# Double hard scattering without double counting

Markus Diehl,[a] Jonathan R. Gaunt[b] and Kay Schönwald[c]

[a]*Deutsches Elektronen-Synchroton DESY, Notkestraße 85, 22607 Hamburg, Germany*

[b]*Nikhef Theory Group and VU University Amsterdam, De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands*

[c]*Deutsches Elektronen-Synchrotron DESY, Platanenallee 6, 15738 Zeuthen, Germany*

*E-mail:* markus.diehl@desy.de, jgaunt@nikhef.nl, kay.schoenwald@desy.de

ABSTRACT: Double parton scattering in proton-proton collisions includes kinematic regions in which two partons inside a proton originate from the perturbative splitting of a single parton. This leads to a double counting problem between single and double hard scattering. We present a solution to this problem, which allows for the definition of double parton distributions as operator matrix elements in a proton, and which can be used at higher orders in perturbation theory. We show how the evaluation of double hard scattering in this framework can provide a rough estimate for the size of the higher-order contributions to single hard scattering that are affected by double counting. In a numeric study, we identify situations in which these higher-order contributions must be explicitly calculated and included if one wants to attain an accuracy at which double hard scattering becomes relevant, and other situations where such contributions may be neglected.

arXiv:1702.06486v2 [hep-ph] 1 Jun 2017

# Contents

---

## 1  Introduction

The precise description of high-energy proton-proton collisions in QCD is imperative for maximising the physics potential of the LHC and of possible future hadron colliders. An important issue in this context is to understand the mechanism of double parton scattering (DPS), in which two pairs of partons undergo a hard scattering in one and the same proton-proton collision. In many situations DPS is suppressed compared with single parton scattering (SPS), but this suppression generically becomes weaker with increasing collision energy. For specific kinematics or specific final states, DPS can become comparable to or even larger than SPS. An overview of recent experimental and theoretical activities in this area can for instance be found in [1, 2].

Consider a DPS process $pp \to Y_1 + Y_2 + X$, where $Y_1$ and $Y_2$ are observed particles or groups of particles produced in two separate hard scattering processes, whilst $X$ denotes all unobserved particles in the final state. A cross section formula has been put forward long ago in the framework of collinear factorisation, where the transverse momenta $\boldsymbol{q}_1$ and $\boldsymbol{q}_2$ of $Y_1$ and $Y_2$ are integrated over [3]. A corresponding expression has been given in [4, 5] for transverse momentum dependent (TMD) factorisation, where $\boldsymbol{q}_1$ and $\boldsymbol{q}_2$ are small compared with the hard scales in $Y_1$ and $Y_2$.

Inside a proton, the two partons that take part in the two hard scatters can originate from the perturbative splitting of one parton. The relevance of this splitting mechanism for the evolution equations of double parton distributions (DPDs) has been realised long ago [6, 7] and studied more recently in [8–10]. However, it was only noted in [4, 5] that the same mechanism dominates DPDs in the limit of small transverse distance between the two partons, and that the splitting contribution leads to infinities when inserted into the DPS cross section formula. These infinities are closely connected with double counting between DPS and SPS in particular Feynman graphs, a problem that had been pointed out earlier in the context of multi-jet production [11].

Different ways of dealing with this issue have been proposed in [12, 13], [14, 15] and [16]. As discussed later in this paper, we find that these proposals have shortcomings either of theoretical or of practical nature. In the present work we present an alternative scheme for computing the cross section in a consistent way, including both DPS and SPS (as well as other contributions). Our scheme allows for a nonperturbative definition of DPDs in terms of operator matrix elements, and it is suitable for pushing the limit of theoretical accuracy to higher orders in $\alpha_s$.

Our paper is structured as follows. In section 2 we recall the theoretical framework for describing DPS and specify the theoretical problems mentioned above. The short-distance behaviour of DPDs is discussed in section 3, since it is essential for the scheme we propose in section 4. We first show how this scheme works at leading order (LO) in $\alpha_s$, before giving examples of its application at next-to-leading order (NLO) in section 5. Collinear DPDs

evolve according to DGLAP equations, and in section 6 we discuss several consequences of this scale evolution. Our scheme is naturally formulated with DPDs that depend on the transverse distance $\boldsymbol{y}$ between the two partons, but we show in section 7 how one may instead use DPDs depending on the transverse momentum conjugate to $\boldsymbol{y}$. This allows us to compare our results with those of [12, 13] and [14, 15], which we do in section 8. Whilst the focus of the present work is theoretical rather than phenomenological, we give in section 9 some quantitative illustrations of our scheme, obtained with a relatively simple ansatz for the DPDs. We summarise our findings in section 10. Some Fourier integrals required in the main text are given in an appendix.

## 2  Setting the scene

In this section we recall theoretical issues originating from the perturbative splitting mechanism in double parton distributions, namely the appearance of ultraviolet divergences in the naive cross section for double parton scattering, the problem of double counting between DPS and single parton scattering, and the treatment of the so-called 2v1 (two versus one) contributions to DPS.[1] We also give some basic definitions and results. Throughout this work we use light-cone coordinates $v^{\pm} = (v^0 + v^3)/\sqrt{2}$ for any four-vector $v$. We write $\boldsymbol{v} = (v^1, v^2)$ for the transverse components and $v = \sqrt{\boldsymbol{v}^2}$ for the length of the transverse vector.

Since the perturbative splitting mechanism in DPDs leads to issues in the ultraviolet (UV) region, renormalisation plays a crucial role in our context. We work in dimensional regularisation and extend the definitions of [5] to $D = 4 - 2\epsilon$ dimensions, changing phase space factors $(2\pi)^4$ into $(2\pi)^D$. Bare two-parton distributions are then given by

$$F_{\text{bare}}(x_i, \boldsymbol{z}_i, \boldsymbol{y}) = 2p^+ \int dy^- \int \frac{dz_1^-}{2\pi} \frac{dz_2^-}{2\pi} e^{i(x_1 z_1^- + x_2 z_2^-)p^+} \langle p | \mathcal{O}_2(0, z_2) \, \mathcal{O}_1(y, z_1) | p \rangle \quad (2.1)$$

with twist-two operators

$$\mathcal{O}_i(y, z_i) = \bar{q}\big(y - \tfrac{1}{2} z_i\big) W^{\dagger}\big(y - \tfrac{1}{2} z_i\big) \Gamma_i W\big(y + \tfrac{1}{2} z_i\big) q\big(y + \tfrac{1}{2} z_i\big)\big|_{z_i^+ = y^+ = 0} \quad (2.2)$$

for quark distributions, where $q$ denotes the bare field. $W(z)$ is a Wilson line from $z$ to infinity along a prescribed path, which we do not recall here. The Dirac matrices $\Gamma_i$ select different quark polarisations. Analogous definitions hold for distributions involving gluons, with quark fields replaced by gluon field strength operators. For ease of writing, we omit colour labels on the operators and distributions throughout this work, bearing in mind that different colour couplings are possible for the four parton fields in (2.1).

In the process of deriving factorisation, one finds that the proton matrix element in (2.1) needs to be multiplied by a combination of so-called soft factors, which are vacuum expectation values of products of Wilson lines. More information for the case of single parton distributions can be found in chapter 13.7 of [18] and in the recent overview [19].

---

[1]We follow here the nomenclature of [17]. The 2v1 contribution is referred to as $4 \times 2$ in [5], as $3 \to 4$ in [12, 13] (where four-jet production is considered), and as $1 \times 2$ in [14, 15].

A brief account for DPDs is given in section 2.1 of [20], and a more detailed discussion will be provided in [21]. The product $F_{\text{bare}} \times \{\text{soft factors}\}$ depends on a parameter $\zeta$ which regulates rapidity divergences. A scheme in which a soft factor appears explicitly in the cross section formula was presented in [22]. Since the treatment of soft gluons does not affect parton splitting in any special way, we will not discuss it further in the present paper. Correspondingly, we will suppress the argument $\zeta$ in all DPDs.

As a final step one performs UV renormalisation, which we assume to be done in the $\overline{\text{MS}}$ scheme. The DPDs obtained from (2.1) are appropriate for TMD factorisation, with the transverse positions $\boldsymbol{z}_i$ being Fourier conjugate to the transverse parton momenta that determine the final state kinematics. These distributions renormalise multiplicatively, with one renormalisation factor for each product of a quark field (or gluon field strength) with the Wilson line at the same transverse position and one factor for each pair of Wilson lines at the same transverse position in the soft factors. Denoting the product of these factors with $Z$, one obtains the final DPD as $F = \lim_{\epsilon \to 0} \big( Z \times F_{\text{bare}} \times \{\text{soft factors}\} \big)$. It obeys a renormalisation group equation which is a straightforward generalisation of the one for TMDs (given e.g. in [19]).

The DPDs needed for collinear factorisation are obtained by setting $\boldsymbol{z}_i = \boldsymbol{0}$ in $F_{\text{bare}}$ and in the associated soft factors, before renormalisation. Setting $\boldsymbol{z}_i = \boldsymbol{0}$ introduces ultraviolet divergences in the operators $\mathcal{O}_1$ and $\mathcal{O}_2$, and in the associated soft factors. The renormalised DPDs are then obtained as $F = \lim_{\epsilon \to 0} \big( Z_1 \otimes Z_2 \otimes F_{\text{bare}} \times \{\text{soft factors}\} \big)$, where $Z_i$ renormalises the operators associated with parton $i$ and where the convolution products are in the momentum fractions $x_i$. In the colour singlet channel, where both operators $\mathcal{O}_i$ in (2.1) are colour singlets, the soft factors reduce to unity and one obtains the renormalised twist-two operators that appear in single parton densities.

Since the operators associated with partons 1 and 2 renormalise independently (both for the TMD and the collinear case) one may choose different renormalisation scales $\mu_1$ and $\mu_2$ in each of them. This is useful when the two hard subprocesses in double parton scattering have widely different hard scales. In particular, one can then approach the kinematics of the so-called underlying event, with a very hard scattering at scale $\mu_1$ and additional jet production at a much lower scale $\mu_2$ (of course $\mu_2$ needs to remain in the perturbative region for our factorisation approach to be justified).

With different scales $\mu_1, \mu_2$ in the collinear DPDs, we have a homogeneous evolution equation

$$\frac{d}{d \log \mu_1^2} \, F_{a_1 a_2}(x_1, x_2, \boldsymbol{y}; \mu_1, \mu_2)$$
$$= \sum_{b_1} \int_{x_1}^{1-x_2} \frac{dx_1'}{x_1'} \, P_{a_1 b_1}\Big(\frac{x_1}{x_1'}, \alpha_s(\mu_1)\Big) \, F_{b_1 a_2}(x_1', x_2, \boldsymbol{y}; \mu_1, \mu_2) \qquad (2.3)$$

in $\mu_1$ and its analogue for $\mu_2$. For colour singlet DPDs, the kernels $P_{a_1 b_1}$ on the r.h.s. are the usual DGLAP splitting functions for single parton densities. In colour non-singlet channels, both the DPDs and the splitting kernels have an additional dependence on the rapidity parameter $\zeta$ [21]. Following the notation of [5], the labels $a_i, b_i$ in (2.3) denote both the species and the polarisation of the partons. The relevant labels are $q, \Delta q, \delta q$ for

unpolarised, longitudinally polarised and transversely polarised quarks, likewise $\bar{q}, \Delta\bar{q}, \delta\bar{q}$ for antiquarks, and $g, \Delta g, \delta g$ for unpolarised, longitudinally polarised and linearly polarised gluons, respectively. Note that the polarisations of the two partons can be correlated with each other, even in an unpolarised proton.

To simplify our presentation, we will consider the production $pp \rightarrow V_1 + V_2 + X$ of two electroweak gauge bosons $V_i = \gamma^*, Z, W$. Our results readily generalise to other processes for which DPS factorisation can be established; in the case of TMD factorisation this requires that the produced particles are colour singlets. We denote the four-momenta of the two bosons by $q_i$, their squared invariant masses by $Q_i^2 = q_i^\mu q_{i\mu}$ and their rapidities by $Y_i = \frac{1}{2}\log(q_i^+/q_i^-)$. We work in the proton-proton centre-of-mass frame, taking the proton with momentum $p$ ($\bar{p}$) to move in the positive (negative) 3 direction. Furthermore we define

$$x_i = \sqrt{\frac{Q_i^2}{s}} \, e^{Y_i} \,, \qquad\qquad \bar{x}_i = \sqrt{\frac{Q_i^2}{s}} \, e^{-Y_i} \,, \qquad (2.4)$$

with $s = (p + \bar{p})^2$. For the phase space of each gauge boson we have

$$d^4 q_i = \frac{s}{2} \, dx_i \, d\bar{x}_i \, d^2 \boldsymbol{q}_i = \frac{1}{2} \, dY_i \, dQ_i^2 \, d^2 \boldsymbol{q}_i \,. \qquad (2.5)$$

The "naive" cross section formulae (not taking into account the UV problems discussed below) read

$$\frac{d\sigma_{\text{DPS}}}{dx_1 \, dx_2 \, d\bar{x}_1 \, d\bar{x}_2 \, d^2\boldsymbol{q}_1 \, d^2\boldsymbol{q}_2} = \frac{1}{C} \sum_{a_1 a_2 b_1 b_2} \hat{\sigma}_{a_1 b_1}(Q_1^2, \mu_1^2) \, \hat{\sigma}_{a_2 b_2}(Q_2^2, \mu_2^2)$$

$$\times \int \frac{d^2 \boldsymbol{z}_1}{(2\pi)^2} \frac{d^2 \boldsymbol{z}_2}{(2\pi)^2} \, d^2\boldsymbol{y} \, e^{-i(\boldsymbol{q}_1 \boldsymbol{z}_1 + \boldsymbol{q}_2 \boldsymbol{z}_2)} \, F_{b_1 b_2}(\bar{x}_i, \boldsymbol{z}_i, \boldsymbol{y}; \mu_i) \, F_{a_1 a_2}(x_i, \boldsymbol{z}_i, \boldsymbol{y}; \mu_i) \qquad (2.6)$$

for TMD factorisation and

$$\frac{d\sigma_{\text{DPS}}}{dx_1 \, dx_2 \, d\bar{x}_1 \, d\bar{x}_2} = \frac{1}{C} \sum_{a_1 a_2 b_1 b_2} \int_{x_1}^{1-x_2} \frac{dx_1'}{x_1'} \int_{x_2}^{1-x_1'} \frac{dx_2'}{x_2'} \int_{\bar{x}_1}^{1-\bar{x}_2} \frac{d\bar{x}_1'}{\bar{x}_1'} \int_{\bar{x}_2}^{1-\bar{x}_1'} \frac{d\bar{x}_2'}{\bar{x}_2'}$$

$$\times \hat{\sigma}_{a_1 b_1}(x_1' \bar{x}_1' s, \mu_1^2) \, \hat{\sigma}_{a_2 b_2}(x_2' \bar{x}_2' s, \mu_2^2) \int d^2\boldsymbol{y} \, F_{b_1 b_2}(\bar{x}_i', \boldsymbol{y}; \mu_i) \, F_{a_1 a_2}(x_i', \boldsymbol{y}; \mu_i) \qquad (2.7)$$

for collinear factorisation. The combinatorial factor $C$ is 1 if the observed final states of the hard scatters are different and 2 if they are identical. For simplicity we will consider the case $C = 1$ throughout this paper, unless mentioned otherwise. As explained in section 2.2.1 of [5], there are further contributions with DPDs that describe the interference of different parton species. They can be discussed in full analogy to the contributions given in (2.6) or (2.7), and we do not treat them explicitly in the present work for ease of notation.

The hard scattering cross sections $\hat{\sigma}_i$ in (2.6) are for the exclusive final state $V_i$ with transverse momentum $\boldsymbol{q}_i$, which must satisfy $q_i \ll Q_i$. By contrast, $\hat{\sigma}_i$ in (2.7) is integrated over $\boldsymbol{q}_i$ and inclusive for final states $V_i + X$. At leading order in $\alpha_s$ it contains $\delta$ functions that enforce $x_i' = x_i$ and $\bar{x}_i' = \bar{x}_i$. The subtractions for collinear and soft regions in $\hat{\sigma}_i$ are different in the two factorisation frameworks, but in both cases they lead to a dependence

on the factorisation scale $\mu_i$ that cancels against the $\mu_i$ dependence of the DPDs, up to powers in $\alpha_s$ beyond the accuracy of the calculation. This happens separately for the two hard scatters ($i = 1, 2$) and by construction works exactly as in the case of single hard scattering.

As was pointed out in [5], the framework discussed so far suffers from problems in the region of small transverse distances between the two partons in a DPD. The leading behaviour of the collinear distributions $F(x_i, \boldsymbol{y})$ at small $y$ can be computed from the perturbative splitting of one parton into two and gives a behaviour like $y^{-2}$ up to logarithmic corrections. When inserted in the factorisation formula (2.7) this gives a quadratically divergent integral at small $y$, which clearly signals an inappropriate treatment of the ultraviolet region. As we will review in section 3.1, the short-distance behaviour of the distributions $F(x_i, \boldsymbol{z}_i, \boldsymbol{y})$ is less singular but still leads to logarithmic divergences in the TMD factorisation formula (2.6).

Instead of using DPDs depending on transverse positions, one may Fourier transform them to transverse momentum space, integrating

$$\frac{1}{(2\pi)^{4-4\epsilon}} \int d^{2-2\epsilon}\boldsymbol{z}_1 \, d^{2-2\epsilon}\boldsymbol{z}_2 \; e^{-i(\boldsymbol{z}_1\boldsymbol{k}_1 + \boldsymbol{z}_2\boldsymbol{k}_2)} \tag{2.8}$$

for TMDs and

$$\int d^{2-2\epsilon}\boldsymbol{y} \; e^{i\boldsymbol{y}\boldsymbol{\Delta}} \tag{2.9}$$

for TMDs and collinear distributions. For collinear distributions, this transformation must be made before subtracting UV divergences and setting $\epsilon \to 0$: with $F(x_i, \boldsymbol{y}) \sim y^{-2}$ in $D = 4$ dimensions, the Fourier integral (2.9) would be logarithmically divergent at $y = 0$. In $D = 4 - 2\epsilon$ dimensions this singularity turns into a pole in $1/\epsilon$, which requires an additional subtraction as we will review in section 7. Rather than being associated with the individual operators $\mathcal{O}_1$ and $\mathcal{O}_2$, this subtraction is related to the singularity arising when the transverse distance $y$ between the two operators goes to zero. It leads to an *inhomogeneous* evolution equation for the DPD $F(x_i, \boldsymbol{\Delta})$ in momentum space, which has been extensively discussed in the literature [6–10] for the case $\boldsymbol{\Delta} = \boldsymbol{0}$. Notice, however, that this extra $\mu$ dependence does *not* cancel in the cross section when (2.7) is rewritten in transverse momentum space. Moreover, the additional UV renormalisation of $F(x_i, \boldsymbol{\Delta})$ does not remove all UV divergences at the cross section level. The singularity of $F(x_i, \boldsymbol{y})$ at small $y$ translates into a behaviour $F(x_i, \boldsymbol{\Delta}) \sim \log(\mu^2/\Delta^2)$ at large $\Delta$ (see section 7), which gives a quadratic divergence for the $\boldsymbol{\Delta}$ integration in the DPS cross section.

It is easy to identify the origin of the UV divergences just discussed. Both in the $\boldsymbol{y}$ and $\boldsymbol{\Delta}$ representations, one has integrated over the full range of the integration variable and thus left the region in which the approximations leading to the DPS cross section formulae are valid, namely the region where $\Delta \ll Q_i$ or, equivalently, $y \gg 1/Q_i$ (see section 2.1.2 in [5]). Outside this region, the DPS approximations are not only unjustified, but they give divergent integrals in the cross section.

This points to another problem, namely that of double counting contributions between SPS and DPS. To see this, let us analyse the graph in figure 1a. Since the transverse boson
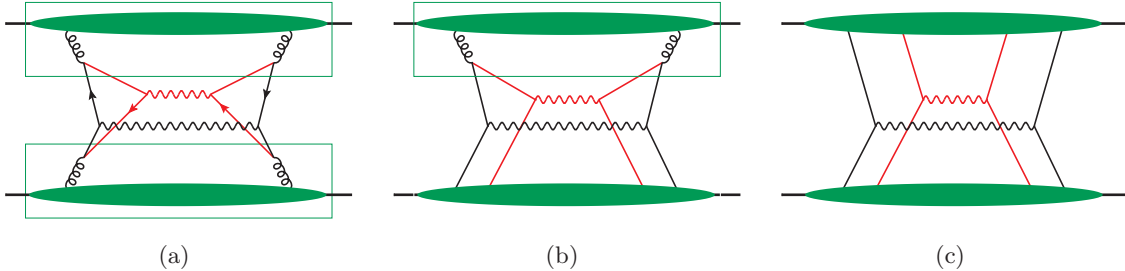
**Figure 1**. (a) A 1v1 contribution to DPS, with perturbative splitting DPDs for each proton (indicated by the boxes). (b) A 2v1 contribution to DPS, with a perturbative splitting DPD in only one proton (indicated by the box). (c) A 2v2 contribution to DPS. Here and in the following it is understood that partons emerging from the oval blobs are approximately collinear to their parent proton. A line for the final state cut across the blobs and the produced vector bosons (wavy lines) is not shown for simplicity.

momenta $\boldsymbol{q}_1$ and $\boldsymbol{q}_2$ are approximately back to back (up to effects from the transverse momenta of the incoming gluons) it is convenient to introduce the combination

$$\boldsymbol{q} = \tfrac{1}{2}(\boldsymbol{q}_1 - \boldsymbol{q}_2). \tag{2.10}$$

For $q \ll Q_i$ the graph gives a leading contribution to $d\sigma/dq^2$ if the transverse quark momenta in the loops are all of order $Q_i$. This carries the quark lines far off shell, so that this contribution is naturally associated with SPS, with $g + g \to V_1 V_2$ as the hard subprocess. A leading contribution is also obtained from the region where all transverse quark momenta are much smaller than $Q_i$. This region is naturally described as DPS, with two disconnected hard scattering processes $q\bar{q} \to V_1$ and $q\bar{q} \to V_2$ and double parton distributions with perturbative $g \to q\bar{q}$ splittings, indicated by the boxes in the graph. We denote this as a 1v1 (1 versus 1) contribution to DPS, emphasising its close relation to SPS.

A double counting problem for this graph obviously arises if one takes the loop integrals in the SPS cross section over all transverse quark momenta (including the DPS region), and likewise if one integrates the DPDs cross section over the full range of transverse positions, which is equivalent to integrating over all transverse momenta in the quark loops. The problem persists if one integrates the cross section over $\boldsymbol{q}$.

Let us now turn to the graph in figure 1b and consider the cross section integrated over $\boldsymbol{q}_1$ and $\boldsymbol{q}_2$. In the region of large $q \sim Q_i$, the quark lines at the top of the graph have transverse momenta of order $q$ and are far off shell. The proper description of this region is in terms of a hard scattering $q\bar{q} + g \to V_1 V_2$, convoluted with a collinear twist-two distribution (i.e. an ordinary PDF) at the top and a collinear twist-four distribution at the bottom of the graph. For brevity we refer to this as the "twist-four contribution" henceforth. In the region $q \ll Q_i$, the $g \to q\bar{q}$ splittings are near collinear and the approximations for DPS are appropriate. We call this the 2v1 contribution to DPS, recalling that there is a $q\bar{q} + g$ subprocess in the graph. Both small and large $q$ give leading contributions to the integrated cross section, and in a naive calculation adding up the twist-four term

and the DPS term has again a double counting problem, as well as divergences in each contribution. The naive DPS cross section has a logarithmic divergence at small $y$, which is seen by inserting the $1/y^2$ splitting behaviour of only one DPD in (2.7). In turn, the hard scattering cross section in the twist-four term contains a collinear divergence in the form of an integral behaving like $dq^2/q^2$ at $q \to 0$, as we will show in section 4.1.2.

Clearly, one needs a consistent scheme for computing the overall cross section, without double counting and without divergences in individual terms. An intuitive approach for evaluating DPS is to separate the "perturbative splitting" part of a DPD from its "intrinsic" nonperturbative part.[2] This has been pursued independently by Blok et al. [12, 13] and by Ryskin and Snigirev [14, 15]. Taking the intrinsic DPD for each proton, one obtains the 2v2 part of DPS, which does not contain any perturbative splitting and is shown in figure 1c. The splitting part of the DPD is explicitly computed in terms of a single parton distribution function (PDF) and a perturbative kernel. This is multiplied with an intrinsic DPD to compute the 2v1 term. Finally, the product of two splitting DPDs is used to compute the 1v1 contribution in the approach of [14, 15], where an ultraviolet cutoff must be imposed to regulate the quadratic divergence we mentioned earlier. By contrast, the authors of [12, 13] advocate to omit this term and replace it entirely with the SPS contribution to the cross section.

We are, however, not able to give a field theoretic definition of the "intrinsic" or "nonperturbative" part of a DPD. The consideration of Feynman graphs in the preceding arguments is instructive, but a satisfactory definition should only appeal to perturbation theory in regions where it is applicable. We regard a nonperturbative definition of DPDs as indispensable for a systematic theory approach, for instance for deriving evolution equations and other general properties.

The setup we propose in this work defines DPDs as operator matrix elements as described above, containing both splitting and intrinsic contributions. UV divergences in the DPS cross section are avoided by introducing (smooth or hard) cutoffs in the integrations over transverse distances. The double counting problems are treated within the subtraction formalism used in standard factorisation theorems, described in detail in sections 10.1 and 10.7 of [18] and briefly recalled in section 4 here. The subtraction terms that avoid double counting also remove the above mentioned collinear divergence in the twist-four term. A distinction between "splitting" and "intrinsic" contributions to a DPD will be made in the limit of small transverse distances, where it can be formulated in terms of an operator product expansion (see section 3.3), and when making a model ansatz for DPDs at large distances, which is of course necessary for phenomenology.

## 2.1 Contributions to the cross section: power behaviour and logarithms

In preparation for later sections, we now recall some results for the power behaviour of different contributions to the cross section, referring to section 2.4 of [5] for a derivation. We also recall which logarithms appear in the lowest order 1v1 and 1v2 graphs. As already

---

[2]The intrinsic part of a DPD may be studied using quark models [23–30], at least in the valence region, or it may be related to the product of two PDFs if correlations between the two partons are neglected.
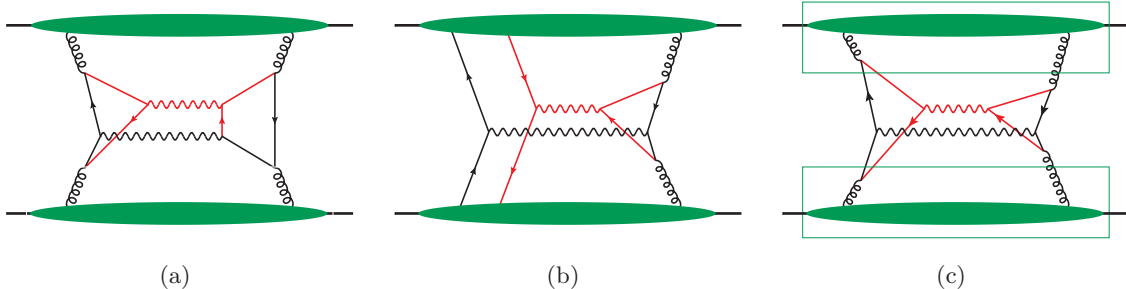
**Figure 2**. (a) A higher-order graph for SPS. The loop on the l.h.s. has a momentum region associated with DPS but the loop on the r.h.s. does not. (b) Interference between DPS in the amplitude and SPS in the complex conjugate amplitude. (c) 1v1 graph in the region where the quark loop on the left is collinear whilst the one on the right is hard. This corresponds to the region of SPS/DPS interference, with the boxes indicating DPDs with perturbative splitting for the parton pair in the amplitude.

stated, we take the process $pp \to V_1 + V_2 + X$ as a concrete example, but the discussion readily generalises to other cases.

The differential cross section $d\sigma/(d^2\boldsymbol{q}_1 d^2\boldsymbol{q}_2)$ in the region $q_i \ll Q_i$ can be computed using TMD factorisation. Here and in the following we write $Q_i$ to denote the generic size of $Q_1$ and $Q_2$, and likewise for $q_i$. The transverse momenta $q_i$ may be of nonperturbative size $\Lambda$ or much larger. In the latter case, further simplifications are possible by expressing transverse momentum dependent distributions in terms of collinear ones [21], but we shall not discuss this here. The leading power behaviour of the cross section is

$$\frac{d\sigma}{d^2\boldsymbol{q}_1 d^2\boldsymbol{q}_2} \sim \frac{1}{Q_i^4 \, q_i^2} \,. \tag{2.11}$$

When $q_i$ goes to zero, it should be replaced by $\Lambda$. Three types of mechanisms contribute to the leading behaviour, namely DPS, SPS, and the interference between SPS and DPS. Corresponding graphs are shown in figures 1c, 2a and 2b, respectively.

As discussed in the previous section, certain graphs contribute both to DPS and to SPS, depending on the kinematics of their internal lines. The 1v1 graph in figure 1 also has leading regions in which one of the loops is hard and the other is collinear. These regions contribute to the SPS/DPS interference, as shown in figure 2c. The double counting problem thus concerns both SPS, DPS and their interference. Note that the SPS graph in figure 2a contributes to the SPS/DPS interference but not to DPS.

Both the amplitude and its conjugate in the 1v1 graph contains a loop integral that behaves like $d^2\boldsymbol{k}/\boldsymbol{k}^2$ in the region $\Lambda, q_i \ll k \ll Q_i$. When integrated over the full phase space, each loop thus builds up a so-called DPS logarithm, namely $\log(Q_i/q_i)$ when $q_i \gtrsim \Lambda$ and $\log(Q_i/\Lambda)$ when $q_i \lesssim \Lambda$. Whether these logarithms reside in SPS, DPS or their interference depends on how exactly one handles the double counting problem. We come back to this issue in section 4.2.

Let us now turn to the cross section integrated over $\boldsymbol{q}_1$ and $\boldsymbol{q}_2$, which can be described using collinear factorisation. The leading power behaviour of the cross section, $\sigma \sim 1/Q_i^2$, is
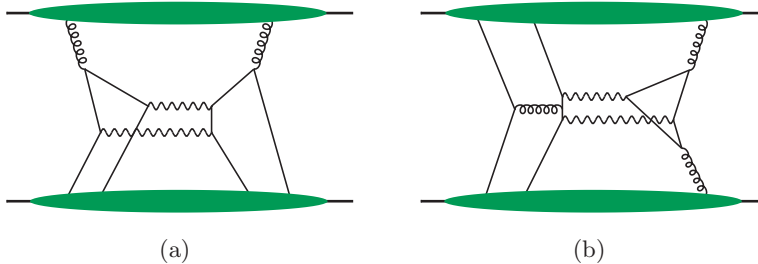
**Figure 3**. (a) A graph with a twist-four distribution for one proton and a twist-two distribution for the other. (b) A graph with twist-three distributions for both protons.

given by the SPS mechanism alone. Several mechanisms contribute at the power suppressed level $\sigma \sim \Lambda^2/Q_i^4$. These are

1. DPS, which is suppressed because it can only populate the region $q_i \ll Q_i$ rather than the full phase space up to $q_i \sim Q_i$,

2. the interference between SPS and DPS, which is suppressed for the same reason,

3. hard scattering with a twist-four distribution for one proton and a twist-two distribution for the other. Example graphs are figure 3a, as well as figure 1b with the box removed.

4. hard scattering with twist-three distributions for both protons. An example graph is figure 3b.

The rationale for considering such contributions is that – whilst being power suppressed compared with SPS – they may be enhanced by higher parton luminosities at small momentum fractions $x$, or by coupling constants in the relevant hard scattering subprocesses. Let us emphasise that a complete calculation of the cross section at the level of $\Lambda^2/Q_i^2$ corrections would be a formidable task, and it is not even established whether factorisation (in particular the cancellation of Glauber gluons) holds at that level.

Notice that in collinear factorisation, the SPS/DPS interference term involves collinear twist three distributions for both protons, because the SPS mechanism forces the two partons in the interfering DPS amplitude to be at same transverse position (see section 2.4.1 in [5]). In this sense, mechanism 2 in the above list may be regarded as a special case of mechanism 4, with a disconnected hard scattering in the amplitude or its conjugate (see figure 2b).

A full treatment of contributions with twist-three or twist-four distributions is beyond the scope of this paper. We remark however that twist-$n$ operators contain $n$ or less than $n$ parton fields, and that different operators are related by the equations of motion. For a detailed discussion we refer to [31–33]. Twist-$n$ operators with $n$ parton fields were called "'quasipartonic" in [34] and involve only the "good" parton fields in the parlance of light-cone quantisation [35]. These are exactly the fields appearing in the definitions

of multiparton distributions, so that graphs with a double counting issue between higher-twist hard scattering and DPS (or the SPS/DPS interference) involve only quasipartonic operators.

