{aligned} \quad (118)$$

On the r.h.s. we recognise the derivative  $df^{i_1}/d \ln \mu^2$  and thus the appropriate derivative of the r.h.s. of the DPD number sum rule (10). For the momentum sum rule, we proceed in full analogy and start with

$$\begin{aligned} \frac{d}{d \ln \mu^2} \sum_{i_2} \int_2 X_2 F^{i_1 i_2} &= \sum_{j_1} P_{i_1, j_1} \otimes_1 \sum_{i_2} \int_2 X_2 F^{j_1 i_2} + \sum_{j_2, i_2} \int X P_{i_2, j_2} \int_2 X_2 F^{i_1 j_2} \\ &\quad + \sum_{j, i_2} \left( \int_2 X_2 P_{i_1 i_2, j} \right) \otimes_1 (X f^j). \end{aligned} \quad (119)$$

The second term on the r.h.s. is zero thanks to the momentum sum rule

$$\sum_i \int X P_{i,j} = 0 \quad (120)$$

for the DGLAP kernels. Using the DPD momentum sum rule at scale  $\mu$  and the relation (115), we get

$$\begin{aligned} \frac{d}{d \ln \mu^2} \sum_{i_2} \int_2 X_2 F^{i_1 i_2} &= \sum_{j_1} P_{i_1, j_1} \otimes_1 [(1-X) f^{j_1}] + \sum_j [(1-X) P_{i_1, j}] \otimes_1 (X f^j) \\ &= \sum_j P_{i_1, j} \otimes_1 f^j - \sum_j (X P_{i_1, j}) \otimes_1 (X f^j) \\ &= (1 - X_1) \sum_j P_{i_1, j} \otimes_1 f^j, \end{aligned} \quad (121)$$

where in the last step we have used the relation (64). The last line of (121) is the scale derivative of the r.h.s. of the DPD momentum sum rule (9), as required. We note that the inhomogeneous term in the double DGLAP equation is essential for the preceding arguments to work. For leading-order evolution, this was already emphasised in [8, 21].

## 8 Conclusion

The sum rules proposed by Gaunt and Stirling [14] present one of the few general constraints on double parton distributions that are currently known. This has motivated us to give a detailed proof for them in QCD. We saw in section 3 that an analysis of Feynman graphs in covariant perturbation theory yields the sum rules in simple cases but quickly becomes complicated for certain types of graphs, which makes this technique unsuitable for a general proof. Instead, we used light-cone perturbation theory in section 5 to show the validity of the DPD sum rules for bare, i.e. unrenormalised, distributions at any order in the coupling. In section 6 we analysed the renormalisation of DPDs and showed that in the  $\overline{\text{MS}}$  scheme this procedure yields renormalised distributions that again satisfy the sum rules.

As by-products of our analysis, we derived in section 7 an all-order evolution equation for DPDs in momentum space and obtained sum rules for the kernel  $P_{i_1 i_2, j}$  that appears in the inhomogeneous term of that equation. These sum rules can be used to verify explicitly that the DPD sum rules are consistent with evolution at any order in perturbation theory. They will also provide valuable cross-checks for the calculation of  $P_{i_1 i_2, j}$  beyond the known order  $\alpha_s$ . Introducing a compact notation and deriving a number of relations for convolution integrals in one or two variables (section 6.1) allowed us to keep the computations for renormalised DPDs reasonably short and transparent.

To construct DPD models that fulfil the sum rules – exactly or approximately – is by far not an easy task [14, 15]. The results of the present work provides an additional motivation for further efforts in this direction. In a forthcoming numerical study [40] we will show how the sum rules can be used to improve existing models for DPDs in position space.

## Acknowledgements

We gratefully acknowledge discussions with Jonathan Gaunt.

## References

- [1] T. Kasemets and S. Scopetta, *Parton correlations in double parton scattering*, *Adv. Ser. Direct. High Energy Phys.* **29** (2018) 49 [1712.02884].
- [2] M. Diehl and A. Schäfer, *Theoretical considerations on multiparton interactions in QCD*, *Phys. Lett.* **B698** (2011) 389 [1102.3081].
- [3] M. G. Ryskin and A. M. Snigirev, *A Fresh look at double parton scattering*, *Phys. Rev.* **D83** (2011) 114047 [1103.3495].
- [4] B. Blok, Yu. Dokshitzer, L. Frankfurt and M. Strikman, *pQCD physics of multiparton interactions*, *Eur. Phys. J.* **C72** (2012) 1963 [1106.5533].
- [5] A. V. Manohar and W. J. Waalewijn, *What is Double Parton Scattering?*, *Phys. Lett.* **B713** (2012) 196 [1202.5034].
- [6] M. G. Ryskin and A. M. Snigirev, *Double parton scattering in double logarithm approximation of perturbative QCD*, *Phys. Rev.* **D86** (2012) 014018 [1203.2330].
- [7] J. R. Gaunt, *Single Perturbative Splitting Diagrams in Double Parton Scattering*, *JHEP* **01** (2013) 042 [1207.0480].
- [8] B. Blok, Yu. Dokshitzer, L. Frankfurt and M. Strikman, *Perturbative QCD correlations in multi-parton collisions*, *Eur. Phys. J.* **C74** (2014) 2926 [1306.3763].
- [9] A. M. Snigirev, N. A. Snigireva and G. M. Zinovjev, *Perturbative and nonperturbative correlations in double parton distributions*, *Phys. Rev.* **D90** (2014) 014015 [1403.6947].
- [10] K. Golec-Biernat and E. Lewandowska, *Electroweak boson production in double parton scattering*, *Phys. Rev.* **D90** (2014) 094032 [1407.4038].
- [11] J. R. Gaunt, R. Maciula and A. Szczurek, *Conventional versus single-ladder-splitting contributions to double parton scattering production of two quarkonia, two Higgs bosons and  $c\bar{c}\bar{c}$* , *Phys. Rev.* **D90** (2014) 054017 [1407.5821].
- [12] M. Rinaldi, S. Scopetta, M. C. Traini and V. Vento, *Correlations in Double Parton Distributions: Perturbative and Non-Perturbative effects*, *JHEP* **10** (2016) 063 [1608.02521].
- [13] M. Diehl, J. R. Gaunt and K. Schönwald, *Double hard scattering without double counting*, *JHEP* **06** (2017) 083 [1702.06486].
- [14] J. R. Gaunt and W. J. Stirling, *Double Parton Distributions Incorporating Perturbative QCD Evolution and Momentum and Quark Number Sum Rules*, *JHEP* **03** (2010) 005 [0910.4347].
- [15] K. Golec-Biernat, E. Lewandowska, M. Serino, Z. Snyder and A. M. Stasto, *Constraining the double gluon distribution by the single gluon distribution*, *Phys. Lett.* **B750** (2015) 559 [1507.08583].
- [16] J. R. Gaunt, C.-H. Kom, A. Kulesza and W. J. Stirling, *Same-sign  $W$  pair production as a probe of double parton scattering at the LHC*, *Eur. Phys. J.* **C69** (2010) 53 [1003.3953].
- [17] R. Kirschner, *Generalized Lipatov-Altarelli-Parisi Equations and Jet Calculus Rules*, *Phys. Lett.* **B84** (1979) 266.