The matrix elements of quasipartonic twist-three operators in an unpolarised target satisfy the important selection rule that the helicities carried by the parton lines must balance . This excludes three-gluon operators since three helicities $\pm 1$ cannot add up to zero. For quark-antiquark-gluon operators it forces the quark and antiquark fields to have opposite chirality, i.e. one only has the operator combination $\bar{q}\sigma^{+j}q$, where the transverse index $j$ is contracted with the polarisation index of the gluon. As for non-quasipartonic twist-three distributions in an unpolarised target, one finds that they are absent in the pure gluon sector [36], whereas the corresponding quark-antiquark distributions are again chiral-odd [37]. Since chiral-odd distributions cannot be generated by gluon ladder graphs, they lack the small $x$ enhancement that is one of the motivations to keep higher twist contributions in the cross section. We will therefore not discuss them further in this work. Note that corresponding selection rules do *not* hold for TMD correlators, where an imbalance in the helicities of the parton fields can be compensated by orbital angular momentum.

Let us finally recall the appearance of DPS logarithms in collinear factorisation. The 2v1 graph (figure 1b) has a behaviour $d\sigma/dq^2 \sim 1/q^2$ in the region $\Lambda \ll q \ll Q_i$, which gives a $\log(Q_i/\Lambda)$ when integrated over the full phase space. Depending on how the double counting between DPS and the twist-four mechanism is resolved, this logarithm can appear in different contributions to the cross section. We will discuss this in section 4.1.2.

## 3   Short-distance limit of DPDs

In this section, we analyse the behaviour of DPDs in the limit where the transverse distance between partons becomes small compared with the scale of nonperturbative interactions. In this region, the splitting of one parton into two becomes dominant. Generalising results in [5] we give expressions in $D = 4 - 2\epsilon$ dimensions, which are necessary in intermediate steps when constructing a factorisation formula for the cross section.

### 3.1   TMDs

A useful choice of position variables for describing the parton splitting mechanism is

$$\boldsymbol{y}_{\pm} = \boldsymbol{y} \pm \tfrac{1}{2}(\boldsymbol{z}_1 - \boldsymbol{z}_2), \qquad\qquad \boldsymbol{Z} = \tfrac{1}{2}(\boldsymbol{z}_1 + \boldsymbol{z}_2) \qquad (3.1)$$

with Fourier conjugate momenta [3]

$$\boldsymbol{k}_{\pm} = \tfrac{1}{2}(\boldsymbol{k}_1 - \boldsymbol{k}_2 \pm \boldsymbol{\Delta}), \qquad\qquad \boldsymbol{K} = \boldsymbol{k}_1 + \boldsymbol{k}_2. \qquad (3.2)$$

The relation between DPDs in position and momentum space reads

$$
\begin{aligned}
F(x_i, &\boldsymbol{y}_{\pm}, \boldsymbol{Z}) \\
&= \frac{1}{(2\pi)^{2-2\epsilon}} \int d^{2-2\epsilon}\boldsymbol{K}\, d^{2-2\epsilon}\boldsymbol{k}_+\, d^{2-2\epsilon}\boldsymbol{k}_-\, e^{i\boldsymbol{Z}\boldsymbol{K} + i(\boldsymbol{y}_+\boldsymbol{k}_- - \boldsymbol{y}_-\boldsymbol{k}_+)}\, F(x_i, \boldsymbol{k}_{\pm}, \boldsymbol{K}) \qquad (3.3)
\end{aligned}
$$

---

[3]The momentum $\boldsymbol{\Delta}$ is called $\boldsymbol{r}$ in [4, 5].
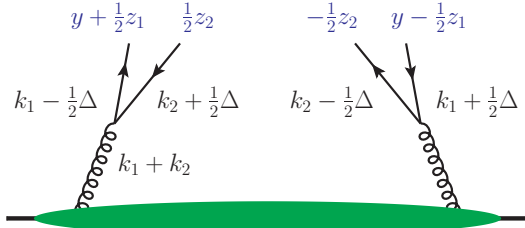
**Figure 4**. The perturbative splitting mechanism for a DPD, with momentum and position assignments. Here and in the following, the line for the final-state cut of the spectator partons is not shown for simplicity.

in $D = 4 - 2\epsilon$ dimensions. As seen in figure 4, one can identify $\boldsymbol{y}_+$ ($\boldsymbol{y}_-$) as the transverse distance between the two partons on the left (right) hand side of the final state cut in the DPD. Correspondingly, the transverse momentum difference between the partons on the left (right) hand side of the cut is $\boldsymbol{k}_-$ ($\boldsymbol{k}_+$). The splitting singularities of the DPDs thus occur at $y_\pm \to 0$ or $k_\pm \to \infty$.

The perturbative splitting contribution $F_{\text{spl,pt}}$ to transverse-momentum dependent DPDs in momentum space has been calculated at leading order in section 5.2.2 of [5]. Generalising these results to $D = 4 - 2\epsilon$ dimensions, we have

$$
F_{a_1 a_2, \text{spl,pt}}(x_i, \boldsymbol{k}_\pm, \boldsymbol{K}) = \frac{\boldsymbol{k}_-^j}{k_-^2} \frac{\boldsymbol{k}_+^{j'}}{k_+^2} \frac{(2\mu)^{2\epsilon}}{\pi^{1-2\epsilon}}
$$
$$
\times \left[ \frac{f_{a_0}(x_1 + x_2, \boldsymbol{K})}{x_1 + x_2} \frac{\alpha_s}{2\pi} T_{a_0 \to a_1 a_2}^{jj'}\left(\frac{x_1}{x_1 + x_2}, \epsilon\right) + \ldots \right], \quad (3.4)
$$

where $j$, $j'$ are transverse Lorentz indices and $f_{a_0}(x_1 + x_2, \boldsymbol{K})$ is an unpolarised single-parton TMD.[4] The ellipsis denotes a term that involves a TMD for polarised partons in an unpolarised proton and depends on $\boldsymbol{K}$ but not on $\boldsymbol{k}_\pm$. In position space we then get

$$
F_{a_1 a_2, \text{spl,pt}}(x_i, \boldsymbol{y}_\pm, \boldsymbol{Z}) = \frac{\boldsymbol{y}_+^j}{y_+^{2-2\epsilon}} \frac{\boldsymbol{y}_-^{j'}}{y_-^{2-2\epsilon}} \mu^{2\epsilon} \frac{\Gamma^2(1-\epsilon)}{\pi^{1-2\epsilon}}
$$
$$
\times \left[ \frac{f_{a_0}(x_1 + x_2, \boldsymbol{Z})}{x_1 + x_2} \frac{\alpha_s}{2\pi} T_{a_0 \to a_1 a_2}^{jj'}\left(\frac{x_1}{x_1 + x_2}, \epsilon\right) + \ldots \right] \quad (3.5)
$$

using the Fourier integral (A.2), where the term denoted by an ellipsis depends on $\boldsymbol{Z}$ but not on $\boldsymbol{y}_\pm$. It is understood that for transverse quark or linear gluon polarisation, both $F_{a_1 a_2}$ and the kernel $T$ carry additional transverse Lorentz indices. $f_{a_0}(x_1 + x_2, \boldsymbol{Z})$ is the Fourier transform of $f_{a_0}(x_1 + x_2, \boldsymbol{K})$. The form (3.4) gives the leading behaviour of the DPD for large $k_\pm \gg \Lambda$, and correspondingly (3.5) gives the leading behaviour for $y_\pm \ll 1/\Lambda$.

If one inserts these results into the cross section formula and sets $D = 4$, logarithmic divergences appear at $y_+ = 0$ and $y_- = 0$. To make them explicit we transform variables

---

[4]Compared with section 5.2 of [5], the kernel $T^{jj'}$ used here has the opposite order of indices $jj'$ and includes a colour factor, e.g. $T_F = 1/2$ for the colour singlet distribution $^1F_{g \to q\bar{q}}$.

to

$$\int d^2\boldsymbol{y}\, d^2\boldsymbol{z}_1\, d^2\boldsymbol{z}_2\, e^{-i(\boldsymbol{q}_1\boldsymbol{z}_1+\boldsymbol{q}_2\boldsymbol{z}_2)} = \int d^2\boldsymbol{Z}\, d^2\boldsymbol{y}_+\, d^2\boldsymbol{y}_-\, e^{-i(\boldsymbol{q}_1+\boldsymbol{q}_2)\boldsymbol{Z}} e^{-i\boldsymbol{q}(\boldsymbol{y}_+-\boldsymbol{y}_-)} \qquad (3.6)$$

with $\boldsymbol{q}$ defined in (2.10). Performing the angular integration in

$$\frac{1}{\pi}\int d^2\boldsymbol{y}\, e^{\pm i\boldsymbol{q}\boldsymbol{y}}\, \frac{\boldsymbol{y}^j\boldsymbol{y}^{j'}}{y^4} = \delta^{jj'}\int \frac{dy}{y}\, J_0(yq) + \left(\delta^{jj'} - \frac{2\boldsymbol{q}^j\boldsymbol{q}^{j'}}{q^2}\right)\int \frac{dy}{y}\, J_2(yq)\,, \qquad (3.7)$$

where $\boldsymbol{y}$ stands for $\boldsymbol{y}_+$ or $\boldsymbol{y}_-$, we see that the integral with $J_0$ is divergent at $y = 0$. Given the range of validity of (3.5) one should impose $y \ll 1/\Lambda$ in (3.7), although the integrals are finite for $y \to \infty$ due to the oscillations of the Bessel functions.

The perturbative splitting contribution to DPDs at higher order in $\alpha_s$ involves graphs with additional partons radiated into the final state as shown in figure 5a, as well as virtual corrections. It is natural to expect that it will again be singular at $y_\pm = 0$. A calculation of this contribution is outside the scope of the present work, so that we will limit our analysis of TMD factorisation to perturbative splitting at LO.

To compute the DPD cross section, we must also consider the case where only one of the distances $y_+$ or $y_-$ is small, whereas the other one remains large. In this case, one has a perturbative splitting only on one side of the final-state cut, as illustrated in figure 5b. We will not discuss the detailed expression of the DPD in this regime, but give its general structure. Setting $D = 4$ for simplicity, we have

$$F_{\alpha_1\alpha_2, y_- \to 0}(x_i, \boldsymbol{y}_\pm, \boldsymbol{Z}) = \frac{\boldsymbol{y}_-^{j'}}{y_-^2}\left[U_{\alpha_0\to\alpha_1\alpha_2}^{j'}(x_i)\right]^* D_{\alpha_1\alpha_2|\alpha_0}(x_i, \boldsymbol{y}_+, \boldsymbol{Z})\,, \qquad (3.8)$$

were $U_{\alpha_0\to\alpha_1\alpha_2}$ is a kernel for the splitting $\alpha_0 \to \alpha_1\alpha_2$ in the amplitude (hence its complex conjugate appears in (3.8)). $D_{\alpha_1\alpha_2|\alpha_0}$ is the position space version of a transverse-momentum dependent twist-three distribution, constructed from the hadronic matrix element

$$\langle p|\phi_0\left(-\tfrac{1}{2}Z\right)\phi_2\left(\tfrac{1}{2}(Z-y_+)\right)\phi_1\left(\tfrac{1}{2}(Z+y_+)\right)|p\rangle$$
$$= \langle p|\phi_0\left(\tfrac{1}{2}(y-Z)\right)\phi_2\left(\tfrac{1}{2}z_2\right)\phi_1\left(y+\tfrac{1}{2}z_1\right)|p\rangle\,, \qquad (3.9)$$

where $\phi_i$ is a "good" field for parton $\alpha_i$ (cf. section 2.1). Distributions $D_{\alpha_0|\alpha_1\alpha_2}(x_i, \boldsymbol{y}_-, \boldsymbol{Z})$ where $\alpha_0$ belongs to the amplitude and $\alpha_1, \alpha_2$ to the complex conjugate amplitude are defined in analogy. In the second step of (3.9) we have used translation invariance and shifted the parton fields to the same position as in the corresponding DPD (see figure 4). The labels $\alpha_i$ denote the parton species; it is understood that in (3.8) and the following equations parton helicities are taken fixed on the l.h.s. and must be appropriately summed over on the r.h.s. Note the difference between this notation and the labels $a_i$, which denote parton species and polarisation (none, longitudinal, transverse or linear) and thus refer to a *pair* of parton legs. The notation with $a_i$ is hence not suitable for distributions with three parton fields.
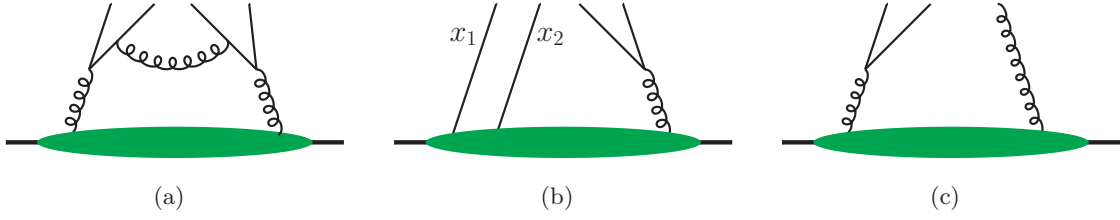
**Figure 5**. (a) A higher order contribution to the splitting part of a DPD. (b) Graph for a DPD with perturbative splitting only to the right of the final-state cut. The blob denotes a distribution $D_{\alpha_1\alpha_2|\alpha_0}$. (c) Graph for perturbative splitting in the distribution $D_{\alpha_1\alpha_2|\alpha_0}$.

If $y_+$ is small, then $D_{\alpha_1\alpha_2|\alpha_0}(x_i, \boldsymbol{y}_+, \boldsymbol{Z})$ itself can be generated by perturbative splitting, as shown in figure 5c. We have

$$D_{\alpha_1\alpha_2|\alpha_0, y_+\to 0}(x_i, \boldsymbol{y}_+, \boldsymbol{Z}) = \frac{\boldsymbol{y}_+^j}{y_+^2} U_{\alpha_0\to\alpha_1\alpha_2}^j(x_i)\, f_{\alpha_0}(x_1 + x_2, \boldsymbol{Z})\,. \tag{3.10}$$

Notice that a quark and antiquark produced by perturbative splitting have opposite helicities, so that the corresponding quark-antiquark operator $\phi_2\,\phi_1$ in $D_{q\bar{q}|g}$ must be chiral-even. Inserting (3.10) into (3.8) we obtain

$$F_{\alpha_1\alpha_2, y_\pm\to 0}(x_i, \boldsymbol{y}_\pm, \boldsymbol{Z})$$
$$= \frac{\boldsymbol{y}_+^j}{y_+^2} \frac{\boldsymbol{y}_-^{j'}}{y_-^2} U_{\alpha_0\to\alpha_1\alpha_2}^j(x_i) \left[U_{\alpha_0\to\alpha_1\alpha_2}^{j'}(x_i)\right]^* f_{\alpha_0}(x_1 + x_2, \boldsymbol{Z})\,. \tag{3.11}$$

Taking appropriate linear combinations of parton helicities, we recover the form of $F_{a_1a_2,\mathrm{spl,pt}}$ in (3.5) at $\epsilon = 0$.

## 3.2 Collinear DPDs: splitting contribution

We now turn to collinear DPDs, i.e. to the case where $\boldsymbol{z}_1 = \boldsymbol{z}_2 = \boldsymbol{0}$. Let us first consider distributions for two unpolarised or two longitudinally polarised partons, so that the DPDs do not carry any transverse Lorentz indices. The lowest order splitting has been computed in [5]. For $4 - 2\epsilon$ dimensions, one finds the general form

$$F_{a_1a_2,\mathrm{spl,pt}}(x_1, x_2, \boldsymbol{y}; \mu)$$
$$= \frac{\mu^{2\epsilon}}{y^{2-4\epsilon}} \frac{\Gamma^2(1-\epsilon)}{\pi^{1-2\epsilon}} \frac{f_{a_0}(x_1 + x_2; \mu)}{x_1 + x_2} \frac{\alpha_s(\mu)}{2\pi} P_{a_0\to a_1a_2}\left(\frac{x_1}{x_1 + x_2}, \epsilon\right)\,. \tag{3.12}$$

The kernel for the splitting $g \to q\bar{q}$ reads for instance

$$P_{g\to q\bar{q}}(u, \epsilon) = \frac{f}{2} \frac{u^2 + (1-u)^2 - \epsilon}{1 - \epsilon} \tag{3.13}$$

with a factor $f = 1$ for the colour singlet and $f = -1/\sqrt{N^2 - 1}$ for the colour octet DPD. In terms of the kernel in (3.4) we have $T_{g\to q\bar{q}}^{jj'} = \delta^{jj'} P_{g\to q\bar{q}}$. We recognise in $P_{g\to q\bar{q}}(u, 0)$ the usual DGLAP splitting kernel without the terms proportional to $\delta(1 - u)$.

Going beyond leading order, one can deduce the general form of the perturbative splitting contribution using dimensional analysis and boost invariance. For colour singlet distributions one finds

$$^{1}F_{a_1 a_2,\mathrm{spl,pt}}(x_1, x_2, \boldsymbol{y}; \mu)$$

$$= \frac{1}{\pi^{1-\epsilon} y^{2-2\epsilon}} \sum_{a_0} \int\limits_{x_1+x_2}^{1} \frac{dv}{v} \frac{f_{a_0}(v; \mu)}{v} V_{a_0 \to a_1 a_2}\left(\frac{x_1}{v}, \frac{x_2}{v}, \alpha_s(\mu), y\mu, \epsilon\right) \qquad (3.14)$$

in $D = 4 - 2\epsilon$ dimensions. The convolution integral over $v$ is familiar from factorisation formulae for hard scattering processes. Both $f$ and $V$ on the right-hand side are understood to include all necessary subtractions, so that they are finite at $\epsilon = 0$. The splitting kernel $V$ is a double series

$$V_{a_0 \to a_1 a_2}(v_1, v_2, \alpha_s, y\mu, \epsilon) = \sum_{n=1}^{\infty} \alpha_s^n \sum_{m=1}^{n} (y\mu)^{2\epsilon m} V_{a_0 \to a_1 a_2}^{(n,m)}(v_1, v_2, \epsilon). \qquad (3.15)$$

The $\mu$ (and thus on dimensional grounds the $y$) dependence of $V$ follows from the fact that the mass parameter of dimensional regularisation appears in graphs only via $\mu^{2\epsilon}\alpha_s(\mu)$; terms with $n > m$ in (3.15) are due to the subtractions of ultraviolet or collinear divergences. At lowest order, the hard splitting graphs are disconnected (with no partons across the final state cut), so that $V^{(1,1)}(v_1, v_2, \epsilon) = \delta(1 - v_1 - v_2) V^{(1)}(v_1, \epsilon)$. Inserted into (3.14) this gives a form consistent with (3.12). Using that at order $\alpha_s^n$ the poles of highest order are $1/\epsilon^{n-1}$, we find

$$V_{a_0 \to a_1 a_2}(v_1, v_2, \alpha_s, y\mu, 0) = \sum_{n=1}^{\infty} \alpha_s^n \sum_{m=0}^{n-1} \log^m(y\mu) V_{a_0 \to a_1 a_2}^{[n,m]}(v_1, v_2). \qquad (3.16)$$

in the physical limit $\epsilon = 0$.

For colour nonsinglet DPDs one must regulate rapidity divergences, which complicates the preceding result. Taking e.g. Wilson lines along non-lightlike paths introduces additional vectors and changes the analysis of boost properties of the kernel. We will not pursue this issue here.

DPDs with transverse quark or linear gluon polarisation carry transverse Lorentz indices. Their perturbative splitting expressions thus have a tensor structure containing additional factors of $\boldsymbol{y}^j/y$ compared with the formulae above. At leading order one readily finds from (3.5) and the appropriate splitting kernels that the factor $1/y^{2-4\epsilon}$ in (3.12) is to be replaced with $\boldsymbol{y}^j \boldsymbol{y}^{j'}/y^{4-4\epsilon}$ times a tensor constructed only from Kronecker deltas.

## 3.3 Collinear DPDs: all contributions

Let us now study the small $y$ behaviour of collinear DPDs in more general terms. We start by writing the relation between unrenormalised DPDs in position and momentum space as

$$(2\pi)^{D-2} F(\boldsymbol{y}) = \int d^{D-2}\boldsymbol{\Delta}\, F(\boldsymbol{\Delta}) + \int d^{D-2}\boldsymbol{\Delta}\, [e^{-i\boldsymbol{y}\boldsymbol{\Delta}} - 1]\, F(\boldsymbol{\Delta}), \qquad (3.17)$$
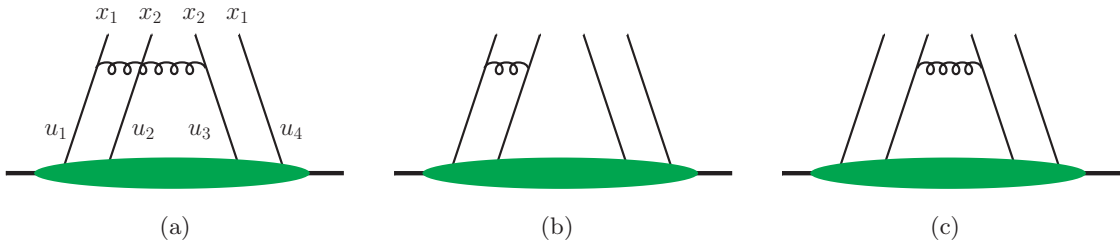
**Figure 6**. Graphs for the short-distance behaviour of a DPD that involve a twist-four distribution. $x_i$ and $u_i$ denote longitudinal momentum fractions.

omitting all arguments other than $\boldsymbol{y}$ and $\boldsymbol{\Delta}$. The first term on the r.h.s. is a collinear twist-four distribution, independent of any transverse variable. For small $y$, the second term is dominated by large $\boldsymbol{\Delta}$, so that one can replace $F(\boldsymbol{\Delta})$ by its approximation for large $\Delta$, following the power counting analysis of section 5.2 in [5]. This leads us to write

$$F_{y\to 0}(\boldsymbol{y}) = F_{\mathrm{spl,pt}}(\boldsymbol{y}) + F_{\mathrm{tw3,pt}}(\boldsymbol{y}) + F_{\mathrm{int,pt}}(\boldsymbol{y}), \qquad (3.18)$$

where the three terms on the r.h.s. will be described shortly. In $D = 4$ dimensions, they respectively go like $y^{-2}$, $y^{-1}$ and $y^0$, up to logarithmic corrections. Further terms from the perturbative expansion of $F(\boldsymbol{\Delta})$ give contributions to $F(\boldsymbol{y})$ that vanish like $y$ or faster.

One may also derive the expansion (3.18) from the operator product expansion, without taking recourse to the transverse momentum representation (3.17). In the definition of collinear DPDs one has a product $\mathcal{O}_2(0, z_2)\, \mathcal{O}_1(y, z_1)$ of operators with $\boldsymbol{z}_1 = \boldsymbol{z}_2 = \boldsymbol{0}$ but nonzero $\boldsymbol{y}$. This can be expanded around $\boldsymbol{y} = \boldsymbol{0}$ in terms of light-ray operators where all fields are at transverse position $\boldsymbol{0}$. These operators have twist 2, 3, 4 for the first, second and third term in (3.18), respectively.

The spitting contribution $F_{\mathrm{spl,pt}}$ is given by graphs as in figures 4 and 5a and has already been discussed in the previous subsection. The term $F_{\mathrm{tw3,pt}}$ originates from two types of graphs. The first type involves a single perturbative splitting and a quasipartonic collinear twist-three distribution as shown in figure 5b. The second type has two splittings as in figure 4 and a twist-three distribution with one "good" and one "bad" parton field. Given the helicity constraints discussed in section 2.1, collinear twist-three distributions in an unpolarised proton involve a quark and antiquark with opposite chirality (and possibly an extra gluon). As announced earlier, we discard twist-three terms in the following, since they are expected to become unimportant at small momentum fractions $x_1, x_2$.

Finally, the term $F_{\mathrm{int,pt}}$ contains contributions without any perturbative splitting; we hence refer to it as the "intrinsic" part of the DPD. It can be written as

$$F_{\mathrm{int,pt}}(x_1, x_2, \boldsymbol{y}; \mu) = G(x_1, x_2, x_2, x_1; \mu) + C(\cdots, \boldsymbol{y}; \mu) \otimes G(\cdots; \mu) + \ldots \qquad (3.19)$$

where $G$ is a quasipartonic collinear twist-four distribution and $C$ a perturbative splitting kernel, corresponding to graphs as in figure 6. The convolution $\otimes$ is in the longitudinal momentum fractions indicated by $\cdots$ (cf. figure 6a). The first term in (3.19) corresponds to

the first term in (3.17) and is the only contribution that does not involve a hard splitting at all. The ellipsis denotes terms with non-quasipartonic twist-four distributions containing three or two parton fields, together with one or two parton splittings. While having the same power behaviour in $y$, one may expect that at small $x_1, x_2$ these terms become less important than the terms with quasipartonic twist-four distributions, which should roughly grow as fast as the square of two parton densities.

We now take a closer look at the second term in (3.19). The kernel $C$ can be determined by computing both sides of (3.19) for a given graph. At order $\alpha_s$ only "non-diagonal" interactions, i.e. interactions connecting partons 1 and 2 as in figure 6a and b, contribute to $C$. The ladder graph in figure 6c is independent of $y$ and thus gives identical contributions to the matrix elements $F_{\text{int,pt}}$ and $G$. As a consequence it does not contribute to $C$.

At this point we can comment on the scale evolution of the different terms in (3.19). The l.h.s. evolves according to the homogeneous double DGLAP equation for DPDs, which describes "diagonal" interactions, either between the partons with final momentum fraction $x_1$ or between those with final momentum fraction $x_2$. By contrast, the evolution of $G(x_1, x_2, x_2, x_1; \mu)$ contains both diagonal and non-diagonal ladder interactions [34]. The non-diagonal interactions in the evolution must thus be cancelled by the $\mu$ dependence of the term $C \otimes G$. At leading order in $\alpha_s$, this dependence comes only from the coefficient function $C$, which indeed contains just non-diagonal interactions as just discussed.

We finally emphasise that an unambiguous decomposition of $F(\boldsymbol{y})$ into splitting, intrinsic and twist-three parts is only possible in the limit of small $y$. If $y$ is of hadronic size, neither the operator product expansion nor the notion of perturbative parton splitting make sense. One may however use the short-distance decomposition (3.18) as a starting point for a model parameterisation of DPDs in the full $y$ range. We describe a simple version of this strategy in section 9.

## 4 A scheme to regulate DPS and avoid double counting

In this section, we present a scheme that regulates the DPS cross section and solves the double counting problem between DPS and SPS, as well as between DPS and the twist-four contribution (figure 1b). Before doing so, we discuss some general considerations that motivate our scheme.

The following properties are in our opinion desirable for any theoretical setup describing double parton scattering.

1. It should permit a field theoretical definition of DPDs, without recourse to perturbation theory. This is the same standard as for the ordinary parton distributions in SPS processes. In particular, it allows one to derive general properties and to investigate these functions using nonperturbative methods, for instance lattice calculations.

   One may object that so far not even ordinary PDFs can be computed to a precision sufficient for phenomenology. However, important progress has been made in the area of lattice computations, and more can be expected for the future. Furthermore,

whereas ordinary PDFs are being extracted with increasing precision from experiment, it is hard to imagine a similar scenario for DPDs, because of their sheer number and because DPS processes are much harder to measure and analyse than most processes from which PDFs are extracted. In such a situation, even semi-quantitative guidance from nonperturbative calculations (such as the relevance of correlations of different types) is highly valuable.

As already discussed in section 2, the requirement of a nonperturbative definition prevents us from separating the "perturbative splitting" contribution of a DPD in a controlled way for all distances $\boldsymbol{y}$ (or equivalently for all conjugate momenta $\boldsymbol{\Delta}$).

2. To pave the way for increased theoretical precision, the scheme should permit a formulation at higher orders in perturbation theory. Furthermore, the complexity of the required higher order calculations should be manageable in practice.

3. For collinear factorisation, one would like to use as much as possible existing higher-order results for SPS processes, namely partonic cross sections and splitting functions. This means that the scheme should not modify the collinear subtractions to be made in hard scattering kernels, nor the validity of standard DGLAP evolution for DPDs in the colour singlet channel.

4. For TMD factorisation, it is desirable not to modify Collins-Soper evolution and the handling of rapidity divergences. This again allows one to re-use calculations done for SPS, although rapidity evolution for DPS is necessarily more complicated due to the complexities caused by colour [5, 21].

5. One would like to keep procedures as similar as possible for collinear and TMD factorisation. This will in particular facilitate the computation of DPS processes at perturbatively large transverse momenta in terms of collinear DPDs [21], adapting the well known procedure for single Drell-Yan production [38].

In principle one can use dimensional regularisation to handle the UV divergences that are induced in the DPS cross section by the perturbative splitting mechanism, as is done with the UV divergences that arise in simpler situations such as single hard scattering. However, contrary to that case, the UV divergences discussed in section 2 arise not at the level of individual DPDs but only when two DPDs are multiplied together and integrated over $\boldsymbol{y}$. This means that if one treats these divergences in dimensional regularisation, only the product of two DPDs is defined in $D = 4$ dimensions but not the DPDs separately. This possibility was explored in [16]. However, DPDs and their products remain nonperturbative functions at large $\boldsymbol{y}$, which according to present knowledge cannot be reduced to simpler quantities in a model independent way. In practice, one therefore needs to model or parameterise them at some starting scale. This is more involved for the product of DPDs than for DPDs themselves, as is the practical implementation of scale evolution. We will come back to this scheme in section 8.

**Ultraviolet regularisation.** We define the regularised DPS cross section by multiplying the integrand in the DPS formula (2.6) for measured transverse momenta with

$\Phi(y_+\nu)\,\Phi(y_-\nu)$ and the integrand in the collinear DPS formula (2.7) with $\Phi^2(y\nu)$. The function $\Phi(u)$ goes to 1 for $u \gg 1$ and to 0 for $u \to 0$, and we can restrict ourselves to the case where $0 \le \Phi(u) \le 1$ for all $u$. More specific requirements are given below.

Collinear and transverse-momentum dependent DPDs are defined as specified in section 2, without any modifications. Constructed from operator matrix elements, they contain both splitting and non-splitting contributions. They quantify specific properties of the proton and have a simple physical interpretation, with the same limitations as single parton densities. (We recall that a literal density interpretation of PDFs and TMDs is hindered by the presence of Wilson lines and of ultraviolet renormalisation.)

**Double counting subtraction.** To treat the double counting between DPS and other contributions, we adapt the recursive subtraction formalism of Collins, which we briefly sketch now (details are given in sections 10.1 and 10.7 of [18]). Consider a graph (or sum of graphs) $\Gamma$ that receives leading contributions from a set of loop momentum regions $R$. An approximation for $\Gamma$ is then given by

$$\Gamma \approx \sum_R C_R \Gamma \qquad \text{with} \quad C_R \Gamma = T_R \Gamma - \sum_{R' < R} T_R C_{R'} \Gamma\,. \qquad (4.1)$$

In each term one integrates over *all* loop momenta. The operator $T_R$ applies approximations designed to work in momentum region $R$. Subtraction terms avoid double counting the contributions from smaller regions $R'$ (regions that are contained in $R$). In these subtraction terms one applies the approximations designed for $R$ and those designed for the smaller regions. One can show [18] that $C_R \Gamma$ then provides a valid approximation in the region $R$ and in all smaller regions, and $\sum_R C_R \Gamma$ gives a valid approximation of the graph in all relevant regions. All approximations discussed here are valid up to power corrections.

In our context, we have graphs in which a set of collinear partons split into partons that can be either collinear (as in DPS) or hard (as in SPS). A slight adaptation of the above formalism is required since we compute DPS using DPDs and a regulating function $\Phi$ that depend on transverse distances $\boldsymbol{y}$ rather than transverse momenta. A collinear splitting region $R'$ then corresponds to large $y$ and the corresponding hard region $R$ to small $y$, but we keep the ordering of regions $R' < R$ from momentum space when implementing (4.1). We will show in section 4.3 that our use of subtractions in position space is equivalent to the one in momentum space up to power suppressed effects.

The subtraction terms for the DPS region turn out to have a very simple form. They can be obtained by replacing the DPDs in the UV regularised DPS cross section with their appropriate limits for small $y_\pm$ in TMD factorisation and for small $y$ in collinear factorisation. Details will be given in sections 4.1, 4.2 and 5.

Criteria 1 and 5 above are obviously satisfied in this scheme. Regarding criterion 2, we note that the higher-order calculations required for the double counting subtraction terms

are for the short-distance limit of DPDs, which involve much simpler Feynman graphs than the full scattering process.

The introduction of a function $\Phi$ in the DPS cross section avoids an explicit modification of the definition of DPDs and thus respects criteria 3 and 4. In particular, the collinear DPDs $F(x_i, \boldsymbol{y})$ in transverse position space follow the homogeneous DGLAP evolution equation (2.3). Since for colour singlet DPDs, the evolution kernels are the familiar DGLAP kernels, the associated scale dependence in the cross section cancels by construction against the one of the hard cross sections computed with the same collinear subtraction as for SPS.

In our scheme, we have introduced an additional momentum scale $\nu$ to separate DPS from SPS and the twist-four contribution. In practical calculations one may take $\nu$ equal to the UV renormalisation scale $\mu$ in DPDs, but we find it useful to keep it separate in the general discussion. As a minimal requirement, $\nu$ must be of perturbative size, so that the double counting subtraction terms remove all contributions from nonperturbative regions in the hard kernels of the SPS and twist-four cross sections, making their calculation consistent. By construction, the $\nu$ dependence in the physical cross section cancels between DPS and the double counting subtraction terms, up to higher orders in $\alpha_s$ that are beyond the accuracy of the computation. We will come back to this issue in section 6.2.

Let us now take a closer look at the properties required for the function $\Phi$.

- The introduction of $\Phi$ in the DPS cross section must not spoil the correct description of the physics in the region where the DPD approximations work. Specifically, we require that the modifications introduced by $\Phi$ in that region should be power suppressed in the large scale. This requires that $\Phi(u)$ must approach 1 sufficiently fast when $u \gg 1$.

  Anticipating the momentum space analysis in section 4.3, we demand that for collinear factorisation the Fourier transformation of $1 - \Phi^2$ exists in $2 - 2\epsilon$ dim (for positive and negative $\epsilon$). Therefore the integral $\int du\, u^{1-2\epsilon} \left[ 1 - \Phi^2(u) \right]$ must converge for $u \to \infty$. The corresponding criterion for TMD factorisation is obtained by replacing $\Phi^2 \to \Phi$.

- $\Phi(u)$ at small $u$ must regulate the UV divergences of the naive DPD cross section in 2 dimensions. This means that $\int du\, u^{-1} \Phi(u)$ must be integrable at $u = 0$ in TMD factorisation. For collinear factorisation we have the stronger requirement that $\int du\, u^{-3} \Phi^2(u)$ be integrable at $u = 0$. These criteria are satisfied if for $u \to 0$

$$\begin{aligned} \Phi(u) &= O(u^\delta) && \text{for TMD factorisation} \\ \Phi(u) &= O(u^{1+\delta}) && \text{for collinear factorisation} \end{aligned} \qquad (4.2)$$

  with some $\delta > 0$.

- To compute the double counting subtraction terms for SPS and the twist-four contribution, we must perform integrals over $\boldsymbol{y}$ in 2 and $2 - 2\epsilon$ transverse dimensions, respectively, as we shall see in the next sections. We choose $\Phi$ such that the required integrals are known analytically. This is especially important for the twist-four contribution, where we must expand the result around $\epsilon = 0$.

A suitable function for both collinear and TMD factorisation is the step function

$$\Phi(u) = \Theta(u - b_0) \qquad\qquad \text{with } b_0 = 2e^{-\gamma}, \qquad (4.3)$$

where $\gamma$ is Euler's constant. This corresponds to a hard lower cutoff $y > b_0/\nu$ in the $y$ integration, where the constant $b_0 \approx 1.12$ is taken to simplify certain analytical results. An alternative choice is

$$\Phi(u) = 1 - \exp(-u^2/4). \qquad (4.4)$$

Important integrals for these functions are given in table 1. Further possible choices for collinear factorisation are $\Phi^2(u) = 1 - \exp(-u^p)$ or $\Phi^2(u) = u^p/(1 + u^p)$ with $p > 2$; the corresponding integrals in the third and fourth row of table 1 can be performed and give Euler $\Gamma$ functions.

| $\Phi(u)$ | $\Theta(u - b_0)$ | $1 - \exp(-u^2/4)$ |
|---|---|---|
| $2\displaystyle\int_0^\infty \frac{du}{u}\, J_0(ur)\,\Phi(u)$ | $\log\dfrac{1}{r^2} + \mathcal{O}(r^2)$ | $\log\dfrac{1}{r^2} - \gamma + \mathcal{O}(r^2)$ |
| $2\displaystyle\int_0^\infty \frac{du}{u^{1-2\epsilon}}\,\Phi^2(u)$ | $-\dfrac{(b_0)^{2\epsilon}}{\epsilon} = -\dfrac{1}{\epsilon} + \mathcal{O}(\epsilon^0)$ | $2^\epsilon\left[1 - 2^{1+\epsilon}\right]\Gamma(\epsilon) = -\dfrac{1}{\epsilon} + \mathcal{O}(\epsilon^0)$ |
| $2\displaystyle\int_0^\infty \frac{du}{u^{3-2\epsilon}}\,\Phi^2(u)$ | $\dfrac{1}{1-\epsilon}\left(\dfrac{1}{b_0}\right)^{2-2\epsilon}$ | $\dfrac{1-2^\epsilon}{2^{1-\epsilon}}\Gamma(\epsilon - 1) = \dfrac{\log 2}{2} + \mathcal{O}(\epsilon)$ |

**Table 1**. Examples for the cutoff function $\Phi$ and relevant integrals. Exact results for the integral in the second row are given in (A.3) and (A.4). The integral in the third row converges for $\epsilon < 0$.

## 4.1 Leading order analysis: collinear factorisation

In this section we show how our formalism works in collinear factorisation, concentrating on the leading order in $\alpha_s$. Following the procedure of all-order factorisation proofs (see for instance [18]), we rewrite individual Feynman graphs in a way consistent with the final factorisation formula. At the end of this procedure, the factors associated with long-distance physics in that formula, such as PDFs and DPDs, are expressed in terms of operator matrix elements and thus defined in a nonperturbative way. During our Feynman graph analysis we can separately consider splitting and intrinsic contributions to a DPD, $F_{\rm spl}$ and $F_{\rm int}$. Likewise, we can consider separate contributions $\sigma_{\rm 1v1}$, $\sigma_{\rm 2v1}$ and $\sigma_{\rm 2v2}$ to the DPS cross section, which respectively involve the combinations $F_{\rm spl}\,F_{\rm spl}$, $F_{\rm int}\,F_{\rm spl} + F_{\rm spl}\,F_{\rm int}$ and $F_{\rm int}\,F_{\rm int}$, corresponding to graphs a, b and c in figure 1. In the final factorisation formula, only the full DPDs $F$ and the full DPS cross section $\sigma_{\rm DPS}$ will appear. For brevity, we write $\sigma$ instead of $d\sigma/(dx_1\,dx_2\,d\bar{x}_1\,d\bar{x}_2)$ here and in the following.

### 4.1.1 Squared box graph

Let us start with the squared box graph in figure 1a, integrated over all transverse momenta in the final state. As discussed in section 2, it receives its leading contribution from the region $q \sim Q$. The quark loops in the amplitude and its conjugate are then in the hard momentum region. This gives a part of the SPS cross section $\sigma_{\mathrm{SPS}}$ and is computed as usual in terms of PDFs and the cross section for $gg$ annihilation into two gauge bosons. Other graphs that contribute to SPS, such as the one in figure 2a, have no overlap with DPS. In the present subsection we restrict $\sigma_{\mathrm{SPS}}$ and other terms in the cross section to refer to the double box graph only.

The DPS region of the graph has near-collinear splittings in both the amplitude and its conjugate, which is only possible at $q \ll Q$. This gives the 1v1 contribution $\sigma_{\mathrm{1v1}}$ to the DPS cross section, which is power suppressed compared with $\sigma_{\mathrm{SPS}}$ because we integrate over the transverse boson momenta. Since the double box graph is integrated over all $q$ in $\sigma_{\mathrm{SPS}}$, we subtract the contribution of the DPS region in order to prevent double counting. (At leading power accuracy this is not strictly necessary, but we will shortly see that it is useful). According to (4.1) this contribution is obtained by applying to the graph the approximations for DPS as well as those for SPS. This gives the DPS expression with the $g \to q\bar{q}$ splittings computed perturbatively, the incoming gluons treated as collinear on-shell partons, and quark masses in the loop set to zero. These approximations are just what goes into the perturbative splitting approximation $F_{\mathrm{spl,pt}}$ of a DPD, which is given in (3.12). The subtraction term thus reads

$$\frac{d\sigma_{\mathrm{1v1,pt}}}{dx_1 \, dx_2 \, d\bar{x}_1 \, d\bar{x}_2} = \sum_{a_1 a_2 b_1 b_2} \hat{\sigma}_{a_1 b_1} \hat{\sigma}_{a_2 b_2} \int d^2\boldsymbol{y} \, \Phi^2(y\nu) \, F_{b_1 b_2, \mathrm{spl,pt}}(\bar{x}_i, \boldsymbol{y}) \, F_{a_1 a_2, \mathrm{spl,pt}}(x_i, \boldsymbol{y}) \,.$$

(4.5)

Note that in $\sigma_{\mathrm{1v1,pt}}$ we use the short-distance approximation $F_{\mathrm{spl,pt}}$ (even for large $\boldsymbol{y}$, where it is not valid), whereas in $\sigma_{\mathrm{1v1}}$ we use the un-approximated splitting part $F_{\mathrm{spl}}$ of each DPD. To ensure that $\sigma_{\mathrm{1v1}}$ and $\sigma_{\mathrm{1v1,pt}}$ receive contributions only from the DPS region, one should take $\nu \ll Q$.

To compute (4.5) we can use the known form (3.12) for the splitting DPDs, with the kernels collected in [39]. With $\sigma_{\mathrm{SPS}}$ being computed in fixed order perturbation theory, we make the same approximations in $\sigma_{\mathrm{1v1,pt}}$, taking the renormalisation and factorisation scales in the splitting kernel and in the collinear PDF fixed. In $D = 2$ dimensions the splitting DPDs then depend on $y$ like $1/y^2$. The $y$ integration in (4.5) is thus readily performed,

$$\int \frac{d^2\boldsymbol{y}}{y^4} \, \Phi^2(y\nu) = \nu^2 \int \frac{d^2\boldsymbol{u}}{u^4} \, \Phi^2(u) \,.$$

(4.6)

This is proportional to $\nu^2$, which is expected since the unregulated integral has a quadratic divergence in $y$ and $\nu$ is the scale regulating this divergence. The relevant integral for different $\Phi(u)$ is obtained from the last row of table 1 by setting $\epsilon = 0$.

We anticipate that the subtraction term at higher orders in $\alpha_s$ involves an integral as in (4.5), with splitting DPDs given by (3.14) and (3.16). Compared with (4.6) the integrand

then has additional powers of $\log(y\mu)$. The integrals can be obtained from those in the last row of table 1 by Taylor expanding around $\epsilon = 0$.

The complete contribution of the double box graph to the cross section is

$$\sigma_{\text{SPS}} - \sigma_{1\text{v}1,\text{pt}} + \sigma_{1\text{v}1} \,. \tag{4.7}$$

As is well known, an appropriate choice for the factorisation scale $\mu$ in collinear factorisation is $\mu \sim Q$. In the same spirit we take $\nu \sim Q$ rather than $\nu \ll Q$, although this extends the $y$ integrals in $\sigma_{1\text{v}1}$ and $\sigma_{1\text{v}1,\text{pt}}$ to values $y \sim 1/Q$ where the DPS approximation does not work. Both $\sigma_{1\text{v}1}$ and $\sigma_{1\text{v}1,\text{pt}}$ are then of the same order as $\sigma_{\text{SPS}}$ rather than power suppressed. Let us see that this still gives the correct approximation of the overall cross section.

Although $\sigma_{\text{SPS}}$ is naturally computed in momentum space, one can perform a Fourier transform w.r.t. the transverse momentum difference $\boldsymbol{\Delta}$ (defined as in figure 4 but for the graph of the cross section). Then $\sigma_{\text{SPS}}$ is a loop integral over the conjugate distance $\boldsymbol{y}$, just as the two other terms in (4.7).

We now show how the subtraction formalism works in this case. In the region of small $y \sim 1/Q$, one has $\sigma_{1\text{v}1} \approx \sigma_{1\text{v}1,\text{pt}}$, because the perturbative approximation of the DPD is designed to work at short distances. The dependence on the cutoff function $\Phi(y\nu)$ cancels between these two terms. One is therefore left with $\sigma_{\text{SPS}}$, which gives the appropriate description of the graph for $y \sim 1/Q$. We note that the DPDs in $\sigma_{1\text{v}1,\text{pt}}$ are computed at fixed order in $\alpha_s$, whereas in $\sigma_{1\text{v}1}$ they should be resummed using the double DGLAP equations. The cancellation between the two terms is therefore only up to higher orders in $\alpha_s$, which are beyond the accuracy of the result. We shall investigate this in detail in section 6.2.

Turning to the region of large $y \gg 1/Q$, one finds that $\sigma_{\text{SPS}} \approx \sigma_{1\text{v}1,\text{pt}}$, because precisely in that region the DPS approximation is designed to work for the graph. We are therefore left with the DPS term $\sigma_{1\text{v}1}$, which is exactly what we want to describe this region. The cutoff function $\Phi(y\nu)$ has no effect here, since we take $\nu \sim Q$ and require $\Phi(u) \approx 1$ for $u \gg 1$. The combination (4.7) thus gives the correct approximation for both small and large $y$.

### 4.1.2   2v1 graph

We now derive a description for the 2v1 graph in figure 1b. Integrated over transverse momenta in the final state, the graph receives leading contributions from small and large transverse momenta of the quarks on the splitting side (i.e. the top of the graph), corresponding to small or large transverse momenta of the produced gauge bosons. The region of small transverse momenta corresponds to DPS, whilst the contribution from large transverse momenta is correctly described by the twist-four mechanism introduced in section 2.

Whereas the SPS contribution in the previous subsection gives a finite result when computed as a perturbative two-loop graph, the twist-four contribution has a collinear divergence when integrated over the full region of phase space. To exhibit this divergence,
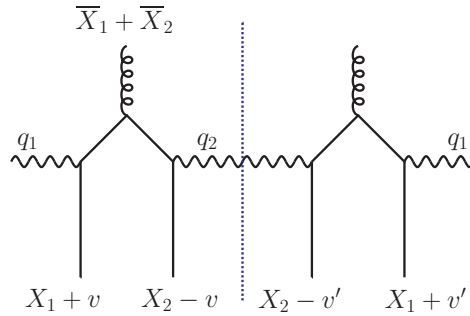
**Figure 7**. Lowest order graph for the 2v1 mechanism, with the blobs connecting partons with hadrons being omitted at the top and the bottom. The vector boson with momentum $q_1$ is understood to cross the final state cut. The four quark lines at the bottom are right moving, the two gluons entering at the top left moving.

we use dimensional regularisation and write

$$\frac{d\sigma_{\text{tw4}}}{dx_1\,dx_2\,d\bar{x}_1\,d\bar{x}_2} = \int d^{D-2}\boldsymbol{q} \int_{-X_1}^{X_2} dv \int_{-X_1}^{X_2} dv' \sum_{a_1 a_2} G_{a_1 a_2}(X_1 + v, X_2 - v, X_2 - v', X_1 + v')$$

$$\times H_{a_1 a_2, g}(\boldsymbol{q}, X_i, v, v', \overline{X}_i, s)\, f_g(\overline{X}_1 + \overline{X}_2)\,. \tag{4.8}$$

As in the previous subsection, we restrict the meaning of $\sigma_{\text{tw4}}$ and other terms in the cross section to the contribution from a single graph. The labels $a_1$ and $a_2$ indicate parton species and polarisation, and $G$ is a collinear twist-four distribution. $H$ includes overall factors and the squared amplitude of the hard scattering process $q\bar{q} + g \to V_1 V_2$. Note that in the hard scattering graph one has $\boldsymbol{q}_1 = -\boldsymbol{q}_2 = \boldsymbol{q}$. Momentum fractions are as shown in figure 7. The plus- and minus-momentum fractions of the produced bosons are respectively given by

$$X_i = \frac{q_i^+}{p^+} = \sqrt{\frac{Q_i^2 + q_i^2}{Q_i^2}}\, x_i\,, \qquad\qquad \overline{X}_i = \frac{q_i^-}{\bar{p}^-} = \sqrt{\frac{Q_i^2 + q_i^2}{Q_i^2}}\, \bar{x}_i \tag{4.9}$$

with $x_i$ and $\bar{x}_i$ defined in (2.4). For the momenta of internal lines in the hard scattering subprocess, minus and transverse components are fixed by the final state kinematics, so that there is only a loop integration over the plus-momentum fractions $v$ and $v'$. Note that by virtue of (4.9) both $X_i$ and $\overline{X}_i$ have an implicit dependence on $q^2$ at given $x_i, \bar{x}_i$. An upper limit on $q^2$ follows from the requirements that $X_1 + X_2 \le 1$ and $\overline{X}_1 + \overline{X}_2 \le 1$.

The expression (4.8) includes a contribution from DPS region, characterised by $q_i \ll Q_i$ and $v, v' \ll 1$. Let us approximate the integrand in that region. We first perform a Fierz transformation for the Dirac indices of the upper quark lines at the gauge boson vertices and retain only the leading Dirac structures. After this, the cross sections $\hat{\sigma}_i$ for $q\bar{q}$ annihilation appear in the expression. Furthermore we set $v = v' = 0$ in $G$ and extend the integrations over $v$ and $v'$ to the full real axis (after which they are easily done using Cauchy's theorem).

Approximating $X_i \approx x_i$ and $\overline{X}_i \approx \bar{x}_i$ and collecting all factors, we obtain

$$\left.\frac{d\sigma_{\text{tw4}}}{dx_1\, dx_2\, d\bar{x}_1\, d\bar{x}_2}\right|_{q \ll Q_i \text{ and } v,v' \ll 1} = \sum_{a_1 a_2 b_1 b_2} \hat{\sigma}_{a_1 b_1} \hat{\sigma}_{a_2 b_2} G_{a_1 a_2}(x_1, x_2, x_2, x_1)$$

$$\times \int\limits_{q \ll Q_i} d^{D-2}\boldsymbol{q}\, \frac{\boldsymbol{q}^j \boldsymbol{q}^{j'}}{q^4}\, \frac{(2\mu)^{4-D}}{\pi^{D-3}}\, \frac{\alpha_s}{2\pi}\, T^{jj'}_{g \to b_1 b_2}\left(\frac{\bar{x}_1}{\bar{x}_1 + \bar{x}_2}\right) \frac{f_g(\bar{x}_1 + \bar{x}_2)}{\bar{x}_1 + \bar{x}_2} \qquad (4.10)$$

with the splitting kernel $T$ from (3.4). If we drop the restriction $q \ll Q_i$, then the expression in the second line of (4.10) becomes the perturbative splitting approximation $F_{\text{spl,pt}}(\bar{x}_i, \boldsymbol{\Delta} = \boldsymbol{0})$ of the transverse-momentum integrated DPD. This DPD is evaluated at $\boldsymbol{\Delta} = \boldsymbol{0}$ in order to fulfil momentum conservation at the gauge boson vertices, because the transverse momenta of the right moving partons are set to zero in the overall hard scattering kernel $H$.

The $\boldsymbol{q}$ integral of (4.10) has an infrared divergence in $D = 4$ dimensions, which must be cancelled by the subtraction term for the DPS region. That term is obtained by applying the DPS approximations to the graph, in addition to those for the twist-four region. This gives again the form (4.10), but with an unrestricted integration over $\boldsymbol{q}$, and a regulator of the corresponding ultraviolet divergence. To implement the regulator, we Fourier transform the collinear DPD, $F_{\text{spl,pt}}(\bar{x}_i, \boldsymbol{\Delta} = \boldsymbol{0}) = \int d^{D-2}\boldsymbol{y}\, F_{\text{spl,pt}}(\bar{x}_i, \boldsymbol{y})$, and then multiply with $\Phi^2(y\nu)$ as prescribed by our formalism. This gives

$$\frac{d\sigma_{\text{2v1,pt}}}{dx_1\, dx_2\, d\bar{x}_1\, d\bar{x}_2}$$

$$= \sum_{a_1 a_2 b_1 b_2} \hat{\sigma}_{a_1 b_1} \hat{\sigma}_{a_2 b_2} G_{a_1 a_2}(x_1, x_2, x_2, x_1) \int d^{D-2}\boldsymbol{y}\, \Phi^2(y\nu)\, F_{b_1 b_2 \text{ spl,pt}}(\bar{x}_i, \boldsymbol{y})\,. \qquad (4.11)$$

Alternatively, one can obtain the subtraction term by starting with the contribution $\sigma_{\text{2v1}}$ of graph 1b to the DPS cross section and applying the additional approximations adequate for the twist-four region. Let us translate the latter from transverse momentum (where the twist-four approximations are formulated) to transverse position (where our UV regulator for the DPS cross section is local). On the l.h.s. of the relation

$$\int \frac{d^{D-2}\boldsymbol{\Delta}}{(2\pi)^{D-2}}\, F_{\text{int}}(x_i, \boldsymbol{\Delta})\, F_{\text{spl}}(\bar{x}_i, -\boldsymbol{\Delta}) = \int d^{D-2}\boldsymbol{y}\, F_{\text{int}}(x_i, \boldsymbol{y})\, F_{\text{spl}}(\bar{x}_i, \boldsymbol{y}) \qquad (4.12)$$

one neglects $\boldsymbol{\Delta}$ in the upper part of the graph, replacing $F_{\text{spl}}(\bar{x}_i, \boldsymbol{\Delta})$ by its value at $\boldsymbol{\Delta} = \boldsymbol{0}$. The integral over $\boldsymbol{\Delta}$ of $F_{\text{int}}(x_i, \boldsymbol{\Delta})$ gives the collinear twist-four distribution $G(x_1, x_2, x_2, x_1)$. In $\boldsymbol{y}$ space this corresponds to replacing $F_{\text{int}}(x_i, \boldsymbol{y})$ by its value at $\boldsymbol{y} = \boldsymbol{0}$ while retaining the $\boldsymbol{y}$ dependence of $F_{\text{spl}}(\bar{x}_i, \boldsymbol{y})$. Including the regulator function $\Phi^2(y\nu)$ under the $\boldsymbol{y}$ integral, one gets $F_{\text{int}}(x_i, \boldsymbol{y} = \boldsymbol{0}) \int d^{D-2}\boldsymbol{y}\, \Phi^2(y\nu)\, F_{\text{spl}}(\bar{x}_i, \boldsymbol{y})$. The twist-four approximation implies computing the $g \to q\bar{q}$ splitting in perturbation theory, with massless quarks and the gluons taken collinear and on shell. This corresponds to replacing $F_{\text{spl}}$ with $F_{\text{spl,pt}}$. For the graph under discussion one has $F_{\text{int}}(x_i, \boldsymbol{y} = \boldsymbol{0}) = G(x_1, x_2, x_2, x_1)$ according to (3.19) and thus obtains (4.11). We see that, as in the case of the 1v1 graph, the subtraction term is obtained from the DPS cross section by replacing the DPDs with their

small $y$ approximation at the appropriate order in $\alpha_s$. In section 5 we will find that this also holds at higher orders, where we have to take into account the second term in (3.19).

Let us finally perform the $\boldsymbol{y}$ integral in the subtraction term (4.11). Collecting all $\boldsymbol{y}$ and $\epsilon$ dependent factors in (3.12), we have

$$
\begin{aligned}
I(\mu, \nu) &= \mu^{2\epsilon} \frac{\Gamma^2(1-\epsilon)}{\pi^{1-2\epsilon}} P(\epsilon) \int \frac{d^{2-2\epsilon}\boldsymbol{y}}{y^{2-4\epsilon}} \, \Phi^2(y\nu) \\
&= \left(\frac{\mu}{\nu}\right)^{2\epsilon} \frac{\Gamma^2(1-\epsilon)}{\pi^{1-2\epsilon}} P(\epsilon) \, \Omega_{2-2\epsilon} \int_0^\infty \frac{du}{u^{1-2\epsilon}} \, \Phi^2(u) \,,
\end{aligned}
\tag{4.13}
$$

where $\Omega_{2n} = 2\pi^n/\Gamma(n)$ is the surface of a sphere in $2n$ dimensions. For brevity we do not display the momentum fraction in the splitting kernel $P$. The $\boldsymbol{y}$ integral is infrared divergent in 2 dimensions and converges for $\epsilon < 0$. Using the third row of table 1, we get in particular

$$
I(\mu, \nu) = \left[ -\frac{1}{\epsilon} - \log(4\pi) + \gamma + \log \frac{\nu^2}{\mu^2} \right] P(0) - P'(0) + \mathcal{O}(\epsilon)
\tag{4.14}
$$

for $\Phi(u) = \Theta(u - b_0)$, where $P' = \partial P/\partial \epsilon$.

The contribution of the graph to the overall cross section is finally given by

$$
\sigma_{\text{tw4}} - \sigma_{\text{2v1,pt}} + \sigma_{\text{2v1}} \,,
\tag{4.15}
$$

in analogy to (4.7). By the same argument as in the previous subsection, one finds that this combination reproduces the graph over the full range of $y$ values, with the second term cancelling against the third term for $y \sim 1/Q$ and against the first term for $y \gg 1/Q$.

As we have just seen, the product of $(\mu/\nu)^{2\epsilon}$ with the $1/\epsilon$ pole gives a logarithm $\log(\nu/\mu)$ in the subtraction term $\sigma_{\text{2v1,pt}}$. In turn, the $\boldsymbol{q}$ integral in the expression (4.10) of $\sigma_{\text{tw4}}$ gives $\log(Q/\mu)$, where $Q$ comes from the upper integration limit and the $\mu$ dependence from the regulated infrared divergence. In the combination $\sigma_{\text{tw4}} - \sigma_{\text{2v1,pt}}$ this gives $\log(Q/\nu)$. The $\nu$ dependence is cancelled by a $\log(\nu/\Lambda)$ from the $\boldsymbol{y}$ integral of the DPS cross section $\sigma_{\text{2v1}}$, where $\nu$ regulates the ultraviolet divergence. Combining all terms, one obtains $\log(Q/\Lambda)$ in (4.15). With the choice $\nu \sim Q$, the large part of the logarithm is contained only in the DPS term.

### 4.1.3 Combining contributions

We can now add all contributions to the physical cross section, including 1v1, 2v1 and 2v2 graphs in DPS. We include the same regulator $\Phi(y\nu)$ in all DPD contributions, so that the splitting and intrinsic contributions to the individual DPDs can be added up to distributions $F = F_{\text{spl}} + F_{\text{int}}$ that are defined by operator matrix elements as described in section 2.

2v2 graphs like the one in figure 1c are dominated by the DPS region, where transverse parton momenta are small, whilst regions with large transverse momenta are subleading. In impact parameter space this means that including the $\Phi$ regulator to suppress the region of small $y$ does not significantly change $\sigma_{\text{2v2}}$. We will see in section 4.3 that the effect of $\Phi$ in the 2v2 term is at the level of power corrections.

The master formula for the cross section with integrated transverse momenta is then

$$\sigma = \sigma_{\text{SPS}} - \sigma_{\text{1v1,pt}} + \sigma_{\text{tw4}} - \sigma_{\text{2v1,pt}} + \sigma_{\text{DPS}}, \tag{4.16}$$

where $\sigma_{\text{DPS}} = \sigma_{\text{1v1}} + \sigma_{\text{2v1}} + \sigma_{\text{2v2}}$ includes all contributions to the DPS cross section and involves the full DPDs $F$. From the discussion in the previous two subsections it follows that the $\nu$ dependence of $\sigma_{\text{DPS}}$ cancels against the one of $\sigma_{\text{1v1,pt}} + \sigma_{\text{2v1,pt}}$, up to terms of higher order in $\alpha_s$ that are beyond the accuracy of the calculation. As discussed in section 2.1 we neglect contributions involving twist-three functions in (4.16).

The terms in (4.16) are now meant to include all contributing graphs, including graphs without a double counting issue. The hard scattering cross sections needed for the DPS term are available for many final states, and $\sigma_{\text{1v1,pt}}$ can be readily obtained from existing results at leading order in $\alpha_s$. For double gauge boson production, the SPS cross section has been computed at high perturbative order, including the double box graphs as well as the first radiative corrections to them [40]. On the other hand, to the best of our knowledge, no calculations of $\sigma_{\text{tw4}}$ have been performed so far. To leading logarithmic accuracy, one can however omit the terms $\sigma_{\text{tw4}} - \sigma_{\text{2v1,pt}}$ in (4.16), because with the choice $\nu \sim Q$ the large DPS logarithm $\log(Q/\Lambda)$ generated by 2v1 graphs is entirely contained in $\sigma_{\text{DPS}}$, as discussed in the previous subsection.

## 4.2 Leading order analysis: TMD factorisation

We now show how our formalism works for TMD factorisation. As explained in section 2.1, the cross section for $q_1, q_2 \ll Q$ receives leading-power contributions from DPS, SPS and from their interference. The 2v1 contribution is power suppressed in this case and will not be discussed further in this subsection. The discussion of leading momentum regions is a bit more involved now, because the two quark loops in the double box graph can be collinear or hard independently of each other.

We start with the SPS/DPS interference, for which a graph is shown in figure 2b. Its contribution to the $V_1 V_2$ production cross section has the form

$$\frac{d\sigma_{\text{DPS/SPS}}}{dx_1\, dx_2\, d\bar{x}_1\, d\bar{x}_2\, d^2\boldsymbol{q}_1\, d^2\boldsymbol{q}_2} \propto \sum_{\substack{\alpha_1,\alpha_2,\alpha_0 \\ \beta_1,\beta_2,\beta_0}} H_{\alpha_1\beta_1}\, H_{\alpha_2\beta_2}\, H^*_{\alpha_0\beta_0} \int d^2\boldsymbol{Z}\, d^2\boldsymbol{y}_+\; e^{-i(\boldsymbol{q}_1+\boldsymbol{q}_2)\boldsymbol{Z} - i\boldsymbol{q}\boldsymbol{y}_+}$$

$$\times\, \Phi(y_+\nu)\, D_{\beta_1\beta_2|\beta_0}(\bar{x}_i, \boldsymbol{y}_+, \boldsymbol{Z})\, D_{\alpha_1\alpha_2|\alpha_0}(x_i, \boldsymbol{y}_+, \boldsymbol{Z}),$$

$$\frac{d\sigma_{\text{SPS/DPS}}}{dx_1\, dx_2\, d\bar{x}_1\, d\bar{x}_2\, d^2\boldsymbol{q}_1\, d^2\boldsymbol{q}_2} \propto \sum_{\substack{\alpha_0,\alpha_1,\alpha_2 \\ \beta_0,\beta_1,\beta_2}} H_{\alpha_0\beta_0}\, H^*_{\alpha_1\beta_1}\, H^*_{\alpha_2\beta_2} \int d^2\boldsymbol{Z}\, d^2\boldsymbol{y}_-\; e^{-i(\boldsymbol{q}_1+\boldsymbol{q}_2)\boldsymbol{Z} + i\boldsymbol{q}\boldsymbol{y}_-}$$

$$\times\, \Phi(y_-\nu)\, D_{\beta_0|\beta_1\beta_2}(\bar{x}_i, \boldsymbol{y}_-, \boldsymbol{Z})\, D_{\alpha_0|\alpha_1\alpha_2}(x_i, \boldsymbol{y}_-, \boldsymbol{Z}) \tag{4.17}$$

for DPS in the amplitude or in its complex conjugate. Here $D$ are the twist-three TMDs introduced in section 3.1 and the proportionality is up to kinematic and numerical factors. We have different hard scattering amplitudes $H$, with $\alpha_0$ and $\beta_0$ being gluons and the remaining partons quarks or antiquarks . The regions of small $y_+$ or $y_-$ are regulated by $\Phi(y_+\nu)$ and $\Phi(y_-\nu)$.

Notice that the graph in figure 2b has a leading contribution when the $g \to q\bar{q}$ splittings in the conjugate amplitude become collinear, in which case we are back in the DPS region. From the first expression in (4.17) one should therefore subtract

$$
\frac{d\sigma_{\mathrm{DPS},\, y_- \to 0}}{dx_1\, dx_2\, d\bar{x}_1\, d\bar{x}_2\, d^2\boldsymbol{q}_1\, d^2\boldsymbol{q}_2} \;\propto\; \sum_{\substack{\alpha_1, \alpha_2 \\ \beta_1, \beta_2}} H_{\alpha_1 \beta_1}\, H_{\alpha_2 \beta_2}\, H^*_{\alpha_1 \beta_1}\, H^*_{\alpha_2 \beta_2}
$$

$$
\times \int d^2\boldsymbol{Z}\, d^2\boldsymbol{y}_+\, d^2\boldsymbol{y}_-\; e^{-i(\boldsymbol{q}_1 + \boldsymbol{q}_2)\boldsymbol{Z} - i\boldsymbol{q}(\boldsymbol{y}_+ - \boldsymbol{y}_-)}\, \Phi(y_+ \nu)\, \Phi(y_- \nu)
$$

$$
\times\, F_{\beta_1 \beta_2,\, y_- \to 0}(\bar{x}_i, \boldsymbol{y}_\pm, \boldsymbol{Z})\, F_{\alpha_1 \alpha_2,\, y_- \to 0}(x_i, \boldsymbol{y}_\pm, \boldsymbol{Z}) \tag{4.18}
$$

with $F_{y_- \to 0}$ given in (3.8) and represented in figure 5b. From the second expression in (4.17) one should subtract an analogous term with $y_+ \to 0$.

According to (3.7), the computation of the subtraction terms involves the integrals

$$
\int_0^\infty \frac{dy}{y}\, J_n(yq)\, \Phi(y\nu) = \int_0^\infty \frac{du}{u}\, J_n(uq/\nu)\, \Phi(u) \qquad \text{with } n = 0, 2, \tag{4.19}
$$

which are given in the appendix. We see in table 1 that for $n = 0$ both our choices for $\Phi$ give a logarithmic behaviour $\log(\nu/q)$ when $\nu \gg q$, as expected for a regulated logarithmically divergent integral.