- [18] V. P. Shelest, A. M. Snigirev and G. M. Zinovev, *The Multiparton Distribution Equations in QCD*, *Phys. Lett.* **B113** (1982) 325.
- [19] A. M. Snigirev, *Double parton distributions in the leading logarithm approximation of perturbative QCD*, *Phys. Rev.* **D68** (2003) 114012 [hep-ph/0304172].
- [20] F. A. Ceccopieri, *An update on the evolution of double parton distributions*, *Phys. Lett.* **B697** (2011) 482 [1011.6586].
- [21] F. A. Ceccopieri, *A second update on double parton distributions*, *Phys. Lett.* **B734** (2014) 79 [1403.2167].
- [22] Gaunt, Jonathan R., *Double parton scattering in proton-proton collisions*, Ph.D. thesis, University of Cambridge, 2012. <https://doi.org/10.17863/CAM.16589>.
- [23] M. G. A. Buffing, M. Diehl and T. Kasemets, *Transverse momentum in double parton scattering: factorisation, evolution and matching*, *JHEP* **01** (2018) 044 [1708.03528].
- [24] A. Vladimirov, *Structure of rapidity divergences in multi-parton scattering soft factors*, *JHEP* **04** (2018) 045 [1707.07606].
- [25] J. C. Collins, D. E. Soper and G. F. Sterman, *Soft Gluons and Factorization*, *Nucl. Phys.* **B308** (1988) 833.
- [26] J. Collins, *Foundations of perturbative QCD*, *Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol.* **32** (2011) 1.
- [27] M. Diehl, J. R. Gaunt, D. Ostermeier, P. Plöchl and A. Schäfer, *Cancellation of Glauber gluon exchange in the double Drell-Yan process*, *JHEP* **01** (2016) 076 [1510.08696].
- [28] S.-J. Chang and S.-K. Ma, *Feynman rules and quantum electrodynamics at infinite momentum*, *Phys. Rev.* **180** (1969) 1506.
- [29] J. B. Kogut and D. E. Soper, *Quantum Electrodynamics in the Infinite Momentum Frame*, *Phys. Rev.* **D1** (1970) 2901.
- [30] T.-M. Yan, *Quantum field theories in the infinite momentum frame. 4. Scattering matrix of vector and Dirac fields and perturbation theory*, *Phys. Rev.* **D7** (1973) 1780.
- [31] S. J. Brodsky and G. P. Lepage, *Exclusive Processes in Quantum Chromodynamics*, *Adv. Ser. Direct. High Energy Phys.* **5** (1989) 93.
- [32] W.-M. Zhang and A. Harindranath, *Light front QCD. 2: Two component theory*, *Phys. Rev.* **D48** (1993) 4881.
- [33] N. E. Ligterink and B. L. G. Bakker, *Equivalence of light front and covariant field theory*, *Phys. Rev.* **D52** (1995) 5954 [hep-ph/9412315].
- [34] Y. V. Kovchegov and E. Levin, *Quantum chromodynamics at high energy*, vol. 33. Cambridge University Press, 2012, 10.1017/CBO9781139022187.
- [35] J. Collins, *The non-triviality of the vacuum in light-front quantization: An elementary treatment*, 1801.03960.
- [36] M. Diehl, D. Ostermeier and A. Schäfer, *Elements of a theory for multiparton interactions in QCD*, *JHEP* **03** (2012) 089 [1111.0910].

- [37] J. Kalinowski, K. Konishi and T. R. Taylor, *Jet Calculus Beyond Leading Logarithms*, *Nucl. Phys.* **B181** (1981) 221.
- [38] J. Kalinowski, K. Konishi, P. N. Scharbach and T. R. Taylor, *Resolving QCD Jets Beyond Leading Order: Quark Decay Probabilities*, *Nucl. Phys.* **B181** (1981) 253.
- [39] M. Diehl, J. R. Gaunt, P. Plöbl and A. Schäfer, *Two-loop splitting in double parton distributions*, 1902.08019.
- [40] M. Diehl, J. R. Gaunt, D. Lang, P. Plöbl and A. Schäfer. *in preparation*.