Let us now turn to the double box graph. It has leading contributions when the loops are both hard (SPS region) or both collinear (DPS region), or when one loop is hard and the other one collinear (region of SPS/DPS interference). The terms to be added to the SPS cross section for removing the contribution from the SPS/DPS interference regions are

$$
- \left[ \frac{d\sigma_{\mathrm{DPS/SPS},\, y_+ \to 0}}{dx_1\, dx_2\, d\bar{x}_1\, d\bar{x}_2\, d^2\boldsymbol{q}_1\, d^2\boldsymbol{q}_2} - \frac{d\sigma_{\mathrm{DPS},\, y_\pm \to 0}}{dx_1\, dx_2\, d\bar{x}_1\, d\bar{x}_2\, d^2\boldsymbol{q}_1\, d^2\boldsymbol{q}_2} \right]
$$

$$
- \left[ \frac{d\sigma_{\mathrm{SPS/DPS},\, y_- \to 0}}{dx_1\, dx_2\, d\bar{x}_1\, d\bar{x}_2\, d^2\boldsymbol{q}_1\, d^2\boldsymbol{q}_2} - \frac{d\sigma_{\mathrm{DPS},\, y_\pm \to 0}}{dx_1\, dx_2\, d\bar{x}_1\, d\bar{x}_2\, d^2\boldsymbol{q}_1\, d^2\boldsymbol{q}_2} \right]. \tag{4.20}
$$

The first terms in the square brackets are obtained from (4.17) by taking the perturbative splitting form for all distributions, the one for $D_{\alpha_1 \alpha_2 | \alpha_0}(x_i, \boldsymbol{y}_+, \boldsymbol{Z})$ being given in (3.10). The second terms are obtained from (4.18) and its analogue for $y_+ \to 0$ by taking both splitting vertices perturbative, using the DPD given in (3.11). Notice that the subtraction terms (4.20) contain subtractions themselves, showing the recursive nature of the construction in (4.1). Finally, one should remove the contribution from the DPS region by adding

$$
- \frac{d\sigma_{\mathrm{DPS},\, y_\pm \to 0}}{dx_1\, dx_2\, d\bar{x}_1\, d\bar{x}_2\, d^2\boldsymbol{q}_1\, d^2\boldsymbol{q}_2}\,, \tag{4.21}
$$

where again both splitting vertices in each DPD are taken as perturbative. Note that in both (4.20) and (4.21) all proton matrix elements are expressed in terms of ordinary twist-two TMDs, as is the case for the SPS cross section.

Adding up contributions and double counting subtractions, we finally obtain

$$\sigma = \sigma_{\text{DPS}} + \big[\sigma_{\text{DPS/SPS}} - \sigma_{\text{DPS},\,y_-\to 0} + \sigma_{\text{SPS/DPS}} - \sigma_{\text{DPS},\,y_+\to 0}\big]$$
$$+ \big[\sigma_{\text{SPS}} - \sigma_{\text{DPS/SPS},\,y_+\to 0} - \sigma_{\text{SPS/DPS},\,y_-\to 0} + \sigma_{\text{DPS},\,y_\pm\to 0}\big], \qquad (4.22)$$

where we have again omitted differentials $dx_1\,dx_2\,\ldots$ for brevity. The DPS cross section $\sigma_{\text{DPS}}$ contains contributions with and without splitting in each of the DPDs, which are defined by operator matrix elements as in the case of collinear factorisation.

Let us now discuss the large logarithms appearing in the different terms. The hard scattering kernel $H_{\alpha_0\beta_0}$ in (4.17) has a DPS logarithm $\log(Q/q)$, and correspondingly there is a $\log^2(Q/q)$ in the unsubtracted SPS cross section. We take $\nu \sim Q$, so that these logarithms are fully removed by the double counting subtractions. They reappear in the logarithmic integrals over $\boldsymbol{y}_+$ and $\boldsymbol{y}_-$ in the DPS cross section and in the SPS/DPS interference terms (4.17). In the full cross section (4.22) one then has a squared logarithm $\log^2(Q/q)$ in the DPS term, and a single $\log(Q/q)$ in the subtracted SPS/DPS interference in the first line, whereas the subtracted SPS contribution in the last line has no leading logarithm. If $q$ is in the nonperturbative region, one has $\log(Q/\Lambda)$ instead of $\log(Q/q)$.

At a practical level, the computation of the SPS/DPS interference requires twist-three TMDs as input, and the DPS term requires transverse-momentum dependent DPDs. In general, the modelling of these functions is very difficult and largely unconstrained. The situation is much better when $q_1, q_2 \gg \Lambda$: in this case one can compute these functions in terms of collinear ones [21].

On the perturbative side, the calculation of the hard scattering cross sections in the DPS term is easiest, while the SPS and interference terms require the full computation of box graphs. If one is satisfied with keeping only the leading double logarithm in $Q/q$, then it is sufficient to retain only the DPS contribution with $\nu \sim Q$.

## 4.3 Subtraction formalism in momentum space

In [18] it was proven that the subtraction formalism leading to (4.1) correctly approximates a graph up to power suppressed terms, provided that each approximant $T_R$ correctly reproduces the graph in its design region $R$, up to power suppressed terms. This proof works with regions in momentum space. However, our UV regulator $\Phi$ for DPS singularities is multiplicative in $\boldsymbol{y}$ space, which gives convolutions in transverse momenta that do not appear in the original Feynman graphs and hence are not part of the proof in [18]. We now show that this does not cause problems, so that our setup gives adequate approximations when transformed to momentum space, with corrections that are subleading in powers of $\Lambda/\nu$. With $\nu \sim Q$ these corrections are of the same order as other approximations in the factorisation formula.

We begin with the case of collinear factorisation, using dimensional regularisation for UV and IR divergences as usual. As mentioned earlier, we choose $\Phi$ such that the Fourier transform

$$\Psi(\boldsymbol{\Delta}, \nu) = \int \frac{d^{D-2}\boldsymbol{y}}{(2\pi)^{D-2}}\, e^{-i\boldsymbol{y}\boldsymbol{\Delta}}\big[1 - \Phi^2(y\nu)\big] = \frac{1}{\nu^{D-2}} \int \frac{d^{D-2}\boldsymbol{u}}{(2\pi)^{D-2}}\, e^{-i\boldsymbol{u}\boldsymbol{\Delta}/\nu}\big[1 - \Phi^2(u)\big].$$
$$(4.23)$$

is finite for all values of $\Delta/\nu$. In the following we show that terms that depend on $\Psi$ in the cross section are power suppressed and can hence be discarded when establishing the validity of the subtraction formalism.

Let us first discuss $\sigma_{2v2}$, which is proportional to

$$(2\pi)^{D-2} \int d^{D-2}\boldsymbol{y}\ \Phi^2(y\nu)\, F_{\text{int}}(\boldsymbol{y})\, F_{\text{int}}(\boldsymbol{y}) = \int d^{D-2}\boldsymbol{\Delta}\, F_{\text{int}}(\boldsymbol{\Delta}) F_{\text{int}}(-\boldsymbol{\Delta})$$
$$- \int d^{D-2}\boldsymbol{\Delta}\, d^{D-2}\boldsymbol{\Delta}'\ \Psi(\boldsymbol{\Delta}-\boldsymbol{\Delta}')\, F_{\text{int}}(\boldsymbol{\Delta})\, F_{\text{int}}(-\boldsymbol{\Delta}')\,. \tag{4.24}$$

The product $\hat{\sigma}_1 \hat{\sigma}_2$ of hard scattering cross sections does not affect our argument and is hence omitted. To keep the notation simple, we only display transverse-momentum arguments in this subsection. For simplicity we perform the power counting for $D = 4$ dimensions; changes in $D = 4 - 2\epsilon$ are by fractional powers and do not alter our conclusions. We then have

$$\Psi(\boldsymbol{\Delta}) \sim 1/\nu^2\,, \qquad\qquad F_{\text{int}}(\boldsymbol{\Delta}) \sim \Lambda^2/\Delta^2 \tag{4.25}$$

for both $\Delta \sim \Lambda$ and $\Delta \sim \nu$, which are the two scales present in (4.24).[5] The large $\Delta$ behaviour of $F_{\text{int}}$ follows from dimensional analysis ($\Lambda^2$ in the numerator comes from a collinear twist-four matrix element and $1/\Delta^2$ from the hard splitting kernel in the Fourier transformed version of (3.19)). The $\Psi$ independent term in (4.24) receives its leading contribution of order $\Lambda^2$ from the region $\Delta \sim \Lambda$, which is the design region of the DPS approximation. By contrast, the $\Psi$ dependent term behaves like $\Lambda^4/\nu^2$, with contributions from both $\Delta \sim \Lambda$ and $\Delta \sim \nu$. We thus see that our UV regulator (which is not necessary in $\sigma_{2v2}$ as noted earlier) only gives changes suppressed by $\Lambda^2/\nu^2$. For $\nu \sim Q$ this does not degrade the overall accuracy of the calculation.

Next we consider the contribution $\sigma_{2v1} - \sigma_{2v1,\text{pt}} + \sigma_{\text{tw4}}$ from 2v1 graphs. Here we have a subtraction term, which is obtained by replacing the DPDs in $\boldsymbol{y}$ space with their perturbative approximations. In momentum space this reads

$$F_{\text{spl}}(\boldsymbol{\Delta}) \to F_{\text{spl,pt}}(\boldsymbol{\Delta})$$
$$F_{\text{int}}(\boldsymbol{\Delta}) \to F_{\text{int,pt}}(\boldsymbol{\Delta}) + \delta^{(D-2)}(\boldsymbol{\Delta}) \int d^{D-2}\boldsymbol{\Delta}'\left[ F_{\text{int}}(\boldsymbol{\Delta}') - F_{\text{int,pt}}(\boldsymbol{\Delta}') \right]. \tag{4.26}$$

The second line is obtained by replacing $F$ with $F_{\text{int}}$ in the first term and with $F_{\text{int,pt}}$ in the second term on the r.h.s. of (3.17) and then taking the Fourier transform with respect

---

[5]In the ultraviolet region $\Delta \gg \nu$ there is a further suppression of $\Psi(\boldsymbol{\Delta})$ compared with $1/\nu^2$, the exact form of which depends on the function $\Phi$.

to $\boldsymbol{y}$. We thus find that $\sigma_{2v1} - \sigma_{2v1,pt}$ is proportional to

$$(2\pi)^{D-2} \int d^{D-2}\boldsymbol{y} \; \Phi^2(y\nu) \left[ F_{spl}(\boldsymbol{y}) F_{int}(\boldsymbol{y}) - F_{spl,pt}(\boldsymbol{y}) F_{int,pt}(\boldsymbol{y}) \right]$$

$$= \int d^{D-2}\boldsymbol{\Delta} \left[ F_{spl}(\boldsymbol{\Delta}) F_{int}(-\boldsymbol{\Delta}) - F_{spl,pt}(\boldsymbol{\Delta}) F_{int,pt}(-\boldsymbol{\Delta}) \right]$$

$$- F_{spl,pt}(\boldsymbol{\Delta}{=}\boldsymbol{0}) \int d^{D-2}\boldsymbol{\Delta}' \left[ F_{int}(\boldsymbol{\Delta}') - F_{int,pt}(\boldsymbol{\Delta}') \right]$$

$$- \int d^{D-2}\boldsymbol{\Delta} \, d^{D-2}\boldsymbol{\Delta}' \, \Psi(\boldsymbol{\Delta} - \boldsymbol{\Delta}') \left[ F_{spl}(\boldsymbol{\Delta}) F_{int}(-\boldsymbol{\Delta}') - F_{spl,pt}(\boldsymbol{\Delta}) F_{int,pt}(-\boldsymbol{\Delta}') \right]$$

$$+ \int d^{D-2}\boldsymbol{\Delta} \, \Psi(\boldsymbol{\Delta}) F_{spl,pt}(\boldsymbol{\Delta}) \int d^{D-2}\boldsymbol{\Delta}' \left[ F_{int}(\boldsymbol{\Delta}') - F_{int,pt}(\boldsymbol{\Delta}') \right]. \tag{4.27}$$

For power counting we use (4.25) and additionally

$$F_{spl}(\boldsymbol{\Delta}) \sim \Delta^0, \tag{4.28}$$

which follows from dimensional analysis of the splitting graph. The perturbative approximations $F_{int,pt}$ and $F_{spl,pt}$ scale like their un-approximated counterparts. The $\Psi$ independent terms in (4.27) are then found to be of order $\Lambda^2$, which after multiplication with $\hat{\sigma}_1 \hat{\sigma}_2$ gives the same power behaviour in the cross section as $\sigma_{tw4}$. To analyse the $\Psi$ dependent terms, we rearrange them as follows:

$$- \int d^{D-2}\boldsymbol{\Delta} \, F_{spl,pt}(\boldsymbol{\Delta}) \int d^{D-2}\boldsymbol{\Delta}' \left[ \Psi(\boldsymbol{\Delta} - \boldsymbol{\Delta}') - \Psi(\boldsymbol{\Delta}) \right] \left[ F_{int}(-\boldsymbol{\Delta}') - F_{int,pt}(-\boldsymbol{\Delta}') \right]$$

$$- \int d^{D-2}\boldsymbol{\Delta} \left[ F_{spl}(\boldsymbol{\Delta}) - F_{spl,pt}(\boldsymbol{\Delta}) \right] \int d^{D-2}\boldsymbol{\Delta}' \, \Psi(\boldsymbol{\Delta} - \boldsymbol{\Delta}') F_{int}(-\boldsymbol{\Delta}'). \tag{4.29}$$

The first line is suppressed by the integration phase space for $\Delta \sim \Lambda$, by $\Psi(\boldsymbol{\Delta} - \boldsymbol{\Delta}') - \Psi(\boldsymbol{\Delta})$ for $\Delta \sim \nu$ and $\Delta' \sim \Lambda$, and by $F_{int}(-\boldsymbol{\Delta}') - F_{int,pt}(-\boldsymbol{\Delta}')$ for $\Delta, \Delta' \sim \nu$. The second line is again suppressed by the integration phase space for $\Delta \sim \Lambda$, and by $F_{spl}(\boldsymbol{\Delta}) - F_{spl,pt}(\boldsymbol{\Delta})$ for $\Delta \sim \nu$. The sum of all $\Psi$ dependent terms is hence power suppressed compared with the $\Psi$ independent terms. One readily finds that the relative suppression is by $\Lambda^2/\nu^2$, as it was in the 2v2 case. Notice that this suppression is obtained for the combination $\sigma_{2v1} - \sigma_{2v1,pt}$ but not for the two terms separately.

Finally, we discuss the contribution $\sigma_{1v1} - \sigma_{1v1,pt} + \sigma_{SPS}$ from 1v1 graphs. The combination $\sigma_{1v1} - \sigma_{1v1,pt}$ is proportional to

$$(2\pi)^{D-2} \int d^{D-2}\boldsymbol{y} \; \Phi^2(y\nu) \left[ F_{spl}(\boldsymbol{y}) F_{spl}(\boldsymbol{y}) - F_{spl,pt}(\boldsymbol{y}) F_{spl,pt}(\boldsymbol{y}) \right]$$

$$= \int d^{D-2}\boldsymbol{\Delta} \left[ F_{spl}(\boldsymbol{\Delta}) F_{spl}(-\boldsymbol{\Delta}) - F_{spl,pt}(\boldsymbol{\Delta}) F_{spl,pt}(-\boldsymbol{\Delta}) \right]$$

$$- \int d^{D-2}\boldsymbol{\Delta} \, d^{D-2}\boldsymbol{\Delta}' \, \Psi(\boldsymbol{\Delta} - \boldsymbol{\Delta}')$$

$$\times \left[ F_{spl}(\boldsymbol{\Delta}) F_{spl}(-\boldsymbol{\Delta}') - F_{spl,pt}(\boldsymbol{\Delta}) F_{spl,pt}(-\boldsymbol{\Delta}') \right]. \tag{4.30}$$

One finds that both the $\Psi$ dependent and the $\Psi$ independent terms on the r.h.s. scale like $\Lambda^2$. Multiplied with $\hat{\sigma}_1 \hat{\sigma}_2$ this is power suppressed by $\Lambda^2/Q^2$ compared with $\sigma_{SPS}$. Notice

that the SPS approximation itself has relative corrections of order $\Lambda^2/Q^2$. This means that the contribution in (4.30) is beyond the accuracy of the calculation of 1v1 graphs. The fact that the $\Psi$ dependent part is of similar size as the $\Psi$ independent one is hence of no concern. We recall from section 2.1 that the rationale to evaluate the 2v2 and 2v1 terms, which are also suppressed by $\Lambda^2/Q^2$, is that they may be enhanced by other factors.

The case of TMD factorisation can be discussed along the same lines. In analogy to (4.23) we define $\Psi$ as the Fourier transform of $[1 - \Phi(y\nu)]$ instead of $[1 - \Phi^2(y\nu)]$. Let us discuss the SPS/DPS interference graph in figure 2b. Its contribution to $\sigma_{\mathrm{DPS/SPS}}$ in (4.17) is proportional to

$$
\int d^2\boldsymbol{Z}\, d^2\boldsymbol{y}_+ \; e^{-i(\boldsymbol{q}_1+\boldsymbol{q}_2)\boldsymbol{Z}-i\boldsymbol{q}\boldsymbol{y}_+}\; \Phi(y_+\nu)\, D(\boldsymbol{y}_+,\boldsymbol{Z})\, D(\boldsymbol{y}_+,\boldsymbol{Z})
$$

$$
= \int d^{D-2}\boldsymbol{K}\, d^{D-2}\boldsymbol{k}_-\; D(\boldsymbol{k}_-,\boldsymbol{K})\, D(\boldsymbol{q}-\boldsymbol{k}_-,\boldsymbol{q}_1+\boldsymbol{q}_2-\boldsymbol{K})
$$

$$
- \int d^{D-2}\boldsymbol{K}\, d^{D-2}\boldsymbol{k}_-\, d^{D-2}\boldsymbol{k}'_-\; \Psi(\boldsymbol{k}_- - \boldsymbol{k}'_-)
$$

$$
\times D(\boldsymbol{k}_-,\boldsymbol{K})\, D(\boldsymbol{q}-\boldsymbol{k}'_-,\boldsymbol{q}_1+\boldsymbol{q}_2-\boldsymbol{K})\,, \tag{4.31}
$$

where on the r.h.s. we have twist-three TMDs in momentum representation. An analogous relation holds for the contribution to $\sigma_{\mathrm{DPS/SPS},\,y_+\to 0}$, which according to (4.22) is to be subtracted from $\sigma_{\mathrm{DPS/SPS}}$ in the overall cross section. In this case one should take the perturbative splitting approximation $D_{\mathrm{spl,pt}}$ of the distributions, which in $\boldsymbol{y}_+$ representation is given by (3.10). For power counting we take $q_1, q_2 \sim \Lambda$ and use

$$
D(\boldsymbol{k}_-,\boldsymbol{K}) \sim 1/\big(k_- K^2\big)\,, \tag{4.32}
$$

where $k_-$ and $K$ can be of order $\Lambda$ or $\nu$. This can be derived by a general analysis of graphs as in section 5.2 of [5]; the $1/k_-$ dependence is also directly obtained by Fourier transforming (3.10). $D_{\mathrm{spl,pt}}$ behaves like $D$, and of course we have $\Psi \sim 1/\nu^2$ as before.

We see that for $k_- \sim K \sim \Lambda$ the $\Psi$ independent term in (4.31) is of order $1/\Lambda^2$, whilst the $\Psi$ dependent one is of order $1/\nu^2$. The only region in which the $\Psi$ dependent term gives a leading contribution is when $k_- \sim k'_- \sim \nu$ and $K \sim \Lambda$. However, in this region one has $D \approx D_{\mathrm{spl,pt}}$, so that its contribution is suppressed in the combination $\sigma_{\mathrm{DPS/SPS}} - \sigma_{\mathrm{DPS/SPS},\,y_+\to 0}$. We thus find that in the overall cross section, $\Psi$ dependent terms are power suppressed, as in the case of collinear factorisation. The same holds of course for $\sigma_{\mathrm{SPS/DPS}} - \sigma_{\mathrm{SPS/DPS},\,y_-\to 0}$. Finally, a corresponding analysis (involving the momenta $\boldsymbol{k}_+$, $\boldsymbol{k}_-$ and $\boldsymbol{K}$) can be given for the combination $\sigma_{\mathrm{DPS}} - \sigma_{\mathrm{DPS},\,y_-\to 0} - \sigma_{\mathrm{DPS},\,y_+\to 0} + \sigma_{\mathrm{DPS},\,y_\pm\to 0}$, which covers all terms in (4.22) that depend on our UV regulator.

## 5 Subtraction terms at higher orders

Our construction of the double counting subtractions for DPS regions is not limited to the leading-order graphs discussed so far. In the present section we illustrate how the formalism works at highers in $\alpha_s$, taking as an example the 2v1 mechanism in collinear factorisation. In the master formula (4.1) we encounter nested subtractions in this case.
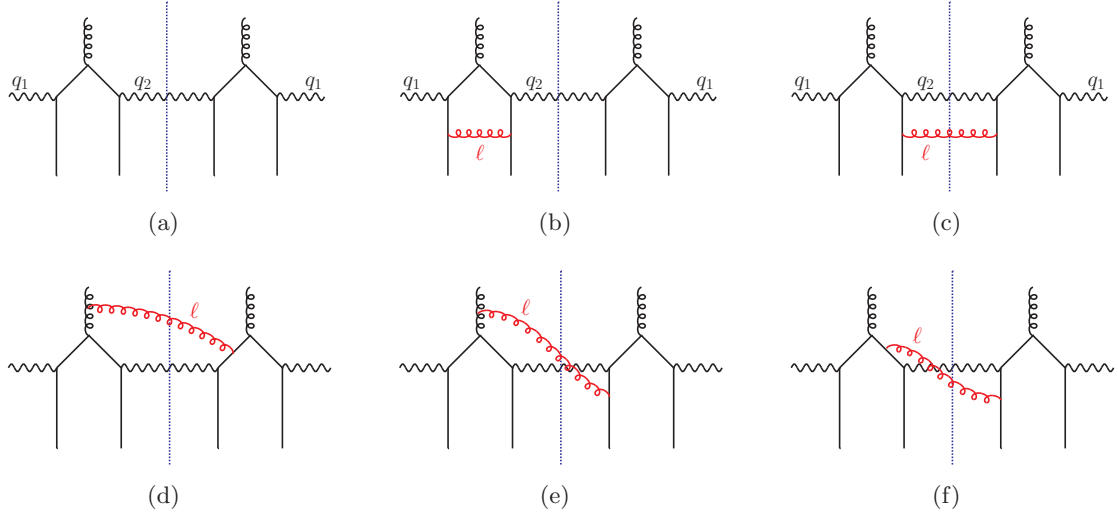
**Figure 8**. LO and NLO graphs for the 2v1 mechanism, with the blobs for the hadronic matrix elements omitted as in figure 7.

We show for selected graphs how the formalism gives the different contributions to the cross section, namely $\sigma_{\mathrm{tw}4}$, $\sigma_{2\mathrm{v}1}$ and the subtraction term $\sigma_{2\mathrm{v}1,\mathrm{pt}}$. In particular we find that the latter can be obtained from $\sigma_{2\mathrm{v}1}$ by replacing the DPDs with their short-distance expansions $F_{\mathrm{int},\mathrm{pt}}$ and $F_{\mathrm{spl},\mathrm{pt}}$, as found at leading order in section 4.1.2.

We focus here on the general structure and use a schematic notation, in particular suppressing indices and arguments that are not essential for the discussion.

**Graph 8a.** To introduce this notation, let us briefly review the leading-order graph in figure 8a, which was discussed in detail in section 4.1.2. It has two leading momentum regions, which we denote by $C$ and $H$. In region $C$ the gauge bosons have small transverse momenta and the $g \to q\bar{q}$ splitting is near collinear. This is the region of DPS, and its approximated contribution to the cross section reads

$$C_C \Gamma^a = \hat{\sigma}_1^{(0)} \hat{\sigma}_2^{(0)} \int_y \Phi^2(y\nu) \, F_{\mathrm{spl}}^{(1)}(y) \, F_{\mathrm{int}}^{(0)}(y) \,. \tag{5.1}$$

Here $\hat{\sigma}_i^{(0)}$ is the tree-level cross section for $q\bar{q} \to V_i$, and we write $\int_y$ as a shorthand for $\int d^{D-2}y$. The contribution of this graph to the splitting DPD at the top is denoted by $F_{\mathrm{spl}}^{(0)}$, and its contribution to the intrinsic DPD at the bottom by $F_{\mathrm{int}}^{(0)}$. Here and in the rest of the section, superscripts $(0), (1), (2)$ indicate the power of $\alpha_s$ in the considered quantity. For simplicity we suppress longitudinal momentum arguments throughout.

In region $H$, the transverse momenta of the gauge bosons are large and thus the $q\bar{q}$ pair into which the gluon splits is far off shell on both sides of the final state cut. This is described by the twist-four contribution and requires a subtraction for the smaller region $C$. The subtracted contribution is

$$C_H \Gamma^a = \int_q f^{(0)} \, H^{(1)} \otimes G^{(0)} - \hat{\sigma}_1^{(0)} \hat{\sigma}_2^{(0)} \int_y \Phi^2(y\nu) \, f^{(0)} \, V^{(1)}(y) \, G^{(0)} \,. \tag{5.2}$$

Here $\int_q$ is shorthand for $\int d^{D-2}\boldsymbol{q}$, the tree-level hard scattering cross section for $q\bar{q} + g \to V_1 V_2$ is denoted by $H^{(1)}$, whilst $f^{(0)}$ and $G^{(0)}$ respectively are the zeroth order contributions to the gluon distribution at the top and the twist-four $q\bar{q}$ distribution at the bottom of the graph. As seen in (4.8), there is a convolution integral over longitudinal momentum fractions in $G$, denoted here by $\otimes$. In the subtraction term, $V^{(1)}$ represents the lowest-order perturbative $g \to q\bar{q}$ splitting kernel, including all terms in (3.12) except for the single-parton density $f_{a_0}$. In this term, the momentum fractions of $G^{(0)}$ are fixed by external kinematics, so that the convolution $\otimes$ is absent.

The perturbative expressions of the DPDs for small $y$ are

$$F_{\text{spl,pt}}^{(1)}(y) = f^{(0)} V^{(1)}(y), \tag{5.3}$$

$$F_{\text{int,pt}}^{(0)}(y) = G^{(0)} \tag{5.4}$$

at leading order, so that the combined contribution of the $C$ and $H$ regions can be written as

$$\Gamma^a = \int_q f^{(0)} H^{(1)} \otimes G^{(0)} - \hat{\sigma}_1^{(0)} \hat{\sigma}_2^{(0)} \int_y \Phi^2(y\nu) F_{\text{spl,pt}}^{(1)}(y) F_{\text{int,pt}}^{(0)}(y)$$

$$+ \hat{\sigma}_1^{(0)} \hat{\sigma}_2^{(0)} \int_y \Phi^2(y\nu) F_{\text{spl}}^{(1)}(y) F_{\text{int}}^{(0)}(y). \tag{5.5}$$

As already mentioned, the double counting subtraction in the second term is simply obtained from the DPS term by replacing each DPD with its small-$y$ expansion at the appropriate order in $\alpha_s$.

**Graph 8b.** We now analyse the NLO graph in figure 8b. The $q\bar{q}$ splitting can be hard or collinear, and likewise the additional gluon with momentum $\ell$. The leading momentum regions thus are $CC$, $HC$ and $HH$, where the first letter refers to the $q\bar{q}$ splitting and the second to the additional gluon. There is also a region where the gluon is hard whereas the $q\bar{q}$ splitting to the right of the cut is collinear. The splitting at the left of the cut is then hard by momentum conservation. One finds that this region is not leading: the loss of phase space for having the gauge bosons at low transverse momenta is stronger than the gain from having one but not both splitting vertices collinear. Let us discuss the leading regions in turn.

In the $CC$ region, the extra gluon is part of the DPD for the right moving particles, and one simply has

$$C_{CC}\Gamma^b = \hat{\sigma}_1^{(0)} \hat{\sigma}_2^{(0)} \int_y \Phi^2(y\nu) F_{\text{spl}}^{(1)}(y) F_{\text{int}}^{(1)}(y), \tag{5.6}$$

where the superscript (1) indicates a contribution to $F_{\text{int}}$ at order $\alpha_s$. Likewise, in the $HC$ region, the extra gluon becomes part of the twist-four distribution $G$, so that one has

$$C_{HC}\Gamma^b = \int_q f^{(0)} H^{(1)} \otimes G_{\text{ren}}^{(1)} - \hat{\sigma}_1^{(0)} \hat{\sigma}_2^{(0)} \int_y \Phi^2(y\nu) f^{(0)} V^{(1)}(y) G_{\text{ren}}^{(1)}. \tag{5.7}$$

Note that the integral over the gluon momentum in the graph giving $G_{\mathrm{ren}}^{(1)}$ contains an ultraviolet divergence. It is understood that $G_{\mathrm{ren}}^{(1)}$ is renormalised, i.e. includes the appropriate UV counterterm, which is indicated by the subscript "ren".

For the $HH$ region, the recursive character of the subtraction formalism becomes evident. We have a subtraction for the next smallest region $HC$, which itself contains a subtraction for the region $CC$. In addition, a $CC$ subtraction has to be made for the overall graph. This gives

$$
\begin{aligned}
C_{HH}\Gamma^b = & \int_q f^{(0)}\, H^{(2)} \otimes G^{(0)} \\
& - \left[ \int_q f^{(0)}\, H^{(1)} \otimes \widetilde{K}_{\mathrm{ren}}^{(1)}(0) - \hat{\sigma}_1^{(0)}\hat{\sigma}_2^{(0)} \int_y \Phi^2(y\nu)\, f^{(0)}\, V^{(1)}(y)\, \widetilde{K}_{\mathrm{ren}}^{(1)}(0) \right] \otimes G^{(0)} \\
& - \hat{\sigma}_1^{(0)}\hat{\sigma}_2^{(0)} \int_y \Phi^2(y\nu)\, f^{(0)}\, V^{(1)}(y)\, \widetilde{K}^{(1)}(y) \otimes G^{(0)} \,.
\end{aligned}
\tag{5.8}
$$

The term in the first line is for the hard region, with $H^{(2)}$ being a one-loop contribution to the cross section for $q\bar{q} + g \to V_1 V_2$. In the second line, the term for the $HC$ region (including its subtraction for the $CC$ region) involves a non-diagonal splitting kernel $\widetilde{K}_{\mathrm{ren}}^{(1)}(0)$ for the exchange of a gluon between the two parton lines with momentum fractions $x_1$ and $x_2$, which are taken at relative transverse position $y = 0$ because this is part of the approximation for the $HC$ region. In the last line, which is the overall subtraction for the $CC$ region, the same kernel is to be taken at relative distance $y$, because in this case the $q\bar{q}$ splitting is not assumed to be harder than the extra gluon emission. One therefore has to take a convolution in transverse momenta, which turns into a product of two $y$ dependent factors $V^{(1)}(y)$ and $\widetilde{K}^{(1)}(y)$ after Fourier transformation.

Notice that in dimensional regularisation the limit $y \to 0$ in $\widetilde{K}^{(1)}(y)$ is not smooth. The graph for $\widetilde{K}_{\mathrm{ren}}^{(1)}(0)$ contains a scaleless momentum integral, which is zero. After renormalisation, i.e. subtraction of the ultraviolet counterterm, one is left with an infrared divergence.[6] By contrast, $\widetilde{K}^{(1)}(y)$ at finite $y$ does not have an ultraviolet divergence (and hence no counterterm), but it does have the same infrared divergence as $\widetilde{K}_{\mathrm{ren}}^{(1)}(0)$.

Adding all contributions and using (5.3) we obtain

$$
\begin{aligned}
\Gamma^b = & \int_q f^{(0)} \left( H^{(1)} \otimes G_{\mathrm{ren}}^{(1)} + \left[ H^{(2)} - H^{(1)} \otimes \widetilde{K}_{\mathrm{ren}}^{(1)}(0) \right] \otimes G^{(0)} \right) \\
& - \hat{\sigma}_1^{(0)}\hat{\sigma}_2^{(0)} \int_y \Phi^2(y\nu)\, F_{\mathrm{spl,pt}}^{(1)}(y)\, F_{\mathrm{int,pt}}^{(1)}(y) \\
& + \hat{\sigma}_1^{(0)}\hat{\sigma}_2^{(0)} \int_y \Phi^2(y\nu)\, F_{\mathrm{spl}}^{(1)}(y)\, F_{\mathrm{int}}^{(1)}(y)
\end{aligned}
\tag{5.9}
$$

with

$$
F_{\mathrm{int,pt}}^{(1)}(y) \underset{\mathrm{graph}\ b}{=} G_{\mathrm{ren}}^{(1)} + \left[ \widetilde{K}^{(1)}(y) - \widetilde{K}_{\mathrm{ren}}^{(1)}(0) \right] \otimes G^{(0)} \,.
\tag{5.10}
$$

---

[6]The UV divergent part of $\widetilde{K}_{\mathrm{ren}}^{(1)}(0)$ gives DGLAP splitting kernels generalised to non-diagonal momentum fractions, known from the evolution of twist-four distributions [34] and of GPDs [41].

In the first line of (5.9) we have an NLO contribution to the twist-four term, with the NLO kernel $H^{(2)}$ having a subtraction for the region where the gluon becomes collinear, as is standard in collinear factorisation. In the DPS subtraction term we have a contribution to the small-$y$ expansion of $F_{\text{int}}$ at NLO, with the three terms in (5.10) corresponding respectively to the three terms in the general expression (3.17). The infrared divergence of the kernel $K$ drops out here, as it must. We see that the DPS subtraction term in (5.9) has the same form as in (5.5), with $F_{\text{int,pt}}^{(1)}$ instead of $F_{\text{int,pt}}^{(0)}$.

In (5.9) all ultraviolet and infrared subtractions have been carried out, so that one can set $D = 4$ in dimensional regularisation. The quantities with subscript "ren" then depend on the associated scale $\mu$. In the first line, this dependence cancels between $G_{\text{ren}}^{(1)}$ (the graph under discussion contributes to the scale evolution of $G$) and the subtracted hard scattering kernel in the second term. In (5.10), the $\mu$ dependence cancels between $G_{\text{ren}}^{(1)}$ and the subtracted kernel of the small-$y$ expansion, i.e. graph 8b does not contribute to the evolution of $F_{\text{int}}$.

**Graph 8c.** We next consider the graph in figure 8c. The contribution from the region $CC$ is

$$C_{CC}\Gamma^c = \hat{\sigma}_1^{(0)}\hat{\sigma}_2^{(0)} \int_y \Phi^2(y\nu)\, F_{\text{spl}}^{(1)}(y)\, F_{\text{int,ren}}^{(1)}(y)\,, \tag{5.11}$$

where in contrast to graph 8b the one-loop contribution to $F_{\text{int}}$ now has an ultraviolet divergence (since the two partons connected by the gluon are at the same transverse position) and needs renormalisation. The contribution from region $HC$ is analogous to (5.7). Unlike for graph 8b, a leading contribution is now also obtained for the momentum region $CH$, where the $g \to q\bar{q}$ splittings are collinear, whilst the gluon carries a large transverse momentum (balanced by the vector boson momentum $\boldsymbol{q}_2$). Including the subtraction for the region $CC$, we have

$$C_{CH}\Gamma^c = \hat{\sigma}_1^{(0)} \int_y \Phi^2(y\nu)\, F_{\text{spl}}^{(1)}(y) \left[\sigma_2^{(1)} - \sigma_2^{(0)} K_{\text{ren}}^{(1)}\right] \otimes F_{\text{int}}^{(0)}(y)\,, \tag{5.12}$$

where $\sigma_2^{(1)}$ is the cross section for $q\bar{q} \to V_2 + g$ and $K_{\text{ren}}^{(1)}$ is a diagonal splitting kernel including the necessary UV counterterm. Its divergent part involves the contribution of this graph to the usual DGLAP splitting kernel $P_{qq}$. Finally, the region $HH$ gives

$$\begin{aligned}
C_{HH}\Gamma^c = &\int_q f^{(0)}\, H^{(2)} \otimes G^{(0)} \\
&- \left[\int_q f^{(0)}\, H^{(1)} \otimes K_{\text{ren}}^{(1)} - \hat{\sigma}_1^{(0)}\hat{\sigma}_2^{(0)} \int_y \Phi^2(y\nu)\, f^{(0)}\, V^{(1)}(y)\, K_{\text{ren}}^{(1)}\right] \otimes G^{(0)} \\
&- \hat{\sigma}_1^{(0)} \int_y \Phi^2(y\nu)\, f^{(0)}\, V^{(1)}(y) \left[\sigma_2^{(1)} - \sigma_2^{(0)} K_{\text{ren}}^{(1)}\right] \otimes G^{(0)} \\
&- \hat{\sigma}_1^{(0)}\hat{\sigma}_2^{(0)} \int_y \Phi^2(y\nu)\, f^{(0)}\, V^{(1)}(y)\, K_{\text{ren}}^{(1)} \otimes G^{(0)}\,,
\end{aligned} \tag{5.13}$$

where the last three lines are the subtraction terms for the regions $HC$, $CH$, $CC$, respectively. Adding all contributions, we obtain

$$\Gamma^c = \int_q f^{(0)} \left( H^{(1)} \otimes G_{\text{ren}}^{(1)} + \left[ H^{(2)} - H^{(1)} \otimes K_{\text{ren}}^{(1)} \right] \otimes G^{(0)} \right)$$

$$- \hat{\sigma}_1^{(0)} \int_y \Phi^2(y\nu)\, F_{\text{spl,pt}}^{(1)}(y) \left( \hat{\sigma}_2^{(0)}\, F_{\text{int,pt}}^{(1)}(y) + \left[ \sigma_2^{(1)} - \sigma_2^{(0)}\, K_{\text{ren}}^{(1)} \right] \otimes F_{\text{int}}^{(0)}(y) \right)$$

$$+ \hat{\sigma}_1^{(0)} \int_y \Phi^2(y\nu)\, F_{\text{spl}}^{(1)}(y) \left( \hat{\sigma}_2^{(0)}\, F_{\text{int,ren}}^{(1)}(y) + \left[ \sigma_2^{(1)} - \sigma_2^{(0)}\, K_{\text{ren}}^{(1)} \right] \otimes F_{\text{int}}^{(0)}(y) \right), \qquad (5.14)$$

where in the second line we used (5.3), (5.4) and

$$F_{\text{int,pt}}^{(1)}(y) \underset{\text{graph } c}{=} G_{\text{ren}}^{(1)}. \qquad (5.15)$$

The latter relation holds because the exchanged gluon in graph 8c gives exactly the same contribution to $F_{\text{int}}(y)$ and to $G$ and hence cancels between the second and third term in the general small-$y$ expansion (3.17) of $F_{\text{int}}(y)$. In each line of (5.14) we have hard scattering cross sections subtracted for the collinear region, as is characteristic for collinear factorisation. The $\mu$ dependence cancels between the subtracted cross sections and the renormalised distributions.

**Graph 8d.** We now discuss figure 8d, where we have an additional gluon in the upper part of the graph. Its leading momentum regions are $CC$, $HC$ and $HH$. As in the case of graph 8b, there is no leading region $CH$. For the $CC$ region we simply have

$$C_{CC}\Gamma^d = \hat{\sigma}_1^{(0)}\hat{\sigma}_2^{(0)} \int_y \Phi^2(y\nu)\, F_{\text{spl}}^{(2)}(y)\, F_{\text{int}}^{(0)}(y). \qquad (5.16)$$

In the region $HC$, the extra gluon in the graph is collinear and couples to an off-shell quark, since the $g \to q\bar{q}$ splitting is hard. Among the approximations to be made in this case is the Grammer-Yennie approximation (see e.g. [5, 20]). After summing over an appropriate set of graphs, one can apply a Ward identity, after which the gluon couples to an eikonal line, as shown in figure 9. We thus can write

$$C_{HC}\Gamma^d + C_{HC}\Gamma^{e+} = \int_q f_{\text{ren}}^{(1)}\, H^{(1)} \otimes G^{(0)} - \hat{\sigma}_1^{(0)}\hat{\sigma}_2^{(0)} \int_y \Phi^2(y\nu)\, f_{\text{ren}}^{(1)}\, V^{(1)}(y)\, G^{(0)}, \qquad (5.17)$$

where $\Gamma^{e+}$ denotes the sum of graph e and of two graphs analogous to graphs d and e, with the gluon coupling to the rightmost quark line. $f_{\text{ren}}^{(1)}$ is an NLO contribution to the gluon density, involving a gluon coupling to an eikonal line. The integral over the gluon transverse momentum diverges in the ultraviolet and hence requires renormalisation. For the $HH$ region we have

$$C_{HH}\Gamma^d - T_{HH}\, C_{HC}\Gamma^{e+} = \int_q f^{(0)}\, H^{(2)} \otimes G^{(0)}$$

$$- \left[ \int_q f^{(0)} \otimes K_{\text{ren}}^{(1)}\, H^{(1)} \otimes G^{(0)} - \hat{\sigma}_1^{(0)}\hat{\sigma}_2^{(0)} \int_y \Phi^2(y\nu)\, f^{(0)} \otimes K_{\text{ren}}^{(1)}\, V^{(1)}(y)\, G^{(0)} \right]$$

$$- \hat{\sigma}_1^{(0)}\hat{\sigma}_2^{(0)} \int_y \Phi^2(y\nu)\, f^{(0)}\, V^{(2)}(y)\, G^{(0)}. \qquad (5.18)$$
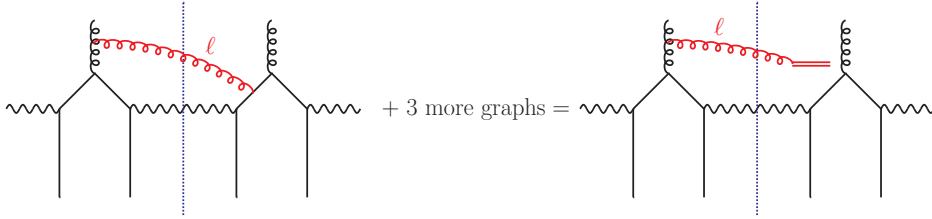
**Figure 9**. Graphical representation of applying the Grammer-Yennie approximation and a Ward identity to the sum of four graphs. The three other graphs required are graph e in figure 8, and the analogues to graphs d and e in which the gluon couples to the quark line at the very right. It is understood that the $g \to q\bar{q}$ splitting is hard, whereas the gluons are collinear left moving.

In the subtraction term for the region $HC$, we have an NLO contribution $K_{\text{ren}}^{(1)}$ to the $gg$ splitting kernel with a gluon coupling to an eikonal line, in analogy to $f_{\text{ren}}^{(1)}$ above. Again, this requires summation over several graphs as specified on the l.h.s. In the last line of (5.18) we have the subtraction term for the region $CC$, in which the gluon and the splitting are regarded as collinear, so that no eikonal approximation is made and one obtains an NLO contribution $V^{(2)}$ to the $g \to q\bar{q}$ splitting kernel. Adding all contributions, we obtain

$$\Gamma^d + (1 - T_{HH}) C_{HC} \Gamma^{e+} = \int_q \left( f_{\text{ren}}^{(1)} H^{(1)} + f^{(0)} \left[ H^{(2)} - K_{\text{ren}}^{(1)} H^{(1)} \right] \right) \otimes G^{(0)}$$
$$- \hat{\sigma}_1^{(0)} \hat{\sigma}_2^{(0)} \int_y \Phi^2(y\nu) F_{\text{spl,pt}}^{(2)}(y) F_{\text{int,pt}}^{(0)}(y)$$
$$+ \hat{\sigma}_1^{(0)} \hat{\sigma}_2^{(0)} \int_y \Phi^2(y\nu) F_{\text{spl}}^{(2)}(y) F_{\text{int}}^{(0)}(y), \qquad (5.19)$$

where

$$F_{\text{spl,pt}}^{(2)}(y) \underset{d \text{ and } e+}{=} f_{\text{ren}}^{(1)} V^{(1)}(y) + f^{(0)} \left[ V^{(2)}(y) - K_{\text{ren}}^{(1)} V^{(1)}(y) \right] \qquad (5.20)$$

is a contribution to the perturbative splitting DPD at NLO. The scale dependence of $f_{\text{ren}}^{(1)}$ and $K_{\text{ren}}^{(1)}$ cancels in (5.20), corresponding to the fact that graph 8d and the other ones required for applying the Ward identity do not contribute to the scale evolution of the DPD.

**Other NLO graphs.** The discussion of graphs where the extra gluon is exchanged between a right-moving and a left-moving line, such as in figure 8e and f, is more complicated because it involves more momentum regions. Depending on whether the $g \to q\bar{q}$ splitting is collinear or hard, the extra gluon can be hard ($H$), collinear left moving ($L$), collinear right-moving ($R$), or soft ($S$). In the three latter cases, one can apply the Grammer-Yennie approximation and, after summation over an appropriate set of graphs, apply Ward identities, similarly to the example we have just discussed. This leads to the gluon coupling to eikonal lines, which in regions $L$ and $R$ are part of the Wilson lines in the operator defining the DPDs, whereas in region $S$ they are part of the Wilson lines in the soft factor. For more details we refer to [20]. Our formulation of the DPS cross section in transverse

position space is well suited for the treatment of these Wilson lines, which are taken at the same transverse positions as the quark or gluon fields in the DPD definitions. In transverse momentum space, one instead has a more complicated structure with convolution integrals.

## 6   Double DGLAP evolution

Collinear DPDs $F(x_i, \boldsymbol{y}; \mu_i)$ evolve with $\mu_i$ according to the double DGLAP equations. In the present section, we discuss several aspects of this scale dependence.

### 6.1   Resummation of large logarithms

As we have seen in section 4.1.2, the 2v1 graph in figure 8a gives rise to a large logarithm $\log(Q/\Lambda)$ when the cross section is integrated over the transverse boson momenta. This logarithm originates from the $g \to q\bar{q}$ splitting vertices in the kinematic region of double parton scattering and is correctly reproduced in the DPS cross section, provided that we make a suitable choice of scales.

Radiative corrections to figure 8a give rise to additional large logarithms of DGLAP type. One can expect that these logarithms are correctly resummed to all orders by the DGLAP evolution of the DPDs, provided again that one makes suitable scale choices. In this section we show for explicit examples that this is indeed the case. We will allow for different hard scales $Q_1, Q_2$ and set the scales as

$$\mu_1 = Q_1, \qquad\qquad \mu_2 = Q_2, \qquad\qquad \nu = Q_{\min}, \qquad\qquad (6.1)$$

where we define

$$Q_{\min} = \min\left(Q_1, Q_2\right). \qquad\qquad (6.2)$$

Of course, our conclusions regarding large logarithms will not change if (6.1) is modified by factors of order unity. For definiteness, we will always take a rigid cutoff $\Phi(u) = \Theta(u - b_0)$ in $y\nu$. For later use we define the scale

$$\mu_y = \frac{b_0}{y}. \qquad\qquad (6.3)$$

DGLAP logarithms are generated by loop integrals in regions of strongly ordered transverse momenta. We will compare the logarithms obtained directly from transverse momentum integrals with the ones obtained from evolved DPDs in $y$ space.

To set the scene, we start with the lowest-order 2v1 graph in figure 8a. In the DPS region, i.e. when the scale of the $g \to q\bar{q}$ splittings is much smaller than $Q_1$ and $Q_2$, the behaviour of a splitting vertex is easily recovered from the computation of the splitting DPD. As can be seen from (3.4), a vertex with quark momentum $k$ gives a factor $\boldsymbol{k}^j/\boldsymbol{k}^2$, where $j$ is a transverse index. The transverse momentum of the splitting vertices must flow through the gauge boson lines. Thus, the two splitting vertices in figure 8a give a
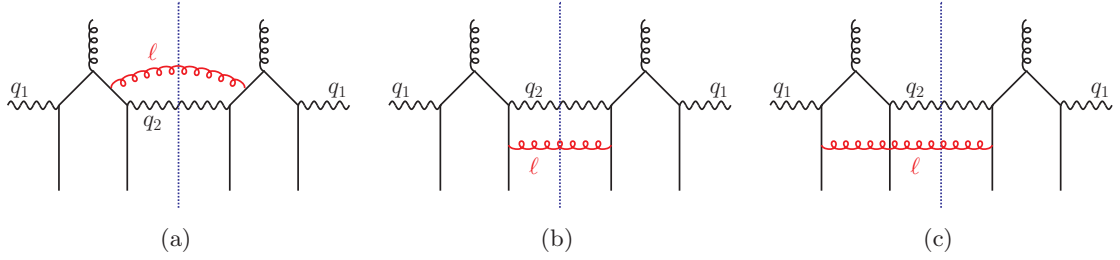
**Figure 10**. Radiative corrections to the hard scattering graph in figure 8a that give rise to DGLAP logarithms.

combined $1/q_1^2$ behaviour for $q_1 \ll Q_{\min}$, as is evident in (4.10) after angular integration. The integrated cross section is then proportional to

$$\int_{\Lambda^2}^{Q_{\min}^2} \frac{dq_1^2}{q_1^2} = \log \frac{Q_{\min}^2}{\Lambda^2} \,. \tag{6.4}$$

Since the $q_1$ integral is logarithmic, we must include an infrared cutoff $\Lambda$, which limits the integration to the perturbative region where the preceding analysis is valid. We take here the usual procedure for extracting the leading behaviour of a logarithmic integral $\int_a^b dk \, \Gamma(k)$, approximating the integrand $\Gamma(k)$ for $a \ll k \ll b$ while integrating over the full range $a \leq k \leq b$.

Let us now see how the logarithm (6.4) arises in the DPS cross section formula. According to the perturbative expressions of the DPDs, $F_{\text{spl,pt}} \sim 1/y^2$ and $F_{\text{int,pt}}$ is $y$ independent at leading order in $\alpha_s$, so that we have

$$\int_{b_0^2/\nu^2}^{b_0^2/\Lambda^2} \frac{dy^2}{y^2} = \log \frac{Q_{\min}^2}{\Lambda^2} \,. \tag{6.5}$$

With our choice for the infrared cutoff on $y$, the expressions (6.4) and (6.5) agree including their overall normalisation.

Now consider the NLO graph in figure 10a. We route the gluon momentum $\ell$ back through the boson line with momentum $q_2$, so that $\ell$ flows through exactly two quark propagators. In the strongly ordered region

$$\Lambda \ll q_1 \ll Q_1 \,, \qquad\qquad q_1 \ll \ell \ll Q_2 \tag{6.6}$$

the graph has the same DPS logarithm as the leading-order result, whilst developing an additional logarithm from the integral $\int d\ell^2/\ell^2$, as is familiar from similar gluon emission graphs in hard scattering processes. In the case $Q_2 < Q_1$ this gives

$$\int_{\Lambda^2}^{Q_2^2} \frac{dq_1^2}{q_1^2} \int_{q_1^2}^{Q_2^2} \frac{d\ell^2}{\ell^2} = \frac{1}{2} \log^2 \frac{Q_2^2}{\Lambda^2} = \frac{1}{2} \log^2 \frac{Q_{\min}^2}{\Lambda^2} \,, \tag{6.7}$$

whereas for $Q_1 < Q_2$ we have

$$\int_{\Lambda^2}^{Q_1^2} \frac{dq_1^2}{q_1^2} \int_{q_1^2}^{Q_2^2} \frac{d\ell^2}{\ell^2} = \frac{1}{2} \left[ \log^2 \frac{Q_2^2}{\Lambda^2} - \log^2 \frac{Q_2^2}{Q_1^2} \right] = \frac{1}{2} \log \frac{Q_{\min}^2}{\Lambda^2} \left[ \log \frac{Q_{\min}^2}{\Lambda^2} + 2 \log \frac{Q_2^2}{Q_1^2} \right] . \tag{6.8}$$

In both cases, we have an additional logarithm $\log(Q_{\min}^2/\Lambda^2)$ compared with the lowest-order result. If $Q_1 \ll Q_2$, then the logarithm of the ratio $Q_2/Q_1$ is also large.

Let us compare these results with the 2v1 part of the DPS cross section formula when we take into account evolution of the splitting DPD. Focusing on the dependence on $y$ and $\mu$, we can write

$$F_{\mathrm{spl,pt}}(x_i, \boldsymbol{y}; \mu_y, \mu_y) \propto \frac{1}{y^2}\, \alpha_s(\mu_y)\, g(x_1 + x_2; \mu_y) \tag{6.9}$$

at leading order, where the choice of scale $\mu_y$ ensures that there are no large logarithms of $y\mu$ in the higher-order corrections. This should be evolved to $F_{\mathrm{spl,pt}}(x_i, \boldsymbol{y}; \mu_1, \mu_2)$ in the cross section. In (6.9) there is a logarithmic dependence on $y$ from the running of $\alpha_s$ and from the scale dependence of the gluon density. Graph 10a does not contribute to either of these, so that the corresponding dependence should not be taken into account here. Rather, the DGLAP logarithm of graph 10a corresponds to the evolution of (6.9) from $\mu_y$ to $\mu_2$, which gives a factor proportional to $\alpha_s \log(\mu_2^2/\mu_y^2)$. The relevant $y$ integral in the DPD cross section thus is

$$\int_{b_0^2/\nu^2}^{b_0^2/\Lambda^2} \frac{dy^2}{y^2}\, \log \frac{y^2 \mu_2^2}{b_0^2} = \frac{1}{2}\left[ \log^2 \frac{Q_2^2}{\Lambda^2} - \log^2 \frac{Q_2^2}{Q_{\min}^2} \right], \tag{6.10}$$

where we have used the scale choice in (6.1). This agrees with both (6.7) and (6.8).

Note the ease with which (6.10) is obtained in $y$ space, compared with the detailed considerations leading to (6.7) and (6.8). We see that with evolved DPDs in $\sigma_{\mathrm{DPS}}$ one includes higher order logarithms while using tree-level partonic cross sections. Treating the 2v1 graphs fully within the collinear twist-four formalism would require the explicit inclusion of the higher-order graphs.

Whereas in the previous example, the gluon is emitted close to the splitting vertices, the opposite is true for the graphs in figure 10b and c. For figure 10b the region generating two large logarithms is

$$\Lambda \ll q_1 \ll Q_{\min}, \qquad\qquad \Lambda \ll \ell \ll Q_2, \tag{6.11}$$

which gives

$$\int_{\Lambda^2}^{Q_{\min}^2} \frac{dq_1^2}{q_1^2} \int_{\Lambda^2}^{Q_2^2} \frac{d\ell^2}{\ell^2} = \log \frac{Q_{\min}^2}{\Lambda^2}\, \log \frac{Q_2^2}{\Lambda^2}. \tag{6.12}$$

To see how these logarithms arise from the evolution of $F_{\mathrm{int,pt}}$, we use the short-distance expansion

$$F_{\mathrm{int,pt}}(x_1, x_2, \boldsymbol{y}; \mu_y, \mu_y) = G(x_1, x_2, x_2, x_1; \mu_y) \tag{6.13}$$

at leading order, which we evaluate at scale $\mu_y$ to avoid large logarithms of $y\mu$ in the higher-order corrections, as we did for $F_{\mathrm{spl,pt}}$ in (6.9). Graph 10b contributes to the scale evolution of $G$, where at leading order it gives $\alpha_s \log(\mu_y^2/\Lambda^2)$ when evolution is started from the generic soft scale $\Lambda^2$. The graph also contributes to the evolution of $F_{\mathrm{int,pt}}$ from $\mu_y$ to

$\mu_2$, where it gives $\alpha_s \log(\mu_2^2/\mu_y^2)$. The two effects add up to a factor $\alpha_s \log(\mu_2/\Lambda^2)$. In the DPS cross section we thus have

$$\int_{b_0^2/\nu^2}^{b_0^2/\Lambda^2} \frac{dy^2}{y^2} \log \frac{\mu_2^2}{\Lambda^2} = \log \frac{Q_{\min}^2}{\Lambda^2} \log \frac{Q_2^2}{\Lambda^2} \tag{6.14}$$

in agreement with (6.12).

The graph in figure 10c also has a region where DGLAP logarithms are generated, as was already discussed in [17]. We route the loop momentum $\ell$ through two quark lines at one of the two $g \to q\bar{q}$ splitting vertices. In the strongly ordered region

$$\Lambda \ll \ell \ll q_1 \ll Q_{\min} \tag{6.15}$$

the two lower fermion lines to which the gluon couples have virtualities of order $\ell^2$, whereas both splitting vertices have virtualities of order $q_1^2 \approx q_2^2$. This gives [7]

$$\int_{\Lambda^2}^{Q_{\min}^2} \frac{dq_1^2}{q_1^2} \int_{\Lambda^2}^{q_1^2} \frac{d\ell^2}{\ell^2} = \frac{1}{2} \log^2 \frac{Q_{\min}^2}{\Lambda^2} . \tag{6.16}$$

Just as graph 10b, graph c contributes to the evolution of $G$ in (6.13), giving a factor $\alpha_s \log(\mu_y^2/\Lambda^2)$ at leading order. There is however no contribution to the evolution of $F_{\text{int,pt}}$ from $\mu_y$ to $\mu_2$ from this graph, so that in the cross section we have

$$\int_{b_0^2/\nu^2}^{b_0^2/\Lambda^2} \frac{dy^2}{y^2} \log \frac{b_0^2}{\Lambda^2 y^2} = \frac{1}{2} \log^2 \frac{Q_{\min}^2}{\Lambda^2} , \tag{6.17}$$

which again reproduces the result (6.16) obtained in momentum space, including the overall factor.

A fully analogous discussion can be given for the graph in figure 8b. One finds the same type of DGLAP logarithms as in (6.16), which is again reproduced correctly in the DPS cross section formula. Note that in this case the additional gluon is a virtual correction rather than emitted into the final state – a somewhat unusual situation in the context of DGLAP logarithms.

In conclusion, higher-order graphs with strong ordering of transverse momenta produce DGLAP type logarithms in 2v1 graphs, which can depend on $Q_{\min}/\Lambda$ or $Q_2/Q_1$. Our expression for the regularised DPS cross section correctly reproduces these logarithms if we make an adequate choice of DPDs and scales. Specifically, we must

- use DPDs that have the correct short-distance behaviour, as discussed in section 3.3. Fixed-order perturbative results for this behaviour can be used reliably at scales of order $\mu_y$, as in (6.9) and (6.13).

- take scales $\mu_1, \mu_2$ and $\nu$ in $\sigma_{\text{DPS}}$ as specified in (6.1), with possible modifications by factors of order unity.

---

[7]Note that in equation (3.13) of [17] a factor $1/2$ is missing on the r.h.s.
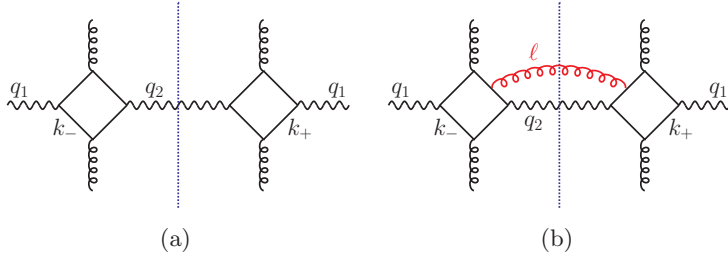
**Figure 11**. (a) Hard scattering part of the double box graph. (b) Higher-order correction giving rise to a large DGLAP logarithm in the DPS region.

The situation is quite different for 1v1 graphs. In the DPS region

$$q_1 \ll k_\pm \ll Q_{\min} \qquad (6.18)$$

the double box in figure 11a develops a squared DPS logarithm in $d\sigma/dq_1^2$,

$$\int_{q_1^2}^{Q_{\min}^2} \frac{dk_+^2}{k_+^2} \int_{q_1^2}^{Q_{\min}^2} \frac{dk_-^2}{k_-^2} = \log^2 \frac{Q_{\min}^2}{q_1^2} \qquad (6.19)$$

which disappears however when one integrates over $q_1$ up to $Q_{\min}$. This is inevitable since the integrated 1v1 cross section depends on a single dimensionful quantity $Q_{\min}$ and hence cannot have any large logarithm. An infrared cutoff $\Lambda$ in the $q_1$ integration does not play any role here since the cross section is dominated by large $q_1 \sim Q_{\min}$, where strong ordering as in (6.18) is not possible.

One finds that the higher-order graph in figure 11b does have a DGLAP logarithm in the strong-ordering regime

$$q_1 \ll k_\pm \ll Q_1 , \qquad\qquad k_\pm \ll \ell \ll Q_2 , \qquad (6.20)$$

due to the additional loop integral $\int d\ell^2/\ell^2$. But the region (6.20) gives only a power suppressed contribution to the cross section integrated over $q_1$, where DPS kinematics is not dominant.[8]

There is an exception to the statement just made. If at least one of the momentum fractions $x_1, x_2$ is very small and the corresponding scale $\mu_1, \mu_2$ is sufficiently large, then the evolution of $F_{\mathrm{spl,pt}}$ from $\mu_y$ to $\mu_1, \mu_2$ can significantly flatten the $1/y^2$ behaviour of the initial condition (6.9). The underlying physics is radiation of soft gluons, which strongly enhances parton distributions at low $x$, as is well known from the evolution of ordinary PDFs. The enhancement by evolution of $F_{\mathrm{spl,pt}}$ increases with the distance between $\mu_y$ and $\mu_1, \mu_2$, i.e. it increases with $y$. In the 1v1 (and also the 2v1) contribution to the DPS cross section, this enhances the importance of the region $y \gg 1/Q$, where the DPS approximation is valid. In this region, the evolution of the DPDs correctly describes the DGLAP logarithms generated by graphs like the one in figure 11b. In the numerical studies presented in section 9.2, we find that this happens for $F_{gg}$, $F_{qg}$ and $F_{\bar{q}g}$, as well as for $F_{q\bar{q}}$ when one momentum fraction is much smaller than the other.

---

[8]This point was overlooked in the discussion of DGLAP logarithms in [42, 43].

## 6.2 Dependence on the cutoff scale

As already discussed in section 4.1, our formalism relies on the fact that in the region of small $y \sim 1/Q$ the subtraction terms in the cross section (4.16) cancel against the DPS contribution, up to corrections that are beyond the accuracy of the calculation. Such corrections include higher-order terms in $\alpha_s$, whose parametric size we now investigate.

Let $F_{\mathrm{fo}}(x_i, \boldsymbol{y}; \mu)$ denote the short distance approximation of a DPD, computed in fixed-order perturbation theory. It depends on $\mu$ via the scales of $\alpha_s$ and of PDFs or twist-four distributions (and at higher orders also via powers of $\log(y\mu)$ in the splitting kernels). In the cross section, this fixed-order approximation is used in two different ways:

1. For the small-$y$ limit of the DPDs used in the DPS cross section, $\sigma_{\mathrm{DPS}}$, one takes $F_{\mathrm{fo}}(x_i, \boldsymbol{y}; \mu_y)$ as initial condition and evolves it to the scales $\mu_1 \sim Q_1$ and $\mu_2 \sim Q_2$ of the two hard scatters.

2. In the subtraction terms, $\sigma_{\mathrm{1v1,pt}}$ and $\sigma_{\mathrm{2v1,pt}}$, one takes $F_{\mathrm{fo}}(x_i, \boldsymbol{y}; \mu_h)$ evaluated at the same scale $\mu_h$ that is used in the hard scattering terms, $\sigma_{\mathrm{SPS}}$ or $\sigma_{\mathrm{tw4}}$, so as to cancel the contribution from the region $y \gg 1/Q$ order by order in perturbation theory.

The scale $\mu_h$ may be a combination of $Q_1$ or $Q_2$ or an independent third scale. For now we assume that $\alpha_s L \ll 1$, where

$$L = \max\left\{|\log(Q_2/Q_1)|, |\log(\mu_h/Q_1)|, |\log(\mu_h/Q_2)|\right\}. \tag{6.21}$$

The case where this is not satisfied because the hard scales are widely different is discussed in the following subsection.

Consider now a calculation at $\mathrm{N}^k\mathrm{LO}$ ($k = 0$ for LO), using DPDs evolved with $\mathrm{N}^k\mathrm{LO}$ kernels $P$ and $\mathrm{N}^k\mathrm{LO}$ expressions for the DPS subprocess cross sections $\hat{\sigma}_{1,2}$ and the short-distance approximations $F_{\mathrm{fo}}$. We have

$$\frac{d}{d\log\mu_1^2}\,\hat{\sigma}_1 + \hat{\sigma}_1 \underset{x_1}{\otimes} P + \hat{\sigma}_1 \underset{\bar{x}_1}{\otimes} P = \mathcal{O}\big(\alpha_s^{k+1}\log^k(\mu_1/Q_1)\big) \times \hat{\sigma}_1 \tag{6.22}$$

and its analogue for $\hat{\sigma}_2$. Here $\otimes$ denotes the usual Mellin convolution for the indicated variable, and we omit labels and sums for parton species. To understand (6.22) we observe that the l.h.s. would be exactly zero if $\hat{\sigma}_1$ and $P$ were calculated to all orders; the convolution of $\hat{\sigma}_1$ with the relevant parton densities would then be $\mu_1$ independent. With $\mathrm{N}^k\mathrm{LO}$ truncations on the l.h.s., the first term beyond the accuracy of the calculation is of order $\alpha_s^{k+1}\hat{\sigma}_1$. Furthermore, the highest power of $\log(\mu_1/Q_1)$ in $\hat{\sigma}_1$ is $k$ according to the general pattern for logarithms related with the renormalisation or factorisation scale. In a similar fashion we can derive

$$\frac{d}{d\log\mu^2}\,F_{\mathrm{spl,fo}}(x_i, \boldsymbol{y}; \mu) = P \underset{x_1}{\otimes} F_{\mathrm{spl,fo}}(\boldsymbol{y}; \mu) + P \underset{x_2}{\otimes} F_{\mathrm{spl,fo}}(\boldsymbol{y}; \mu)$$
$$+ \mathcal{O}\big(\alpha_s^{k+1}\log^k(y\mu)\big) \times F_{\mathrm{spl,fo}}(\boldsymbol{y}; \mu), \tag{6.23}$$

using that $k$ is the highest power of $\log(y\mu)$ in the N$^k$LO approximation, according to (3.16). Finally, it is easy to see that for a quantity satisfying

$$\frac{d}{d\log\mu^2}\,S(\mu) = \mathcal{O}\big(\alpha_s^{k+1}\log^k(\mu/\mu_0)\big)\times S(\mu) \tag{6.24}$$

one has

$$S(\mu_b) - S(\mu_a) = \mathcal{O}\big(\alpha_s^{k+1}\log^{k+1}(\mu_a/\mu_0),\alpha_s^{k+1}\log^{k+1}(\mu_b/\mu_0)\big)\times S(\mu_a)\,, \tag{6.25}$$

where the integration of the differential equation (6.24) between $\mu_a$ and $\mu_b$ gives an extra logarithm on the r.h.s. We thus obtain for the small-$y$ contribution to the 1v1 cross section

$$
\begin{aligned}
\big[\sigma_{\mathrm{1v1}}\big]_{y\sim 1/Q} &= \hat{\sigma}_1(\mu_1)\otimes\hat{\sigma}_2(\mu_2)\otimes\int\limits_{y\sim 1/Q} d^2\boldsymbol{y}\,F_{\mathrm{spl,pt}}(\boldsymbol{y};\mu_1,\mu_2)\,F_{\mathrm{spl,pt}}(\boldsymbol{y};\mu_1,\mu_2)\\
&= \int\limits_{y\sim 1/Q} d^2\boldsymbol{y}\,\hat{\sigma}_1(\mu_y)\otimes\hat{\sigma}_2(\mu_y)\otimes F_{\mathrm{spl,fo}}(\boldsymbol{y};\mu_y)\,F_{\mathrm{spl,fo}}(\boldsymbol{y};\mu_y)\\
&\quad\times\left[1+\mathcal{O}\big(\alpha_s^{k+1}\log^{k+1}(Q_1/\mu_y),\,\alpha_s^{k+1}\log^{k+1}(Q_2/\mu_y)\big)\right]\\
&= \hat{\sigma}_1(\mu_h)\otimes\hat{\sigma}_2(\mu_h)\otimes\int\limits_{y\sim 1/Q} d^2\boldsymbol{y}\,F_{\mathrm{spl,fo}}(\boldsymbol{y};\mu_h)\,F_{\mathrm{spl,fo}}(\boldsymbol{y};\mu_h)\\
&\quad\times\left[1+\mathcal{O}\big(\alpha_s^{k+1}\log^{k+1}(y\mu_h),\,\alpha_s^{k+1}\log^{k+1}(yQ_1),\,\alpha_s^{k+1}\log^{k+1}(yQ_2)\big)\right]\,, 
\end{aligned}\tag{6.26}
$$

using our initial condition $F_{\mathrm{spl,pt}} = F_{\mathrm{spl,fo}}$ at scale $\mu_y$. The longitudinal momentum arguments and convolutions are as in (2.7). The first line of the last expression in (6.26) is just the double counting subtraction term. We thus find

$$
\begin{aligned}
\big[\sigma_{\mathrm{1v1}} - \sigma_{\mathrm{1v1,pt}}\big]_{y\sim 1/Q} &= \big[\sigma_{\mathrm{1v1,pt}}\big]_{y\sim 1/Q}\times\mathcal{O}\big(\alpha_s^{k+1}L^{k+1}\big)\\
&= \big[\sigma_{\mathrm{SPS}} - \sigma_{\mathrm{1v1,pt}}\big]\times\mathcal{O}\big(\alpha_s^{k+1}L^{k+1}\big)\,. 
\end{aligned}\tag{6.27}
$$

with $L$ defined in (6.21). In the second step we have used that the full SPS cross section is of the same order of magnitude as the DPS subtraction term, regarding its power behaviour in $Q^2$ as well as the relevant coupling constants and PDFs. In full analogy one can derive

$$
\begin{aligned}
\big[\sigma_{\mathrm{2v1}} - \sigma_{\mathrm{2v1,pt}}\big]_{y\sim 1/Q} &= \big[\sigma_{\mathrm{2v1,pt}}\big]_{y\sim 1/Q}\times\mathcal{O}\big(\alpha_s^{k+1}L^{k+1}\big)\\
&= \big[\sigma_{\mathrm{tw4}} - \sigma_{\mathrm{2v1,pt}}\big]\times\mathcal{O}\big(\alpha_s^{k+1}L^{k+1}\big)\,. 
\end{aligned}\tag{6.28}
$$

Notice that the difference $\sigma_{\mathrm{tw4}} - \sigma_{\mathrm{2v1,pt}}$ is by construction dominated by short distances $y\sim 1/Q$, whilst the individual terms receive contributions from a wide range of $y$. We thus find that, for both 1v1 and 2v1 contributions, the DPS term cancels against the double counting subtractions in the region $y\sim 1/Q$, up to corrections of relative order $\alpha_s^{k+1}$. This is beyond the accuracy of an N$^k$LO calculation, as it should be. The only logarithms multiplying $\alpha_s$ in this context are logarithms of the ratios between different hard scales, which at this point we must require to be of moderate size.

We can now also quantify the dependence of the overall cross section $\sigma$ in (4.16) on the scale $\nu$. With two choices $\nu_1 \sim \nu_2 \sim Q$ the difference $\sigma(\nu_1) - \sigma(\nu_2)$ is just of the size given by (6.27) and (6.28). In analogy to the variation of renormalisation and factorisation scales, one may thus use the variation of $\nu$ to estimate uncalculated higher-order corrections.

An important consequence of (6.27) is that

$$\big[\sigma_{1v1}\big]_{y \sim 1/Q} \sim \sigma_{1v1,pt} \sim \sigma_{SPS} \,, \tag{6.29}$$

where we used that the subtraction term $\sigma_{1v1,pt}$ is dominated by $y \sim 1/Q$, given the $1/y^2$ behaviour of $F_{spl,fo}$. This means that we can use the $\nu$ variation of the 1v1 contribution to $\sigma_{DPS}$ to estimate the size of the subtraction term $\sigma_{1v1,pt}$ and thus of the SPS contribution to the cross section at the corresponding order in $\alpha_s$. This is of great practical value, given the considerable difficulty to compute SPS at this order. We will come back to this point in section 9.2.

## 6.3 Multiscale problems

The arguments leading to (6.27) and (6.28) are based on an expansion in $\alpha_s$ and break down when $\alpha_s L$ is not small, for instance because there is a large hierarchy $Q_1 \ll Q_2$ between the two hard scales. We now show that a modified choice of DPDs in the subtraction terms $\sigma_{1v1,pt}$ and $\sigma_{2v1,pt}$ can cope with such a situation.

Let us for definiteness describe the 1v1 subtraction; the 2v1 case can be treated in full analogy. We use the notation $F_{spl,pt}(x_i, \boldsymbol{y}; \mu_b, \mu_c | \mu_a)$ for the distribution that is initialised to

$$F_{spl,pt}(x_i, \boldsymbol{y}; \mu_a, \mu_a | \mu_a) = F_{spl,fo}(x_i, \boldsymbol{y}; \mu_a) \tag{6.30}$$

and evolved according to the double DGLAP equations to the scale $\mu_b$ ($\mu_c$) for the parton with momentum fraction $x_1$ ($x_2$). The small-$y$ limit of the distributions used in $\sigma_{DPS}$ is $F_{spl,pt}(x_i, \boldsymbol{y}; \mu_1, \mu_2 | \mu_y)$ and its analogue for $F_{int,pt}$. Let us now introduce a profile function $p(u; \mu_a, \mu_b)$ that interpolates monotonically between the limiting values $\mu_a$ at $u = 0$ and $\mu_b$ at $u \to \infty$, satisfying

$$\begin{aligned} p(u; \mu_a, \mu_b) &\approx \mu_a && \text{for } u \sim 1, \\ p(u; \mu_a, \mu_b) &\approx \mu_b && \text{for } u \gg 1. \end{aligned} \tag{6.31}$$

We can then take distributions

$$F_{spl,pt}\big(x_i, \boldsymbol{y}; p(y\nu; \mu_1, \mu_h), p(y\nu; \mu_2, \mu_h) \,|\, p(y\nu; \nu, \mu_h)\big) \tag{6.32}$$

in $\sigma_{1v1,pt}$. These distributions tend to $F_{spl,fo}(x_i, \boldsymbol{y}; \mu_h)$ for $y \gg 1/\nu$, so that the subtraction term cancels the contribution of $\sigma_{SPS}$ in this region. For $y \sim 1/\nu$ the distributions tend to $F_{spl,pt}(x_i, \boldsymbol{y}; \mu_1, \mu_2 | \nu) \approx F_{spl,pt}(x_i, \boldsymbol{y}; \mu_1, \mu_2 | \mu_y)$, so that in that region the subtraction term cancels against $\sigma_{DPS}$. Distances $y \ll 1/\nu$ are of course removed in the cross section by the cutoff function $\Phi(y\nu)$. Using the same technique as in section 6.2, one can show that the cancellations just described are up to corrections of order $\alpha_s^{k+1}$ at $N^k LO$, without

any large logarithms. In particular, this ensures that the $\nu$ dependence of the overall cross section is beyond the accuracy of the computation.

As an alternative, one may replace $\nu$ in (6.32) with $\min(\mu_1, \mu_2)$, which requires less computation when $\nu$ is varied in the cross section (while $\mu_1, \mu_2, \mu_h$ are kept fixed). With this choice, the DPDs in the subtraction term reduce to $F_{\mathrm{spl,fo}}(x_i, \boldsymbol{y}; \mu_h)$ in the case of equal hard scales, $\mu_1 = \mu_2 = \mu_h$.

Profile functions for the renormalisation scale have been used in other contexts, see e.g. [44, 45]. An suitable function for our purpose is

$$p(u; \mu_a, \mu_b) = \begin{cases} \mu_a & \text{for } u \leq u_1 \\ \dfrac{\mu_b + \mu_a}{2} + \dfrac{\mu_b - \mu_a}{2}\left[\dfrac{4(u - u_2)^3}{(u_1 - u_3)^3} - \dfrac{3(u - u_2)}{u_1 - u_3}\right] & \text{for } u_1 < u < u_3 \\ \mu_b & \text{for } u \geq u_3 \end{cases} \quad (6.33)$$

where $u_1 \sim 1$, $u_3 \gg 1$ and $u_2 = \frac{1}{2}(u_1 + u_3)$. Functions that are piecewise polynomials have been used for instance in [45, 46]. To estimate theoretical uncertainties, one can vary the profile function $p$, for instance by varying $u_1$ and $u_3$. For a detailed discussion we refer to [46].

## 6.4 Approximation of the intrinsic distribution

As discussed earlier, the correct small-$y$ limit $F_{\mathrm{int,pt}}$ at leading order in $\alpha_s$ is obtained by making the ansatz (6.13) at scale $\mu_y$ and then evolving $F_{\mathrm{int,pt}}$ to the scales $\mu_1, \mu_2$ relevant for the DPS cross section. In practical terms, this requires an ansatz or model for the twist-four distribution $G$ at some reference scale $\mu_0$ and subsequent evolution of $G$ to the scale $\mu_y$. The technical implementation of twist-four evolution is rather involved, and one may wish to avoid it in situations where the corresponding loss of precision is tolerable.

In the model used later in this paper we indeed follow this path, taking as limiting form for small $y$

$$F_{\mathrm{int,pt}}(x_1, x_2, \boldsymbol{y}; \mu_0, \mu_0) = G(x_1, x_2, x_2, x_1; \mu_0) \qquad (6.34)$$

at some fixed reference scale $\mu_0$ and then evolving $F_{\mathrm{int,pt}}$ to the final scales $\mu_1, \mu_2$. For the r.h.s. we will take a form proportional to the product $f(x_1, \mu_0) f(x_2, \mu_0)$ of ordinary PDFs (this model is restricted to DPDs in the colour singlet channel). Multiplying with a suitable function of $y$, one readily obtains a model for the intrinsic DPD at any value of $y$. This is indeed standard in the existing literature.

As follows from the discussions below (3.19) and below (6.13), evolution of the ansatz (6.34) to scales $\mu_1, \mu_2$ correctly takes into account diagonal interactions as in figure 6c whilst neglecting non-diagonal interactions as in figure 6a and b, which in the correct treatment contribute to the evolution of $G$ from $\mu_0$ to $\mu_y$. For $\sigma_{2v2}$, these interactions are irrelevant since the small-$y$ region is power suppressed. Furthermore, we can expect them to be of limited importance for $\sigma_{2v1}$ in the situation described at the end of section 6.1. Evolution then leads to a $y$ dependence of $F_{\mathrm{spl,pt}}$ that is significantly flatter than the $1/y^2$ behaviour obtained at fixed order. This shifts the dominant region of the $y$ integration in

$\sigma_{2v1}$ to larger values, so that on the $F_{\text{int,pt}}$ side there remains little room for the evolution of $G$ between $\mu_0$ to $\mu_y$. The limited numerical importance of non-diagonal interactions for 2v1 diagrams in such situations was also observed in [17], where calculations were done in transverse momentum space.

# 7 Collinear DPDs in momentum space

As we argued in section 4, there are several advantages to work with DPDs in $\boldsymbol{y}$ space when setting up a factorisation formula for the overall cross section. We now consider collinear DPDs depending on the Fourier conjugate transverse momentum $\boldsymbol{\Delta}$, showing that they can be connected with our formalism in a useful way. The motivation for this is twofold.

- Sum rules for momentum space DPDs at $\Delta = 0$ have been proposed in [9], where it was stated that they are preserved by the leading-order inhomogeneous evolution equations in [6–10]. If it can be shown that they are valid at some starting scale, this will provide valuable (albeit nontrivial) constraints on DPDs. Of course, no such sum rules can hold for $\int d^2\boldsymbol{y}\, F(x_i, \boldsymbol{y})$ since this integral is divergent at small $y$.

- Previous work on DPS, in particular on the perturbative splitting contributions [12–15], is formulated in transverse momentum space. The results derived below will facilitate the comparison of that work with ours, which is the subject of section 8.

Both aspects only concern colour singlet DPDs, to which the rest of this section is restricted. Colour non-singlet DPDs have an additional rapidity dependence described by Collins-Soper equations, which are multiplicative in $\boldsymbol{y}$ space (see [21]) and hence more complicated in momentum space.

The function $\Phi$ that regulates the DPS cross section in our formalism can also be used to introduce momentum space DPDs. We define

$$F_\Phi(x_i, \boldsymbol{\Delta}; \mu, \nu) = \int d^2\boldsymbol{y}\, e^{i\boldsymbol{y}\boldsymbol{\Delta}}\, \Phi(y\nu) F(x_i, \boldsymbol{y}, \mu)\,, \tag{7.1}$$

omitting the superscript indicating the colour singlet channel for brevity. With the conditions on $\Phi$ stated in (4.2), the integral converges at $y = 0$. Throughout this section, we restrict ourselves to the two choices of $\Phi$ in table 1 and assume that $\nu \gg \Lambda$. The dependence of $F_\Phi$ on $\nu$ is then given by

$$\frac{dF_\Phi(x_i, \boldsymbol{\Delta}; \mu, \nu)}{d\log \nu^2} = \int d^2\boldsymbol{y}\, e^{i\boldsymbol{y}\boldsymbol{\Delta}}\, \frac{d\Phi(y\nu)}{d\log \nu^2}\, F_{\text{spl,pt}}(x_i, \boldsymbol{y}; \mu)\,. \tag{7.2}$$

The derivative of $\Phi(y\nu)$ strongly suppresses the nonperturbative region $y \sim 1/\Lambda$ in the integral (or removes it altogether if $\Phi$ is a step function). This allows us to replace $F(x_i, \boldsymbol{y}, \mu)$ with its perturbative splitting part on the r.h.s. The contribution from $F_{\text{int,pt}}(x_i, \boldsymbol{y}, \mu)$ is suppressed by $\Lambda^2/\nu^2$. Inserting the expressions (3.12) or (3.14) for $F_{\text{spl,pt}}(x_i, \boldsymbol{y}, \mu)$ at $\epsilon = 0$ one obtains

$$\frac{dF_\Phi(x_i, \boldsymbol{\Delta}; \mu, \nu)}{d\log \nu^2} = \frac{f(x_1 + x_2; \mu)}{x_1 + x_2}\, \frac{\alpha_s(\mu)}{2\pi}\, P\left(\frac{x_1}{x_1 + x_2}, 0\right) I_0(\Delta, \mu, \nu) \tag{7.3}$$

at leading order and

$$\frac{dF_\Phi(x_i, \boldsymbol{\Delta}; \mu, \nu)}{d \log \nu^2} = \int\limits_{x_1+x_2}^{1} \frac{dv}{v} \frac{f(v; \mu)}{v} \sum_{n=1}^{\infty} \alpha_s^n(\mu) \sum_{m=0}^{n-1} V^{[n,m]}\left(\frac{x_1}{v}, \frac{x_2}{v}\right) I_m(\Delta, \mu, \nu) \qquad (7.4)$$

in general, where

$$I_m(\Delta, \mu, \nu) = \int\limits_{0}^{\infty} du \, \frac{d\Phi(u)}{du} \, J_0(u\Delta/\nu) \, \log^m(u\mu/\nu) \,. \qquad (7.5)$$

For brevity, we have omitted labels and sums for parton species and polarisations.

For $\Delta = 0$ we have $I_0 = 1$ and recognise in (7.3) the inhomogeneous term of the evolution equation for a conventional momentum space DPD $F(x_i, \boldsymbol{\Delta}, \mu)$ defined by $\overline{\text{MS}}$ subtraction. At leading order, our momentum space DPD $F_\Phi(x_i, \mathbf{0}, \mu, \mu)$ thus obeys the same evolution equation as $F(x_i, \boldsymbol{\Delta}, \mu)$. At small $\Delta$ the inhomogeneous term is multiplied by $I_0(\Delta, \mu, \mu) = 1 + \mathcal{O}(\Delta^2/\mu^2)$. The higher-order version (7.4) at $\Delta = 0$ has the same structure as the inhomogeneous term in the evolution equation proposed in [10], although the evolution kernel may not be the same.

Let us now compute the momentum space analogue of $F_{\text{spl,pt}}(x_i, \boldsymbol{y}, \mu)$, both with the cutoff regularisation (7.1) and with $\overline{\text{MS}}$ renormalisation of the splitting singularity. This gives the correct result for the full momentum-space DPD in the limit $\Delta \gg \Lambda$, when the exponent $e^{i\boldsymbol{y}\boldsymbol{\Delta}}$ suppresses large $y \sim 1/\Lambda$ in the Fourier integral, so that the perturbative splitting approximation can be used.

Restricting ourselves to leading order, we start with (3.12) and use the Fourier integrals given in (A.3) and (A.4). For $\Delta \ll \nu$ we find

$$\frac{F_{\Phi,\text{spl,pt}}(x_i, \boldsymbol{\Delta}; \mu, \nu)}{R(x_i, \mu)} \approx \begin{cases} \log \dfrac{\nu^2}{\Delta^2} & \text{for } \Phi(u) = \Theta(u - b_0) \\[2mm] \log \dfrac{\nu^2}{\Delta^2} - \gamma & \text{for } \Phi(u) = 1 - e^{-u^2/4} \end{cases} \qquad (7.6)$$

with corrections of $\mathcal{O}(\Delta^2/\nu^2)$, whereas for $\Delta \gg \nu$ we obtain

$$\frac{F_{\Phi,\text{spl,pt}}(x_i, \boldsymbol{\Delta}; \mu, \nu)}{R(x_i, \mu)} \approx \begin{cases} \dfrac{e^{3\gamma/2}}{\sqrt{\pi}} \left(\dfrac{\nu}{\Delta}\right)^{3/2} \cos\left(b_0 \dfrac{\Delta}{\nu} + \dfrac{\pi}{4}\right) & \text{for } \Phi(u) = \Theta(u - b_0) \\[3mm] \dfrac{\nu^2}{\Delta^2} \exp\left[-\dfrac{\Delta^2}{\nu^2}\right] & \text{for } \Phi(u) = 1 - e^{-u^2/4} \end{cases} \qquad (7.7)$$

with

$$R(x_i, \mu) = \frac{f(x_1 + x_2; \mu)}{x_1 + x_2} \frac{\alpha_s(\mu)}{2\pi} P\left(\frac{x_1}{x_1 + x_2}, 0\right). \qquad (7.8)$$

To compute the usual momentum space DPD, we take the Fourier transform (2.9) of $F_{\text{spl,pt}}(x_i, \boldsymbol{y})$ in $D = 4 - 2\epsilon$ dimensions, subtract the UV divergence and then set $\epsilon = 0$. This gives

$$F_{\text{spl,pt}}(x_i, \boldsymbol{\Delta}; \mu) = \lim_{\epsilon \to 0} \mu^{2\epsilon} \int d^{2-2\epsilon}\boldsymbol{y} \, e^{i\boldsymbol{y}\boldsymbol{\Delta}} \frac{1}{y^{2-4\epsilon}} \frac{\Gamma^2(1-\epsilon)}{\pi^{1-2\epsilon}} \frac{\alpha_s}{2\pi} P(v, \epsilon) \frac{f(x)}{x} - C_\epsilon \,, \qquad (7.9)$$

where we have abbreviated $x = x_1 + x_2$ and $v = x_1/(x_1 + x_2)$. Here $C_\epsilon$ denotes the $\overline{\text{MS}}$ subtraction term for the splitting singularity. Using (A.1) we get

$$
\begin{aligned}
F_{\text{spl,pt}}(x_i, \boldsymbol{\Delta}; \mu) &= \lim_{\epsilon \to 0} \left( \frac{\mu}{\Delta} \right)^{2\epsilon} \frac{\Gamma(\epsilon)\,\Gamma^2(1-\epsilon)}{\Gamma(1-2\epsilon)} \, (4\pi)^\epsilon \, \frac{\alpha_s}{2\pi} \, P(v, \epsilon) \frac{f(x)}{x} - C_\epsilon \\
&= \lim_{\epsilon \to 0} \left[ \frac{1}{\epsilon} + \log(4\pi) - \gamma + \log \frac{\mu^2}{\Delta^2} \right] \frac{\alpha_s}{2\pi} P(v,0) \frac{f(x)}{x} + \frac{\alpha_s}{2\pi} P'(v,0) \frac{f(x)}{x} - C_\epsilon, \quad (7.10)
\end{aligned}
$$

where $P'(v,\epsilon) = \partial P(v,\epsilon)/\partial\epsilon$. The $\overline{\text{MS}}$ counterterm is equal to

$$
C_\epsilon = \left[ \frac{1}{\epsilon} + \log(4\pi) - \gamma \right] \frac{\alpha_s}{2\pi} P(v,0) \frac{f(x)}{x}, \tag{7.11}
$$

so that we obtain

$$
F_{\text{spl,pt}}(x_i, \boldsymbol{\Delta}; \mu) = \left[ \log \frac{\mu^2}{\Delta^2} + \frac{P'(v,0)}{P(v,0)} \right] R(x_i, \mu). \tag{7.12}
$$

We can use the preceding results to compute the matching between the DPD defined with a $\Phi$ cutoff and the one obtained by $\overline{\text{MS}}$ renormalisation. Since the integral in (7.1) is finite, it can be smoothly continued to $D = 4 - 2\epsilon$ dimensions. We then have

$$
F(x_i, \boldsymbol{\Delta}; \mu) - F_\Phi(x_i, \boldsymbol{\Delta}; \mu, \nu) = \lim_{\epsilon \to 0} \int d^{2-2\epsilon} \boldsymbol{y} \, e^{i\boldsymbol{y}\boldsymbol{\Delta}} \, F(x_i, \boldsymbol{y}; \mu) \left[ 1 - \Phi(y\nu) \right] - C_\epsilon. \tag{7.13}
$$

The factor $[1 - \Phi(y\nu)]$ strongly suppresses the region $y \gg 1/\nu$, so that on the r.h.s. we can use the fixed-order perturbative splitting contribution, evaluated in $4 - 2\epsilon$ dimensions. This gives

$$
F(x_i, \boldsymbol{\Delta}; \mu) - F_\Phi(x_i, \boldsymbol{\Delta}; \mu, \nu) = F_{\text{spl,pt}}(x_i, \boldsymbol{\Delta}; \mu) - F_{\Phi,\text{spl,pt}}(x_i, \boldsymbol{\Delta}; \mu, \nu) \tag{7.14}
$$

for all $\boldsymbol{\Delta}$, and we have a smooth limit for $\boldsymbol{\Delta} \to 0$. For $\Delta \ll \nu$ we can use (7.6) for $\Phi(u) = \Theta(u - b_0)$ with (7.12) and obtain

$$
F(x_i, \boldsymbol{\Delta}; \mu) - F_\Phi(x_i, \boldsymbol{\Delta}; \mu, \nu) = \left[ \log \frac{\mu^2}{\nu^2} + \frac{P'(v,0)}{P(v,0)} \right] R(x_i, \mu) + \mathcal{O}\left( \frac{\Delta^2}{\nu^2} \right) + \mathcal{O}(\alpha_s^2). \tag{7.15}
$$

At $\Delta = 0$ we can in particular achieve equality of the two DPDs by choosing the scale

$$
\nu = \mu \exp \left[ \frac{P'(v,0)}{2P(v,0)} \right]. \tag{7.16}
$$

With the splitting kernel $P_{g \to q\bar{q}}$ from (3.13) this gives for instance a scale varying between $\nu = \mu e^{-1/2} \approx 0.6\mu$ and $\nu = \mu$. These values respectively pertain to $v = 1/2$ (i.e. $x_1 = x_2$) and to $v = 0$ or $1$ (i.e. $x_1 = 0$ or $x_2 = 0$).

## 8 Comparison with other work

Three groups have previously studied the consequences of the perturbative splitting mechanism on DPS cross sections, namely Blok et al. [12, 13], Ryskin and Snigirev [14, 15], and Manohar and Waalewijn [16]. We refer to them as BDFS, RS and MW, respectively. In this section we compare the results obtained in our formalism with theirs. We restrict ourselves to collinear factorisation. To begin with, we rearrange our master formula (4.16) as

$$\sigma = \sigma_{\text{SPS}} + (\sigma_{\text{tw4}} - \sigma_{\text{2v1,pt}}) + (\sigma_{\text{DPS}} - \sigma_{\text{1v1,pt}}) \tag{8.1}$$

and recall that one should take $\nu \sim \min(Q_1, Q_2)$ for the cutoff parameter. Let us briefly characterise the different terms in (8.1).

In all approaches just mentioned, the SPS cross section $\sigma_{\text{SPS}}$ is defined in the standard way. It is obtained by multiplying two PDFs with a hard scattering cross section and summing over all parton channels.

The term $\sigma_{\text{tw4}} - \sigma_{\text{2v1,pt}}$ does not have any equivalent in the work of BDFS and of RS. As discussed in section 4.1.2, this term is lacking the large $\log(Q/\Lambda)$ that is contained in $\sigma_{\text{DPS}}$ and built up by parton splitting taking place over a wide region $1/Q \ll y \ll 1/\Lambda$ in the 2v1 graph 1b. The approaches of BDFS and RS can thus only be valid at leading logarithmic accuracy. This limitation is explicitly stated in the work of BDFS [12]. The same limitation holds of course in our approach if $\sigma_{\text{tw4}} - \sigma_{\text{2v1,pt}}$ is not computed.

The term $\sigma_{\text{DPS}} - \sigma_{\text{1v1,pt}}$ includes the 2v2 and 2v1 parts of DPS, whereas the 1v1 part has its small $y$ approximation subtracted. As discussed at the end of section 6.1 and illustrated in section 9.2, there are situations where strongly ordered parton emissions in 1v1 graphs enhance the region $y \gg 1/Q$ for the initial parton splitting. In our approach, this effect is included via the evolution of the DPDs in $\sigma_{\text{DPS}}$ (which thus provides a small $x$ enhancement for $\sigma_{\text{2v2}}$, $\sigma_{\text{2v1}}$ and $\sigma_{\text{1v1}}$).

If such an enhancement does not take place, then the 1v1 part of DPS has an integrand going like $1/y^4$ at small $y$, so that the integral is dominated by the region $y \sim 1/Q$. In $\sigma_{\text{DPS}} - \sigma_{\text{1v1,pt}}$ the contribution from this region then cancels up to higher orders in $\alpha_s$, as shown in section 6.2. In turn, the contribution from $y \sim 1/\Lambda$ is suppressed by $\Lambda^2/Q^2$ compared with $\sigma_{\text{SPS}}$ and thus of the same order as other power corrections beyond the accuracy of the calculation. In this situation, the dominant terms in $\sigma_{\text{DPS}} - \sigma_{\text{1v1,pt}}$ are the 2v2 and 2v1 parts of DPS, provided that they have a small $x$ enhancement compared with generic power corrections to SPS.

### 8.1 The approach of Manohar and Waalewijn

The master formula for the overall cross section in the MW approach is given in eq. (14) of [16]. Adapted to our notation, it reads

$$\sigma = \sigma_{\text{SPS}} + \hat{c}_3 \left[ f F(\boldsymbol{y} = \boldsymbol{0}) \right] + \hat{c}_3 \left[ F(\boldsymbol{y} = \boldsymbol{0}) f \right] + \hat{c}_4 \left[ \int d^2\boldsymbol{y} \, F(\boldsymbol{y}) F(\boldsymbol{y}) \right], \tag{8.2}$$

where $\hat{c}_3$ and $\hat{c}_4$ are hard scattering coefficients and $[\cdots]$ indicates UV renormalisation of the enclosed quantity. This means that the product $F(\boldsymbol{y})F(\boldsymbol{y})$ of DPDs is integrated over

$\boldsymbol{y}$ in $2 - 2\epsilon$ transverse dimensions, and the resulting poles in $\epsilon$ are subtracted as usual in dimensional regularisation (we assume $\overline{\text{MS}}$ renormalisation, although this is not explicitly stated in [16]). The terms going with $\hat{c}_3$ involve the renormalised product of a twist-two distribution in one proton and a twist-four distribution in the other. In the following, the terms with $\hat{c}_4$ and $\hat{c}_3$ in (8.2) will be called DPS and twist four terms, respectively. Via the perturbative splitting mechanism in one of the two protons, the operator in the DPS term mixes under renormalisation with the operators in the twist-four terms. This is discussed at LO in [16], but we see no obstacle to extending the analysis to higher orders.

In section 4.1.2 we have seen that the computation of the hard scattering coefficient for graph 7 has a divergence from the region where the $g \to q\bar{q}$ splittings become collinear. In our formalism the term $\sigma_{2\text{v1,pt}}$ in (8.1) subtracts the contribution from this region. In the MW approach, the general subtraction formalism recalled in section 4 will yield an $\overline{\text{MS}}$ subtraction term for the collinear divergence, resulting from the renormalisation procedure described in the previous paragraph. For the choice $\Phi(u) = \Theta(u - b_0)$ we see in (4.14) that $\sigma_{2\text{v1,pt}}$ is proportional to the $\overline{\text{MS}}$ expression $1/\epsilon + \log(4\pi) - \gamma$ plus an extra term $\log(\mu^2/\nu^2) + P'(0)/P(0)$, where $P(\epsilon)$ is the relevant splitting function in $4 - 2\epsilon$ dimensions. This extra term corresponds to the difference between the momentum space DPD $F_{\Phi,\text{spl,pt}}$ in (7.6) and its $\overline{\text{MS}}$ counterpart in (7.12). Up to this difference, the term $\sigma_{\text{tw4}} - \sigma_{2\text{v1,pt}}$ in our master formula (8.1) is thus equivalent to the twist four term in (8.2). The mismatch between the two versions may be understood as a scheme difference.

It remains to compare the DPS term in the MW approach with our term $\sigma_{\text{DPS}} - \sigma_{1\text{v1,pt}}$. As the scale evolution of $\left[ \int d^2\boldsymbol{y}\, F(\boldsymbol{y})F(\boldsymbol{y}) \right]$ derived by MW reflects perturbative splitting in one of the two protons, their DPS term should contain the contributions from 2v2 and 2v1 graphs. Our approach also contains the contribution of 1v1 graphs, with a subtraction of their small $y$ approximation. As discussed above, this part is negligible if DPS evolution effects are moderate. Both approaches should then give a valid representation of the overall cross section, with the possibilities to go beyond the leading logarithmic approximation and to systematically include higher orders in $\alpha_s$. On the other hand, strongly ordered emissions can significantly enhance the the region $y \gg 1/Q$ in 1v1 graphs. In our formalism, such emissions are resummed to all orders by the evolution of the DPDs, but we see no counterpart of this in the MW formulation.

MW argue that under renormalisation $\left[ \int d^2\boldsymbol{y}\, F(\boldsymbol{y})F(\boldsymbol{y}) \right]$ does not mix with the product of two PDFs by pointing out that the 1v1 graph in figure 1a gives a quadratically divergent scaleless integral, which is zero in dimensional regularisation. This argument hinges on using the perturbative approximation for the $g \to q\bar{q}$ splitting at all values of $\boldsymbol{y}$ (or equivalently of the Fourier conjugate momentum $\boldsymbol{\Delta}$). This includes the infrared region, where the perturbative approximation does not work. It would be interesting to investigate this point further, but this is beyond the scope of the present paper.

## 8.2 The approaches of Blok et al. and of Ryskin and Snigirev

Since BDFS and RS use the same $\sigma_{\text{SPS}}$ as we do, and since they have no counterpart to our $\sigma_{\text{tw4}} - \sigma_{2\text{v1,pt}}$ (which can be neglected at leading logarithmic accuracy), it remains to

compare our term $\sigma_{\mathrm{DPS}} - \sigma_{\mathrm{1v1,pt}}$ with their expressions for the DPS cross section. In this comparison, we can take the leading logarithmic approximation of our result.

The approaches of BDFS and of RS are both based on separating the intrinsic and splitting contributions to a DPD at all values of $\boldsymbol{\Delta}$. As discussed in section 4, we avoid such a separation when formulating the factorisation formula for the cross section. When modelling the DPDs – which is of course unavoidable for phenomenology – it is however convenient to use this separation. In $\boldsymbol{y}$ space we thus make the ansatz

$$F(x_i, \boldsymbol{y}, \mu_i) = F_{\mathrm{int}}(x_i, \boldsymbol{y}, \mu_i) + F_{\mathrm{spl}}(x_i, \boldsymbol{y}, \mu_i) \tag{8.3}$$

for all $\boldsymbol{y}$. We demand that $F_{\mathrm{int}}$ and $F_{\mathrm{spl}}$ separately follow the homogeneous evolution equation (2.3) and that for $y \ll 1/\Lambda$ they tend to the correct short-distance limits $F_{\mathrm{int,pt}}$ and $F_{\mathrm{spl,pt}}$. At small $y$ the form (8.3) is hence unambiguous, whereas for large $y$ it represents a model. For $F_{\mathrm{int,pt}}$ we make the approximation discussed in section 6.4. Then $F_{\mathrm{int}}$ has a finite limit for $y = 0$ rather than a weak singularity induced by non-diagonal twist-four evolution. Notice that $F_{\mathrm{int}}$ and $F_{\mathrm{spl}}$ now denote two terms in a DPD model, whereas in previous sections they were associated with different Feynman graphs in the discussion of factorisation.

BDFS and RS only consider DPDs for unpolarised partons in the colour singlet sector, neglecting spin and colour correlations in both the intrinsic and the splitting parts. We will therefore do the same when comparing to our formalism. We note that spin correlations tend to be washed out by evolution to high scales [39] and that colour correlations are suppressed by Sudakov logarithms [22, 47]. For sufficiently large hard scales, the importance of such correlations will therefore be reduced.

In terms of the momentum space DPDs introduced in section 7, the DPS cross section at leading order in $\alpha_s$ is given by

$$\int \frac{d^2\boldsymbol{\Delta}}{(2\pi)^2} F_\Phi(x_i, \boldsymbol{\Delta}; \mu_i, \nu) F_\Phi(\bar{x}_i, -\boldsymbol{\Delta}; \mu_i, \nu) \tag{8.4}$$

times the appropriate subprocess cross sections and a combinatorial factor. As was done in the papers cited above, we allow for different factorisation scales $\mu_1, \mu_2$. We note that with the behaviour (7.7) of the splitting DPD at large $\boldsymbol{\Delta}$, the integral (8.4) converges in the ultraviolet for both choices of $\Phi$. For definiteness, we take $\Phi(u) = \Theta(u - b_0)$ in the following. We now analyse the 2v2, 2v1 and 1v1 contributions in turn.

The ultraviolet cutoff can be neglected in the intrinsic contribution:

$$F_{\Phi,\mathrm{int}}(x_i, \boldsymbol{\Delta}; \mu_i, \nu) = \int d^2\boldsymbol{y}\, e^{i\boldsymbol{y}\boldsymbol{\Delta}}\, \Phi(y\nu) F_{\mathrm{int}}(x_i, \boldsymbol{y}, \mu_i) \approx \int d^2\boldsymbol{y}\, e^{i\boldsymbol{y}\boldsymbol{\Delta}}\, F_{\mathrm{int}}(x_i, \boldsymbol{y}, \mu_i). \tag{8.5}$$

This is because $F_{\mathrm{int}}(x_i, \boldsymbol{y})$ is of generic size $1/\Lambda^2$ for $y \sim 1/\Lambda$ and has a smooth behaviour for $y \ll 1/\Lambda$. A cutoff $y > b_0/\nu$ in the integral therefore only removes a contribution of order $\Lambda^2/\nu^2$. If we make the same ansatz for the intrinsic DPD as BDFS or RS, we will thus obtain the same result for $\sigma_{\mathrm{2v2}}$ as they do, up to power corrections that are beyond the accuracy of all all approaches discussed here.

For a typical ansatz, $F_{\Phi,\text{int}}(x_i, \boldsymbol{\Delta})$ strongly decreases when $\Delta \gg \Lambda$. The 2v1 contribution to the integral (8.4) is therefore dominated by the region $\Delta \sim \Lambda$. In this region, the integral

$$F_{\Phi,\text{spl}}(x_i, \boldsymbol{\Delta}; \mu_i, \nu) = \int d^2\boldsymbol{y} \, e^{i\boldsymbol{y}\boldsymbol{\Delta}} \, \Phi(y\nu) \, F_{\text{spl}}(x_i, \boldsymbol{y}; \mu_i) \tag{8.6}$$

receives leading contributions from both small and large $y$. However, a large logarithm $\log(\nu/\Delta)$ is only built up in the region $1/\nu \ll y \ll 1/\Delta$, where we can use the perturbative form $F_{\text{spl,pt}}(x_i, \boldsymbol{y}) \sim 1/y^2$. To leading logarithmic accuracy, we can therefore compute the splitting contribution without a nonperturbative model, even for $\Delta \sim 1/\Lambda$, which we now do. We want to improve on our result (7.6) by including evolution effects. To do so, we start from the first order perturbative expression (3.12) at scale $\mu_y = b_0/y$ and then evolve using the Green function $D(x, \mu; \mu_y)$, which solves the single DGLAP equation in the scale $\mu$ with the initial condition $D(x, \mu_y; \mu_y) = \delta(1 - x)$. For simplicity we omit labels for and sums over parton species. We then have

$$F_{\Phi,\text{spl}}(x_i, \boldsymbol{\Delta}; \mu_i, \nu) \approx \int_{b_0^2/\nu^2}^{\infty} \frac{dy^2}{y^2} \, J_0(y\Delta) E(x_i; \mu_i, \mu_y) \tag{8.7}$$

with

$$E(x_i; \mu_i, \mu_y) = \int_{x_1}^{1-x_2} \frac{dz_1}{z_1} \int_{x_2}^{1-z_1} \frac{dz_2}{z_2} \, D\left(\frac{x_1}{z_1}, \mu_1; \mu_y\right) D\left(\frac{x_2}{z_2}, \mu_2; \mu_y\right)$$
$$\times \frac{\alpha_s(\mu_y)}{2\pi} P\left(\frac{z_1}{z_1 + z_2}, 0\right) \frac{f(z_1 + z_2, \mu_y)}{z_1 + z_2} . \tag{8.8}$$

The function $E(x_i; \mu_i, \mu_y)$ depends weakly on $y$ due to logarithmic corrections from evolution and the running of $\alpha_s$. To make this explicit, we expand the quantities depending on $\mu_y$ (the Green functions, $\alpha_s$ and the parton density) around $\nu$. Keeping only leading logarithms, we obtain

$$E(x_i; \mu_i, \mu_y) = \sum_{n=0}^{\infty} c^{(n)}(x_i, \mu_i, \nu) \, \alpha_s^{n+1}(\nu) \, \log^n\left(\frac{y\nu}{b_0}\right). \tag{8.9}$$

We show in appendix A that with such a $y$ behaviour, the Bessel function $J_0(y\Delta)$ in (8.7) can be replaced by an upper cutoff $\Theta(b_0 - y\Delta)$, up to relative corrections in $\Delta^2/\nu^2$ and up to subleading logarithms $\alpha_s^{n+1}(\nu) \log^{n-2}(\nu/\Delta)$. We then have

$$F_{\Phi,\text{spl}}(x_i, \boldsymbol{\Delta}; \mu_i, \nu) \approx \int_{b_0^2/\nu^2}^{b_0^2/\Delta^2} \frac{dy^2}{y^2} \, E(x_i; \mu_i, \mu_y) = \int_{\Delta^2}^{\nu^2} \frac{d\mu^2}{\mu^2} \, E(x_i; \mu_i, \mu) \tag{8.10}$$

for $\Delta \ll \nu$. With the scale choice (6.1) our result (8.10) corresponds to the expressions of the perturbative splitting DPD in equation (18) of [12] and in equation (8) of [15]. Note that at leading logarithmic accuracy, the lower integration limit on $k$ in [12] is equivalent to $\Delta$ in the region $\Delta \sim \Lambda$ that dominates the 2v1 cross section.

To leading logarithmic accuracy, the results for $\sigma_{2v1}$ obtained by both BDFS [12, 13] and RS [14, 15] are therefore consistent with our approach (provided of course we make a

suitable ansatz for $F_{\text{int}}$). In equation (6) of [15] an upper cutoff $\min(Q_1, Q_2)$ is put on the $\Delta$ integration, but this is of no concern since as discussed above, the 2v1 term is dominated by $\Delta \sim \Lambda$.

The situation is different for $\sigma_{1\text{v}1}$, which RS compute using the form (8.10) we just derived for $\Delta \ll \nu$. The integral over $\Delta$ in the cross section then requires an upper cutoff, which RS choose as $\min(Q_1, Q_2)$, as we do for our cutoff $\nu$ in $y$ space. The dominant integration region for the 1v1 cross section is $\Delta \sim \nu$ in their approach and in ours, given that up to logarithms the splitting DPD is of order unity for both $\Delta \sim \Lambda$ and $\Delta \sim \nu$. For $\Delta \sim \nu$ our distribution $F_{\Phi,\text{spl}}$ differs from (8.10) even if we neglect evolution effects, which readily follows from the result (A.3) for the $y$ integral in (8.7). Furthermore, $F_{\Phi,\text{spl}}$ strongly depends on the choice of $\Phi$ in that region. This just highlights that $\sigma_{1\text{v}1}$ has no physical meaning by itself. In our formalism, the contribution from $\Delta \sim \nu$ in $\sigma_{1\text{v}1}$ is cancelled by the subtraction term $\sigma_{1\text{v}1,\text{pt}}$ , as it should be since the DPS approximation is not valid in that region at all.

The DPS cross section in the approach of BDFS is given entirely by 2v2 and 2v1 terms. To leading logarithmic accuracy, these terms agree with the ones in our approach. The term $\sigma_{\text{DPS}} - \sigma_{1\text{v}1,\text{pt}}$ in our approach has an additional contribution $\sigma_{1\text{v}1} - \sigma_{1\text{v}1,\text{pt}}$, which as already mentioned may be neglected *unless* DPD evolution enhances the region $y \gg 1/Q$ (or equivalently $\Delta \ll Q$) in $\sigma_{1\text{v}1}$. This possible enhancement is not captured by the BDFS approach.

In the RS approach, the DPS cross section includes 2v2, 2v1 and 1v1 terms, the first two of which agree with ours in the leading logarithmic approximation. If DPD evolution effects are moderate, then the 1v1 term of RS is dominated by large $\Delta \sim Q$, strongly depends on the upper cutoff imposed on the $\Delta$ integral, and leads to a double counting problem with $\sigma_{\text{SPS}}$ that is not solved in the RS approach. In the sense we have discussed in section 6.2, one may at best take the cutoff dependence of the 1v1 cross section as an estimate for the size of $\sigma_{\text{SPS}}$ and thus as an indicator for how serious the double counting problem is. Only if DPD evolution enhances the region $\Delta \ll Q$ so strongly that it *dominates* the 1v1 term does the RS approach give a valid representation of the overall cross section. The cutoff dependence of the 1v1 cross section is then reduced, as was emphasised in [14].

## 9 Quantitative illustrations

We have set up a general formalism for calculating and combining the contributions from SPS, DPS and other mechanisms to the physical cross section. Whilst a detailed phenomenological application is beyond the scope of this work, we wish to give some quantitative illustrations of the formalism. This will allow us to estimate the relative size of contributions in specific channels and kinematics, and to explore the interplay between SPS or DPS terms and the double counting subtractions. We limit ourselves to collinear factorisation, where the construction of a model for DPDs is much simpler than in the TMD case.

Convenient quantities for our purpose are double parton luminosities

$$\mathcal{L}_{a_1 a_2 b_1 b_2}(x_i, \bar{x}_i, \mu_i, \nu) = \int d^2\boldsymbol{y}\, \Phi^2(y\nu)\, F_{a_1 a_2}(x_i, \boldsymbol{y}; \mu_i)\, F_{b_1 b_2}(\bar{x}_i, \boldsymbol{y}; \mu_i)\,, \qquad (9.1)$$

which directly appear in the DPS cross section. For definiteness, we use a sharp cutoff $\Phi(u) = \Theta(u - b_0)$ to remove the UV region. As a model for colour singlet DPDs we take the sum of an intrinsic and a splitting part, as specified in (8.3) and in the paragraph following that equation. The luminosity (9.1) then naturally splits into separate terms for 2v2, 2v1 and 1v1. Colour non-singlet channels will not be considered, except in subsection 9.3, where they can be computed without recourse to a model.

## 9.1 Simplified analytic estimates

We shall see that DPD evolution has important quantitative effects on double parton luminosities. Nevertheless, let us start by deriving simple analytic expressions where evolution is neglected, but where the power behaviour, logarithms and other important factors can be easily identified.

For the intrinsic contribution we make the well-known product ansatz

$$F_{\text{int}}(x_i, \boldsymbol{y}) = \frac{\Lambda^2}{\pi}\, e^{-y^2 \Lambda^2}\, f(x_1)\, f(x_2) \qquad (9.2)$$

with a Gaussian form for the $y$ dependence. For the splitting contribution we take

$$F_{\text{spl}}(x_i, \boldsymbol{y}) = \frac{1}{\pi y^2}\, e^{-y^2 \Lambda^2}\, \frac{f(x_1 + x_2)}{x_1 + x_2}\, \frac{\alpha_s}{2\pi}\, P\!\left(\frac{x_1}{x_1 + x_2}\right), \qquad (9.3)$$

which is the perturbative expression at $O(\alpha_s)$ times an exponential to suppress the large $y$ region. Sums and labels for the parton type are omitted here. For simplicity we use the same exponential in $F_{\text{int}}$ and $F_{\text{spl}}$. Since we ignore evolution for the time being, no scale dependence is taken into account in (9.2) and (9.3). To simplify even further, let us take equal momentum fractions $x_i = \bar{x}_i = x$ in both protons.

The parton luminosity for the 2v2 contribution then is

$$\mathcal{L}_{\text{2v2}} = \frac{\Lambda^4}{\pi}\, \big[f(x)\big]^4 \int_{b_0^2/\nu^2}^{\infty} dy^2\, e^{-2y^2 \Lambda^2} = \frac{\Lambda^2}{2\pi}\, e^{-2 b_0^2 \Lambda^2 / \nu^2}\, \big[f(x)\big]^4$$

$$= \frac{\Lambda^2}{2\pi}\, \big[f(x)\big]^4 + \mathcal{O}(\Lambda^4 / \nu^2)\,. \qquad (9.4)$$

As mentioned earlier, the UV regulator $\Phi(y\nu)$ only gives rise to a power suppressed effect. The 2v1 luminosity reads

$$\mathcal{L}_{\text{2v1}} = \frac{2\Lambda^2}{\pi}\, \big[f(x)\big]^2\, \frac{f(2x)}{2x}\, \frac{\alpha_s}{2\pi}\, P(1/2) \int_{b_0^2/\nu^2}^{\infty} \frac{dy^2}{y^2}\, e^{-2y^2 \Lambda^2}$$

$$= \frac{2\Lambda^2}{\pi}\, \log\!\left(\frac{\nu^2}{4 b_0 \Lambda^2}\right) \frac{\alpha_s}{2\pi}\, \big[f(x)\big]^2\, \frac{f(2x)}{2x}\, P(1/2) + \mathcal{O}(\Lambda^4 / \nu^2)\,. \qquad (9.5)$$

This is enhanced over $\mathcal{L}_{2\mathrm{v}2}$ by $\log(\nu^2/\Lambda^2)$ but suppressed by $\alpha_s/(2\pi)$. Furthermore, there are different factors regarding the parton densities. We note that the size of the splitting function strongly depends on the parton channel, as is exemplified by the cases $P_{g\to q\bar{q}}(1/2) = 1/4$, $P_{q\to qg}(1/2) = 10/3$ and $P_{g\to gg}(1/2) = 27/2$. For the 1v1 luminosity we finally obtain

$$\mathcal{L}_{1\mathrm{v}1} = \frac{1}{\pi}\left[\frac{f(2x)}{2x}\frac{\alpha_s}{2\pi}P(1/2)\right]^2 \int_{b_0^2/\nu^2}^{\infty} \frac{dy^2}{y^4} e^{-2y^2\Lambda^2}$$

$$= \frac{\nu^2}{\pi b_0^2}\left(\frac{\alpha_s}{2\pi}\right)^2 \left[\frac{f(2x)}{2x}P(1/2)\right]^2 + \mathcal{O}\left(\Lambda^2\log(\nu^2/\Lambda^2)\right). \tag{9.6}$$

We see the power enhancement over the 2v2 and 2v1 terms, as well as the quadratic dependence on the cutoff $\nu$.

It is easy to repeat the above exercise for our alternative regulator function $\Phi(u) = 1 - e^{-u^2/4}$. The integrals can be performed writing $\Phi(u) = u^2 \int_0^{1/4} d\tau\, e^{-\tau u^2}$. For the leading term of the expansion in $\Lambda^2/\nu^2$, one obtains the same result for $\mathcal{L}_{2\mathrm{v}2}$ as in (9.4), whereas in (9.5) and (9.6) one should replace

$$\log\left(\frac{\nu^2}{4b_0\Lambda^2}\right) \to \log\left(\frac{\nu^2}{16\Lambda^2}\right), \qquad\qquad \frac{\nu^2}{b_0^2} \to \frac{\nu^2\log 2}{2}, \tag{9.7}$$

respectively. The leading logarithm $\log(\nu^2/\Lambda^2)$ in $\mathcal{L}_{2\mathrm{v}1}$ is independent of the choice for $\Phi(u)$, but not the constant term. The corresponding ambiguity is compensated by the DPS subtraction term, as is the change of the numerical prefactor in $\mathcal{L}_{1\mathrm{v}1}$.

## 9.2 Collinear Parton Luminosities

We now present a numerical study of the parton luminosities, using a slightly refined model for the input DPDs, and including evolution effects. We start by considering the unpolarised luminosities only, turning to the issue of polarised DPDs and luminosities at the very end of the section.

### 9.2.1 Setup

We use scale evolution at leading order in $\alpha_s$ and take equal scales $\mu_1 = \mu_2 = \mu$ for the two hard processes throughout. In this illustrative study, we fix the number of active quark flavours to $n_f = 3$. This is sufficient for exploring the broad qualitative features of the DPDs and luminosities. Implementation of heavy quark flavours, as required for a realistic phenomenological investigation, would require some dedicated work on flavour thresholds and on memory management in our numerical code, which we postpone to future studies.

As in section 9.1, we write each DPD as a sum of intrinsic and splitting pieces:

$$F_{a_1 a_2}(x_1, x_2, \boldsymbol{y}; \mu, \mu) = F_{a_1 a_2,\mathrm{int}}(x_1, x_2, \boldsymbol{y}; \mu, \mu) + F_{a_1 a_2,\mathrm{spl}}(x_1, x_2, \boldsymbol{y}; \mu, \mu). \tag{9.8}$$

We initialise $F_{\mathrm{int}}$ at a low scale $\mu_0 = 1\,\mathrm{GeV}$ using a product ansatz similar to (9.2):

$$F_{a_1 a_2,\mathrm{int}}(x_1, x_2, \boldsymbol{y}; \mu_0, \mu_0) = \frac{1}{4\pi h_{a_1 a_2}}\exp\left[-\frac{y^2}{4h_{a_1 a_2}}\right] f_{a_1}(x_1; \mu_0)\, f_{a_2}(x_2; \mu_0)$$

$$\times (1 - x_1 - x_2)^2 (1 - x_1)^{-2}(1 - x_2)^{-2}. \tag{9.9}$$

For the single PDFs $f_a(x, \mu_0)$, we take the MSTW2008LO distributions [48]. In (9.9) we have multiplied the PDFs by a function of the $x_i$ that does not affect the DPD at small $x_i$, but smoothly cuts it off near the kinematic boundary $x_1 + x_2 = 1$. The function we use is that given in equation (3.12) of [9], with $n = 2$. We also take a different $y$ dependence for different parton species. For this we use a simplified version of the model in section 4.1 of [39], taking the width $h$ to be $x$ independent (corresponding to $h(x_1, x_2)$ of [39] evaluated at $x_1 = x_2 = 10^{-3}$) and setting each $h$ with $q^-$ indices to be the same as the one with $q^+$. Thus we have

$$h_{a_1 a_2} = h_{a_1} + h_{a_2} \tag{9.10}$$

with

$$h_g = 2.33 \, \text{GeV}^{-2} , \qquad\qquad h_q = h_{\bar{q}} = 3.53 \, \text{GeV}^{-2} . \tag{9.11}$$

For the splitting piece of the DPD we generalise our ansatz in (9.3), choosing an initialisation scale that goes to $b_0/y$ at small $y$ but freezes to a constant value $b_0/y_{\text{max}}$ when $y$ exceeds a value $y_{\text{max}}$ that marks the transition between perturbative and nonperturbative behaviour. This ensures that the single PDF and $\alpha_s$ in the splitting expression are never evaluated at too low scales. A suitable choice of scale is

$$\mu_y = \frac{b_0}{y^*} , \qquad\qquad y^* = \frac{y}{\sqrt{1 + y^2/y_{\text{max}}^2}} . \tag{9.12}$$

Such a prescription is very similar to the $b^*$ prescription used in TMD phenomenology [38, 49]. Here we take $y_{\text{max}} = 0.5 \, \text{GeV}^{-1}$, which is one of the values considered in the recent TMD study [50]. Using the same parton dependent Gaussian damping as in (9.13), we have

$$
\begin{aligned}
&F_{a_1 a_2, \text{spl}}(x_1, x_2, \boldsymbol{y}; \mu_y, \mu_y) \\
&= \frac{1}{\pi y^2} \exp\left[-\frac{y^2}{4 h_{a_1 a_2}}\right] \frac{f_{a_0}(x_1 + x_2; \mu_y)}{x_1 + x_2} \frac{\alpha_s(\mu_y)}{2\pi} P_{a_0 \to a_1 a_2}\left(\frac{x_1}{x_1 + x_2}\right) .
\end{aligned}
\tag{9.13}
$$

The coupling $\alpha_s(\mu_y)$ is determined by 3 flavour running, starting with the MSTW2008LO value $\alpha_s(\mu_0) = 0.68183$. The single PDFs are obtained by taking the MSTW2008LO distributions at $\mu_0$ and evolving them according to the DGLAP equations for $n_f = 3$.

To obtain the splitting and intrinsic DPDs at scale $\mu$, as in (9.8), the input forms just discussed must be evolved, starting from $\mu_0$ for $F_{\text{int}}$ and from $\mu_y$ for $F_{\text{spl}}$, according to the homogeneous double DGLAP equations. For this we use a modified version of the code developed in [9]. The modified code works on a grid in the $x_i$, $\mu$ and $\mu_y$ directions (the grid of the original code is in $x_i$ and $\mu$ only). The grid points in the $x_i$ directions are evenly spaced in $\log(x_i/(1 - x_i))$, whilst those in the $\mu$ and $\mu_y$ directions are evenly spaced in $\log \mu$ or $\log \mu_y$. The integrals appearing in the double DGLAP equations are computed from points in the $x_i$ grids using Newton-Cotes rules (for details see Appendix A of [9]), and evolution from one point in the $\mu$ grid to the next is carried out using the Runge-Kutta method.

The grid is set up with 88 points in each $x_i$ direction, spanning $5 \times 10^{-5} < x_i < 1$, with 60 points in the $\mu$ direction, spanning $\mu_0 < \mu < 170\,\text{GeV}$, and with 60 points in the $\mu_y$ direction, spanning $b_0/y_{\text{max}} < \mu_y < 340\,\text{GeV}$. According to the studies made in [9], this suggests an error on the level of a few per cent in the DPD values obtained after evolution, which is tolerable in this first study. The DPDs computed on this grid are used together with an interpolation code to produce numerical values for the investigations below.

### 9.2.2 Numerical results

We begin with a study where the scale $\mu$ is equal to $Q_1 = Q_2 = 80\,\text{GeV}$ (as in the production of a $W$ boson pair). Taking the collider energy to be $\sqrt{s} = 14\,\text{TeV}$, we set $x_1$ and $\bar{x}_1$ in (9.1) to correspond to central production of the first system and $x_2$ and $\bar{x}_2$ to correspond to the production of the second system with rapidity $Y$ (all rapidities refer to the $pp$ centre of mass). This gives

$$x_1 = \bar{x}_1 = 5.7 \times 10^{-3}\,, \quad x_2 = 5.7 \times 10^{-3}\exp(Y)\,, \quad \bar{x}_2 = 5.7 \times 10^{-3}\exp(-Y)\,. \quad (9.14)$$

In figure 12 we plot $\mathcal{L}_{a_1a_2b_1b_2}(Y)$ in the range $0 \leq Y \leq 4$ for the parton combinations $a_1a_2b_1b_2 = u\bar{u}\bar{u}u + \bar{u}uu\bar{u}$, $a_1a_2b_1b_2 = gggg$ and $a_1a_2b_1b_2 = u\bar{d}\bar{d}u + \bar{d}uu\bar{d}$. The first parton combination appears e.g. in $ZZ$ production, the second is important in four-jet production, and the last appears in $W^+W^+$. For ease of language, we will refer to these parton combinations as the $u\bar{u}$, $gg$ and $u\bar{d}$ channels, respectively. We split the overall luminosity into contributions from 1v1 ($F_{\text{spl}} \times F_{\text{spl}}$), from 2v1 ($F_{\text{spl}} \times F_{\text{int}} + F_{\text{int}} \times F_{\text{spl}}$) and from 2v2 ($F_{\text{int}} \times F_{\text{int}}$). We vary $\nu$ by a factor of 2 up and down around a central value of 80 GeV, in order to see how DPS alone is affected by variation of this cutoff. For each contribution, the line in the plots denotes the luminosity with $\nu = 80\,\text{GeV}$, whilst the band is generated from the envelope of the functions with $\nu = 40\,\text{GeV}$, $\nu = 80\,\text{GeV}$ and $\nu = 160\,\text{GeV}$. In the case of the 2v2 contribution, the $\nu$ scale variation is negligible (as expected from basic considerations, see (9.4)), so this appears as a dashed line in each plot.

For the 1v1 contribution of the $gg$ and $u\bar{u}$ channels, we also plot a band generated by varying the prediction with $\nu = 80\,\text{GeV}$ by a factor 4 up and down. This corresponds to a strictly quadratic cutoff dependence of $\mathcal{L}$, i.e. to the variation of $\nu$ by a factor 2 in the naive formula (9.6), where DPD evolution is neglected. Any discrepancy between this band and the actual 1v1 band is therefore due to evolution effects. We do not plot this band for the $u\bar{d}$ channel: there is no LO splitting process giving $u\bar{d}$, so that the scale variation (and central value) from (9.6) is zero in this case.

We immediately notice in figures 12a and b that the 1v1 contribution is generally much larger than the 2v2 and 2v1 contributions, and that it has an enormous $\nu$ variation. To obtain a sensible prediction in these cases, one must include the SPS corrections up to the order that includes the double box graph in figure 1a, so that the associated subtraction term can approximately cancel the $\nu$ dependence of DPS. We also notice that at central rapidities in the $u\bar{u}$ channel, the 1v1 $\nu$ variation band essentially fills the band corresponding to a quadratic $\nu$ dependence, indicating small evolution effects in this case. By contrast, the 1v1 $\nu$ variation band for $gg$ is clearly smaller than the naive expectation, which indicates significant evolution effects. We shall explore the reasons for these differences below.
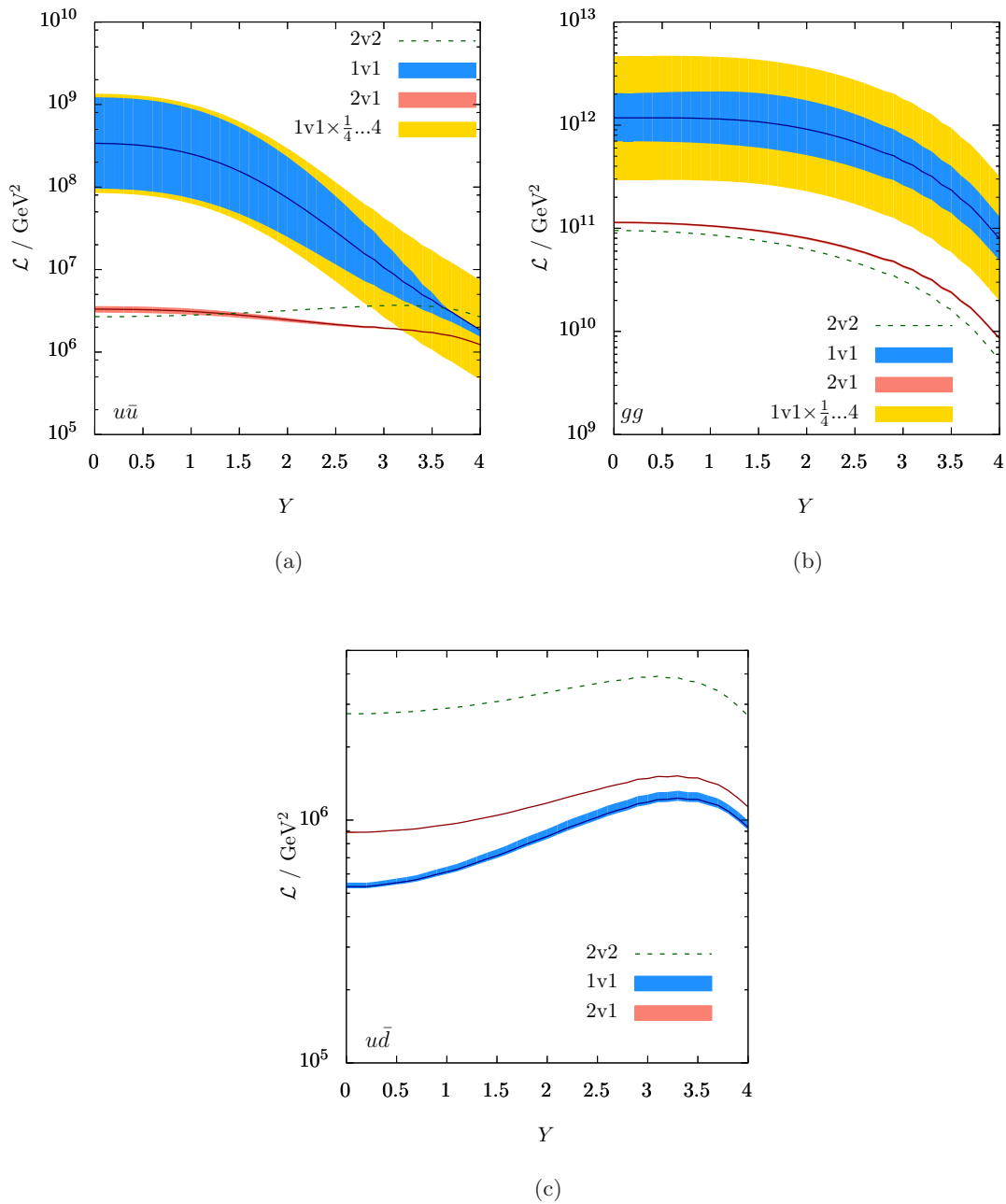
**Figure 12**. Double parton scattering luminosities $\mathcal{L}_{a_1 a_2 b_1 b_2}(Y)$ for the production of two systems with $Q_1 = Q_2 = 80\,\text{GeV}$ at $\sqrt{s} = 14\,\text{TeV}$, one with central rapidity and the other with rapidity $Y$. The parton combinations $a_1 a_2 b_1 b_2$ are $u\bar{u}\bar{u}u + \bar{u}uu\bar{u}$ (a), $gggg$ (b) and $u\bar{d}\bar{d}u + \bar{d}uu\bar{d}$ (c).

In figure [12]c, the 1v1 contribution is small compared to the 2v2 piece and has a small $\nu$ dependence. This is because, as already mentioned, there is no LO splitting directly giving $u\bar{d}$ (generation of a $u\bar{d}$ pair requires at least two steps, such as $u \to u + g \to u + d + \bar{d}$). In this case, there is less of a need to compute $\sigma_{\text{SPS}}$ and the subtraction term $\sigma_{\text{1v1,pt}}$ up to

the order that contains the lowest-order DPS-type loop (in both amplitude and conjugate). This is fortunate, since this order is two powers of $\alpha_s$ higher than for graph 1a (two-step rather than one-step splittings are required in both protons), and the corresponding SPS calculation will not be available for some time.

In the $gg$ channel, the 2v1 and 2v2 contributions are of similar magnitude and shape. This is consistent with what has been found in previous phenomenological studies [13, 51] at similar scales and $x$ values. In the $u\bar{u}$ channel, the magnitude of 2v1 and 2v2 is broadly the same but their shape is somewhat different. In the $u\bar{d}$ channel, the shape of the two contributions is similar, but 2v1 lies well below 2v2 (and is close to 1v1).

We can examine the degree to which the $\nu$ variation in the $gg$ and $u\bar{u}$ channels is reduced by adding the remaining fixed-order terms in the cross section, even without considering a specific final state or having to compute $\sigma_{\text{SPS}}$. This is because only $\sigma_{\text{DPS}}$ and the subtraction terms depend on $\nu$. Just as we did for $\sigma_{\text{DPS}}$, we can factor the two parton-level cross sections out of $\sigma_{1v1,\text{pt}}$ and introduce a "subtraction luminosity" $\mathcal{L}_{1v1,\text{pt}}$. This is defined as in (9.1) except that the DPDs are replaced by the fixed-order splitting expression evaluated at scale $\mu$,

$$F_{a_1 a_2,\text{spl,fo}}(x_1, x_2, \boldsymbol{y}; \mu) = \frac{1}{\pi y^2} \frac{f_{a_0}(x_1 + x_2; \mu)}{x_1 + x_2} \frac{\alpha_s(\mu)}{2\pi} P_{a_0 \to a_1 a_2}\left(\frac{x_1}{x_1 + x_2}\right). \qquad (9.15)$$

For a given parton combination $a_1 a_2 b_1 b_2$, one can directly subtract $\mathcal{L}_{1v1,\text{pt}}$ from the DPS luminosity $\mathcal{L}_{\text{DPS}}$ and compare the $\nu$ dependence of the result to that of $\mathcal{L}_{\text{DPS}}$ alone. Up to multiplication by the appropriate parton-level cross sections, the $\nu$ dependence of $\mathcal{L}_{\text{DPS}} - \mathcal{L}_{1v1,\text{pt}}$ is equal to that of the full cross section. On the other hand, the overall magnitude of $\mathcal{L}_{\text{DPS}} - \mathcal{L}_{1v1,\text{pt}}$ (which can be positive or negative) has no direct significance, since one needs to add the SPS contribution to obtain the physical cross section.

We plot $\mathcal{L}_{\text{DPS}}$ and $\mathcal{L}_{\text{DPS}} - \mathcal{L}_{1v1,\text{pt}}$ for the $u\bar{u}$ and $gg$ channels in figure 13. The lines are the values for $\nu = 80\,\text{GeV}$, whilst the bands show the variation when $\nu$ is varied by a factor of 2 up and down. Note that for $\mathcal{L}_{\text{DPS}} - \mathcal{L}_{1v1,\text{pt}}$, the curve with $\nu = 80\,\text{GeV}$ is always at the very top of the scale variation band. We see that the $\nu$ scale variation of $\mathcal{L}_{\text{DPS}} - \mathcal{L}_{1v1,\text{pt}}$ is indeed reduced compared to $\mathcal{L}_{\text{DPS}}$, with the reduction being much stronger for $u\bar{u}$ than for $gg$. The latter is consistent with our previous observation that evolution effects are weaker in the $u\bar{u}$ channel at central rapidities: if evolution effects are weak, $F_{\text{spl}}$ and $F_{\text{spl,fo}}$ have a similar $y$ dependence, so that $(F_{\text{spl}} \times F_{\text{spl}} - F_{\text{spl,fo}} \times F_{\text{spl,fo}})$ is flat in $y$ space and $\mathcal{L}_{\text{DPS}} - \mathcal{L}_{1v1,\text{pt}}$ varies weakly with $\nu$.

In section 6.2 we argued on generic grounds that the $\nu$ variation of the 1v1 contribution to $\sigma_{\text{DPS}}$ should be of the same order as the subtraction term $\sigma_{1v1,\text{pt}}$. This is confirmed in figure 13, where the size of $\mathcal{L}_{1v1,\text{pt}}$ (evaluated at central $\nu$) can be read off from the distance between the central curve for $\mathcal{L}_{\text{DPS}}$ and the top of the band for $\mathcal{L}_{\text{DPS}} - \mathcal{L}_{1v1,\text{pt}}$. This finding allows us to sharpen the argument we already made in the discussion of figure 12: not only does a large $\nu$ variation of the 1v1 contribution to $\sigma_{\text{DPS}}$ indicate the need to include SPS and the associated subtraction term in the cross section, but the size of the $\nu$ variation may even serve as a rough estimate for the size of these terms.
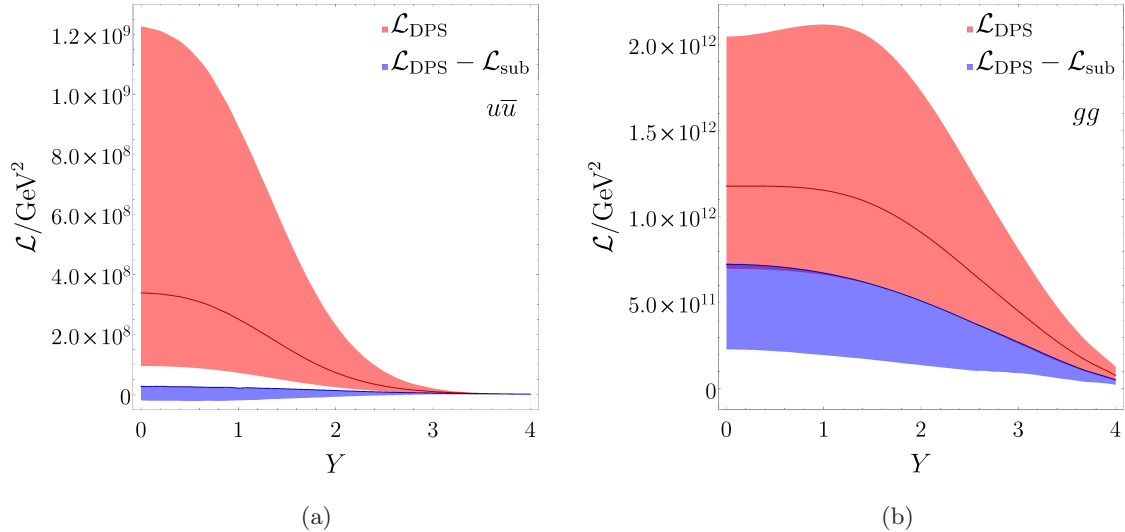
**Figure 13**. DPS luminosity $\mathcal{L}_{\text{DPS}}$ and the difference $\mathcal{L}_{\text{DPS}} - \mathcal{L}_{\text{1v1,pt}}$ between the DPS and subtraction luminosities, shown for the $u\bar{u}$ (a) and $gg$ (b) channels. Kinematic conditions are as in figure 12. In the case of $\mathcal{L}_{\text{DPS}} - \mathcal{L}_{\text{1v1,pt}}$, the value for $\nu = 80\,\text{GeV}$ is always at the top of the scale variation band.

To better understand the behaviour we have seen in the 1v1 luminosities for the $u\bar{u}$ and $gg$ channels, we now take a closer look at how evolution affects the $y$ dependence of splitting DPDs. In figure 14 we plot $F_{\text{spl}}(\boldsymbol{y}; \mu)$ against $y$, setting $x_1 = x_2 = 5.7 \times 10^{-3}$ so that the $u\bar{u}$ and $gg$ DPDs are the ones used for the point $Y = 0$ in figures 12 and 13. We also plot the $ug$ DPD, which mixes with $u\bar{u}$ and $gg$ under evolution. For comparison we show $F_{\text{spl,fo}}(\boldsymbol{y}; \mu)$, as well as the initial condition $F_{\text{spl,ini}}(\boldsymbol{y}) = F_{\text{spl}}(\boldsymbol{y}; \mu_y)$ given in (9.13), from which $F_{\text{spl}}(\boldsymbol{y}; \mu)$ is obtained by evolution. The distributions are plotted in the range $b_0/(160\,\text{GeV}) < y < b_0/(40\,\text{GeV})$, i.e. in the range over which the lower integration limit in $y$ is varied when we vary $\nu$ between $40\,\text{GeV}$ to $160\,\text{GeV}$. In this region, the exponential damping factor in our DPD model is irrelevant, so that $F_{\text{spl,ini}}$ and $F_{\text{spl,fo}}$ only differ by the scales taken in $\alpha_s$ and $f_{a_0}(x_1 + x_2)$.

For the initial conditions $F_{\text{spl,ini}}$ we note that the $u\bar{u}$ and $ug$ distributions are of similar size; the former is initialised by a larger PDF ($f_g$ instead of $f_u$) but has a smaller splitting coefficient $P(1/2)$ as we noted before (9.6). By contrast, the $gg$ distribution is much bigger; here both the initialising PDF and the splitting coefficient are large.

By construction, all three curves in each plot are equal at $y = b_0/(80\,\text{GeV})$, when $\mu_y = \mu$. In all plots, $F_{\text{spl,ini}}$ has a more shallow $y$ behaviour than $F_{\text{spl,fo}}$. This difference is mainly driven by $\alpha_s$, as the PDFs do not vary so much between $\mu_y$ and $\mu$ at momentum fraction $x_1 + x_2 = 1.14 \times 10^{-2}$. One observes that the difference between $F_{\text{spl}}$ and $F_{\text{spl,ini}}$ is more significant for $gg$ and $ug$ than for $u\bar{u}$, i.e. that DPD evolution has a much stronger effect on the former two channels. This is to be expected at small $x$, since the $1/v$ behaviour of the splitting kernels $P_{gg}(v)$ and $P_{gq}(v)$ at small $v$ favours the radiation of a gluon with much smaller momentum than its parent parton, whereas the kernels $P_{qq}(v)$ and $P_{qg}(v)$
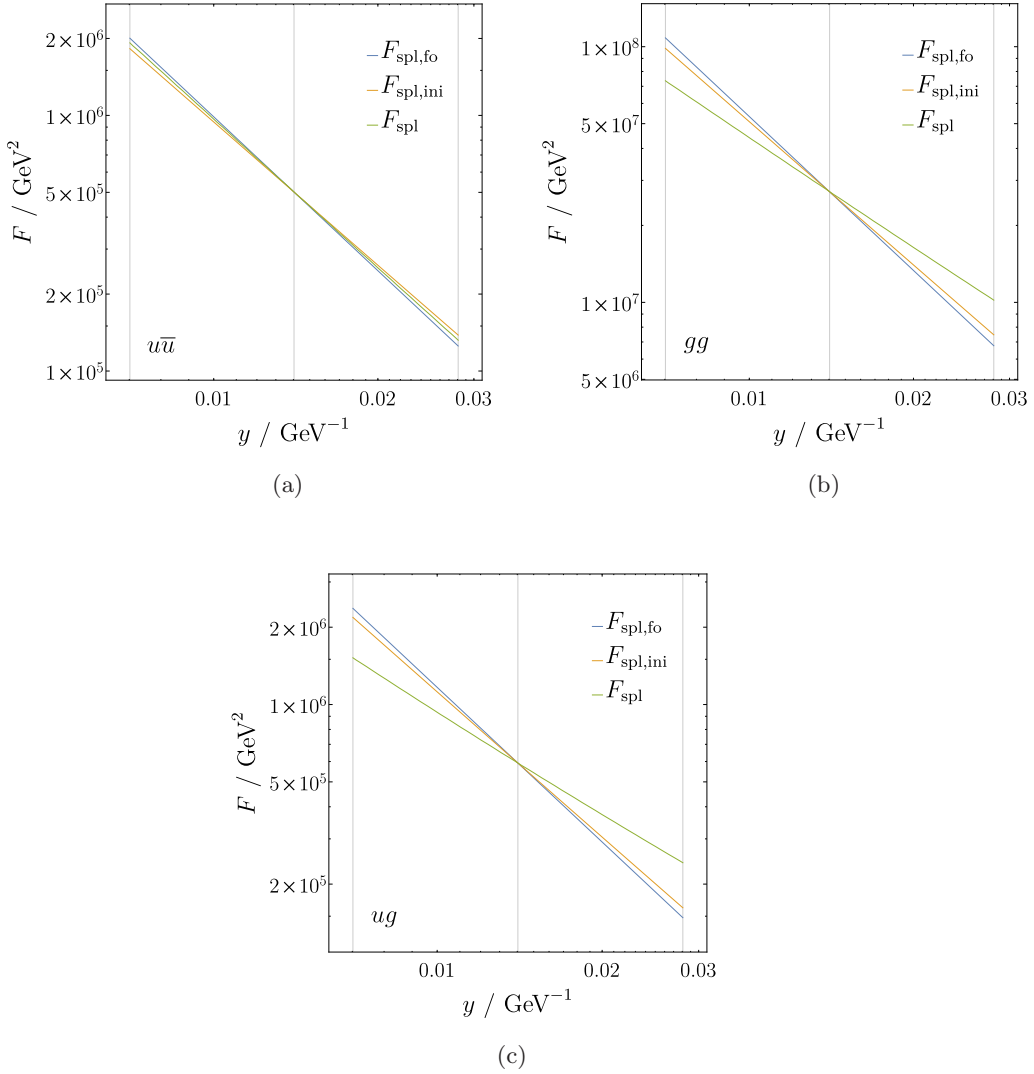
**Figure 14**. Splitting DPDs $F_{\text{spl}}$ as a function of $y$, with $x_1 = x_2 = 5.7 \times 10^{-3}$ and $\mu = 80\,\text{GeV}$. The parton combinations shown are $u\bar{u}$ (a), $gg$ (b) and $ug$ (c). The vertical grey lines correspond to $y = b_0/(160\,\text{GeV})$, $y = b_0/(80\,\text{GeV})$ and $y = b_0/(40\,\text{GeV})$.

giving a quark stay finite for $v \to 0$.

An interesting point to note is that, whilst for $gg$ and $ug$ the curves for $F_{\text{spl}}$ are more shallow than for $F_{\text{spl,ini}}$, in the case of $u\bar{u}$ the $F_{\text{spl}}$ curve is actually steeper. The latter is surprising since — based on the experience with single PDFs at small $x$ — one may expect that forward evolution for $y > b_0/(80\,\text{GeV})$ would always increase a DPD, and backward evolution for $y < b_0/(80\,\text{GeV})$ would always decrease it. This indeed happens in the $gg$ and $ug$ channels, whilst forward evolution results in a decrease of the $u\bar{u}$ DPD. The reason for this is that $F_{ug,\text{spl}}$ and $F_{g\bar{u},\text{spl}}$, which directly feed into $F_{u\bar{u},\text{spl}}$ during evolution, are comparatively small. In the case of PDFs, $f_g$ is much larger than $f_u$ and hence can

drive its small-$x$ evolution although the splitting function $P_{qg}$ does not favour the radiation of low-momentum quarks. The evolution of the $gg$ and $ug$ DPDs is driven by the large distribution $F_{gg,\mathrm{spl}}$ and enhanced by the $P_{gg}$ splitting function.

Note that in all three channels, $F_{\mathrm{spl}}$ has a smaller slope in $y$ than $F_{\mathrm{spl,fo}}$. This implies that the $y$ integrand for the computation of $\mathcal{L}_{\mathrm{DPS}} - \mathcal{L}_{\mathrm{1v1,pt}}$ is positive for $y > b_0/(80\,\mathrm{GeV})$ and negative for $y < b_0/(80\,\mathrm{GeV})$. Therefore, the $\mathcal{L}_{\mathrm{DPS}} - \mathcal{L}_{\mathrm{1v1,pt}}$ curves for $\nu = 40\,\mathrm{GeV}$ and $\nu = 160\,\mathrm{GeV}$ lie below that for $\nu = 80\,\mathrm{GeV}$. The curve for the central $\nu$ value thus lies at the top of the $\nu$ variation band, as we already observed in figure 13.

For $u\bar{u}$, the evolution from $\mu_y$ to $\mu$ turns out to have much the same quantitative effect as adjusting the scale in the fixed order expression from $\mu_y$ to $\mu$, such that $F_{\mathrm{spl}}$ and $F_{\mathrm{spl,fo}}$ end up extremely close together. This explains why at central rapidities the $\nu$ variation of the 1v1 $u\bar{u}$ contribution in figure 12a coincides almost exactly with the naive prediction from (9.6), which assumes that $F_{\mathrm{spl}}$ depends on $y$ like $y^{-2}$ (as does $F_{\mathrm{spl,fo}}$). Furthermore, if $F_{\mathrm{spl}}$ and $F_{\mathrm{spl,fo}}$ are very close to each other, then $\mathcal{L}_{\mathrm{DPS}} - \mathcal{L}_{\mathrm{1v1,pt}}$ is much smaller than $\mathcal{L}_{\mathrm{DPS}}$, which we indeed see in figure 13a.

In the $gg$ channel, the modification of the $y$ slope by evolution is significant. In the range of figure 14, the $y$ dependence is changed from $y^{-2}$ at the starting scale to approximately $y^{-1.4}$. One may expect this flattening effect to become even stronger as the $x$ values in the DPDs decrease. The dependence of $\mathcal{L}_{\mathrm{DPS}}$ on the scale $\nu$ should then decrease. This was already anticipated in [14, 15], where studies in the double leading logarithm approximation were performed.

Here we investigate this effect using our full LO DGLAP set-up. We consider the $gg$ channel with all $x$ values set equal and study the DPS luminosity as a function of the common $x$ value. We fix the collider energy $\sqrt{s}$, so that the scale $\mu = Q_1 = Q_2 = x\sqrt{s}$ also varies with $x$. We make one plot with $\sqrt{s} = 14\,\mathrm{TeV}$ and $5 \times 10^{-4} < x < 0.2$ (corresponding to $7\,\mathrm{GeV} < \mu < 2800\,\mathrm{GeV}$) and another one with higher collider energy $\sqrt{s} = 100\,\mathrm{TeV}$ and $10^{-4} < x < 0.02$ (corresponding to $10\,\mathrm{GeV} < \mu < 2000\,\mathrm{GeV}$). In our numerical code this requires a new DPD grid with larger $\mu$ and $\mu_y$ ranges, which was generated using 60 points in the $\mu$ and $\mu_y$ directions (as in the original grid). The resulting 1v1 luminosity is shown in figure 15, with bands for the $\nu$ variation and its naive version as in figure 12. For comparison we also show the 2v1 and 2v2 luminosities.

We see that the $\nu$ variation does indeed become progressively smaller compared to the central value as $x$ (and $\mu$) decreases. At the lowest $x$ values it is much smaller than the naive expectation. At larger $x$, where evolution has a smaller effect on the DPDs, the $\nu$ variation becomes larger: for $\sqrt{s} = 14\,\mathrm{TeV}$ the actual and naive $\nu$ bands essentially coincide towards the right of the plot, where values of the order of $x \sim 0.1$ are reached. Towards the left of each plot, where $x$ and $\mu$ become small, the 2v2 and 2v1 contributions begin to dominate over 1v1. For given $\mu$, this effect is more pronounced for $\sqrt{s} = 100\,\mathrm{TeV}$ than for $\sqrt{s} = 14\,\mathrm{TeV}$, since the $x$ values in the former case are smaller. We note that in the $u\bar{u}$ channel (not shown here) the actual and naive $\nu$ bands remain very similar throughout the kinematics of figure 15.

For the small $x$ values where the $\nu$ variation is small in figure 15, the $gg$ splitting DPD is significantly shallower in $y$ than the fixed-order $y^{-2}$ form. For $\sqrt{s} = 14\,\mathrm{TeV}$ we find that
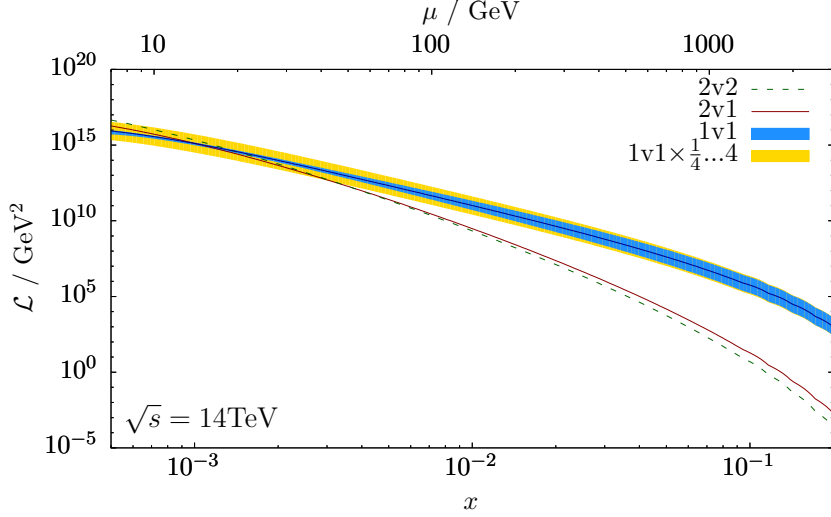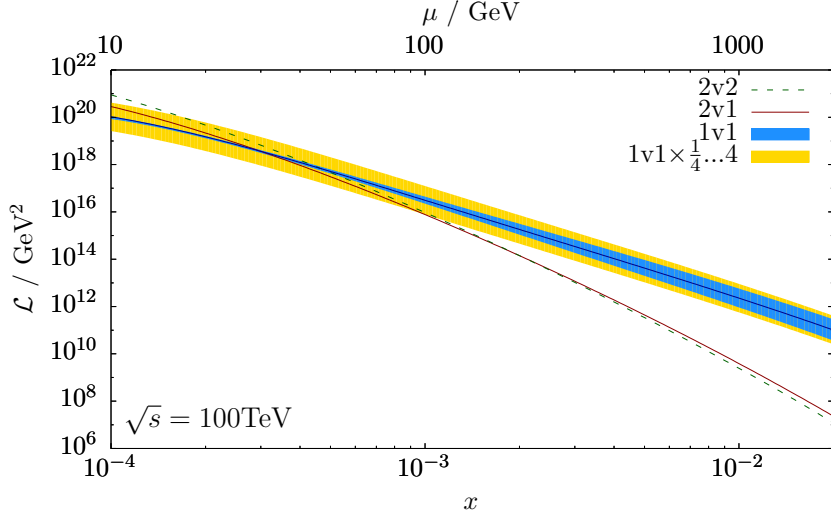
**Figure 15**. DPS luminosities in the $gg$ channel with all $x$ values in the DPDs set equal. The collider energy $\sqrt{s}$ is fixed to $14\,\text{TeV}$ (a) or to $100\,\text{TeV}$ (b). The scales of the DPDs, set to $\mu = x\sqrt{s}$, are given at the top of the plot.

around $x = 2 \times 10^{-3}$ it reaches $y^{-1}$, implying a logarithmic dependence on $\nu$ in $\mathcal{L}_{DPS}$, whilst a behaviour like $y^{-0.5}$ is reached around $5 \times 10^{-4}$. For $\sqrt{s} = 100\,\text{TeV}$ the $y$ behaviour is like $y^{-1}$ around $x = 4 \times 10^{-4}$ and like $y^{-0.5}$ around $x = 10^{-4}$. In such a regime, the bulk of the contribution to the 1v1 part of $\mathcal{L}_{\text{DPS}}$ comes from large distances $y \gg 1/\nu$, where the DPS approximations used to derive $\sigma_{\text{DPS}}$ are valid. Also, the 1v1 part of $\sigma_{\text{DPS}}$ becomes clearly larger than its $\nu$ variation. As we argued earlier, the $\nu$ variation should be of the same size as the subtraction term. This is confirmed in figure 16, which is analogous to figure 13b
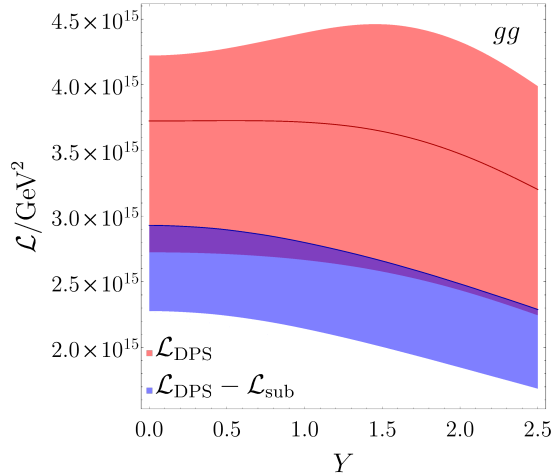
**Figure 16**. As in figure 13b, but for different kinematics, with $\mu = 10\,\mathrm{GeV}$ and all $x$ set equal to $7.1 \times 10^{-4}$, corresponding to $Q_1 = Q_2 = 10\,\mathrm{GeV}$ at $\sqrt{s} = 14\,\mathrm{TeV}$. Notice the suppressed zero on the $\mathcal{L}$ axis.

but for the kinematics of the point with $\mu = 10\,\mathrm{GeV}$ in figure 15a. Using our argument that $\sigma_{1v1,pt}$ is of the same magnitude as the order of $\sigma_{SPS}$ that contains the lowest-order DPS-type loop, one may in this situation justifiably make predictions that include the DPS piece but omit the order of SPS just specified, as well as the associated subtraction term. This is encouraging, for instance in the context of four-jet production, where the computation of the relevant SPS order (namely NNLO) is well beyond the current state of the art. Note that lower orders in SPS should be computed and included, if possible.

Notice that when the $y$ behaviour of $F_{spl}$ becomes flatter than $y^{-1}$, the dominant $y$ region in the 1v1 part of $\mathcal{L}_{DPS}$ shifts to values $y \sim 1/\Lambda$, where one cannot compute the splitting DPD in perturbation theory and must rely on a model. Likewise, the 2v1 part of $\mathcal{L}_{DPS}$ becomes increasingly sensitive to the region $y \sim 1/\Lambda$ as soon as $F_{spl}$ becomes flatter than the fixed-order form $y^{-2}$.

Another kinematic regime where the 1v1 contribution to $\mathcal{L}_{DPS}$ becomes large compared with its $\nu$ variation and compared with $\mathcal{L}_{1v1,pt}$ is when the two hard systems have a large separation in rapidity. This reduction of the $\nu$ variation can be seen as $Y$ approaches 4 in figure 12, but only in the $u\bar{u}$ and not in the $gg$ channel. In both channels, the effect becomes more pronounced once the rapidity separation of the hard systems is increased beyond 4. To illustrate this, we make plots similar to figure 12 but now with one hard system at rapidity $Y$ and the other at rapidity $-Y$ (rather than one at $Y$ and the other at 0), such that a given value of $Y$ corresponds to a rapidity separation of $2Y$. The results are shown in figure 17, where we see that the $\nu$ variation in the 1v1 contribution is strongly reduced towards the right hand side of the plots, becoming hardly visible in the $u\bar{u}$ channel. Also notable is the fact that for large $Y$, the 2v2 contribution in this channel exceeds 1v1, which strongly decreases between $Y = 0$ and 2.

This reduction in $\nu$ dependence can be explained by the fact that at large $Y$ we probe
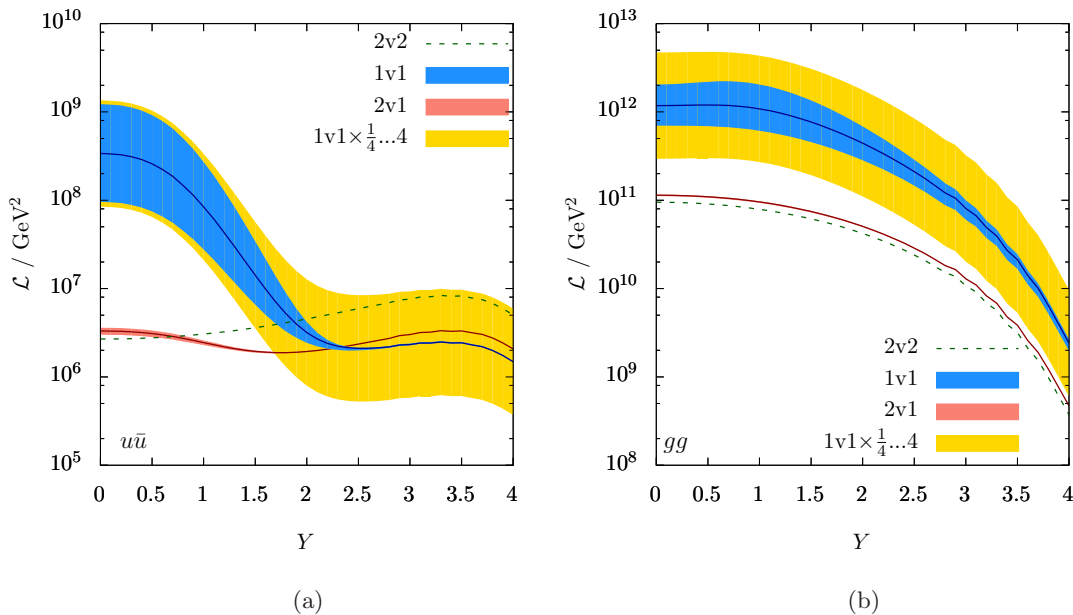
**Figure 17**. Double parton scattering luminosities $\mathcal{L}_{a_1 a_2 b_1 b_2}(Y)$ for the production of two systems with $Q_1 = Q_2 = 80\,\text{GeV}$ at $\sqrt{s} = 14\,\text{TeV}$, one with rapidity $Y$ and the other with rapidity $-Y$. The parton combinations $a_1 a_2 b_1 b_2$ are $u\bar{u}\bar{u}u + \bar{u}uu\bar{u}$ (a) and $gggg$ (b).

splitting DPDs with one large $x$ and one small $x$ parton. From the point of view of small $x$ logarithms, it is preferable to generate such a configuration by having the $1 \to 2$ splitting at large $x$, generating directly the large $x$ parton plus a gluon with smaller $x$, the latter of which splits in a number of stages into smaller $x$ gluons, eventually yielding the small $x$ parton. This increasingly happens with increasing $y$, since the "evolution distance" between the initial and final scales, $\mu_y$ and $\mu$, is increased. Thus, evolution again flattens the DPD compared to the naive $y^{-2}$ expectation and reduces the $\nu$ dependence of the DPS luminosities.

The effect is particularly drastic in the $u\bar{u}$ channel, because the lowest-order splitting $g \to u\bar{u}$ is inefficient at generating a pair with very different $x$ values (as $P_{g \to q\bar{q}}(v)$ goes to a constant at small $v$). Therefore, the repeated splitting mechanism described in the previous paragraph is strongly preferred, even though its last step is penalised by the lack of small-$v$ enhancement in $P_{gq}(v)$. In figure 18 we plot the $u\bar{u}$ DPD in the $x_1 \ll x_2$ configuration relevant for $Y = 4$ in figure 17a. It has a similar $y$ dependence as the $u\bar{d}$ DPD in equal kinematics, with the curve for $u\bar{d}$ crossing zero exactly at $y = b_0/\mu$ (by construction), and the curve for $u\bar{u}$ crossing zero rather close to this point. In the former case, the lowest-order $1 \to 2$ splitting process is forbidden, whereas in the latter it is numerically almost irrelevant. The situation for $x_1 \gg x_2$ is the same.

To end this section, we study polarised distributions and luminosities. We limit ourselves to the splitting part $F_{\text{spl}}$ of the DPD and to the 1v1 contribution to $\mathcal{L}_{\text{DPS}}$. In fact, we have little guidance for modelling the intrinsic part $F_{\text{int}}$, where a product ansatz as
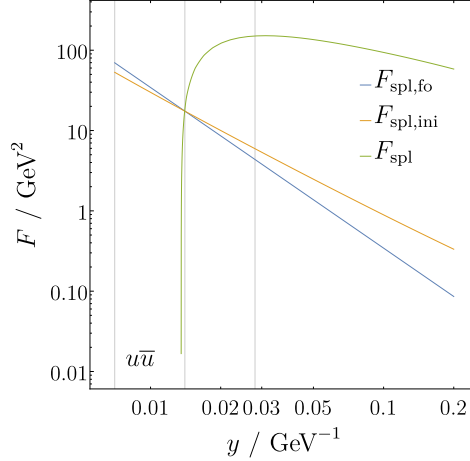
**Figure 18**. As figure 14a but for asymmetric kinematics $x_1 = 10^{-4}$, $x_2 = 0.31$ at $\mu = 80\,\text{GeV}$, which is needed for the point $Y = 4$ in figure 17a.

in (9.9) makes no sense. We note that in [39], different ansätze were made for polarised DPDs, and the effects of evolution on these distributions were investigated. It was found that, in many cases, evolution to higher scales leads to a suppression of the polarised DPD with respect to its unpolarised counterpart.

We initialise the polarised $F_{\text{spl}}$ at $\mu_y$ using the expression in (9.13) with the unpolarised $1 \to 2$ splitting functions replaced by their polarised counterparts (given in Appendix B of [39]). These distributions are then evolved to the scale $\mu$ using the appropriate polarised double DGLAP equations (the required polarised splitting functions are collected in Appendix A of [52]). The settings we used for the evolution code are the same as specified in section 9.2.1.

We consider the same scenario as in figure 12, i.e. the production of two systems with $Q_1 = Q_2 = 80\,\text{GeV}$ at $\sqrt{s} = 14\,\text{TeV}$, one with rapidity 0 and the other with rapidity $Y$, and now include polarised 1v1 contributions. In figure 19a, we reproduce figure 12a but include 1v1 luminosities for the polarised parton combinations $a_1 a_2 b_1 b_2 = \Delta u \Delta \bar{u} \Delta \bar{u} \Delta u + \Delta \bar{u} \Delta u \Delta u \Delta \bar{u}$ and $\delta u \delta \bar{u} \delta \bar{u} \delta u + \delta \bar{u} \delta u \delta u \delta \bar{u}$. For brevity we refer to these combinations as $\Delta u \Delta \bar{u}$ and $\delta u \delta \bar{u}$ in the following, recalling that $\Delta q$ and $\delta q$ indicate longitudinal and transverse quark polarisation, respectively. These polarised luminosities appear in the DPS cross section for double $Z$ production, see section 3.1 of [53] (combinations like $u\bar{u}\Delta\bar{u}\Delta u$ also appear but are not shown here). To avoid cluttering the plots, we omit the unpolarised 2v1 luminosity, reminding the reader that it is of the same magnitude as 2v2. In figure 19b we repeat the exercise for the pure gluon channel, including the 1v1 luminosities for the polarised combinations $a_1 a_2 b_1 b_2 = \Delta g \Delta g \Delta g \Delta g$ and $\delta g \delta g \delta g \delta g$, referred to as $\Delta g \Delta g$ and $\delta g \delta g$, where $\Delta g$ and $\delta g$ respectively denote longitudinal and linear gluon polarisation.

We see that the 1v1 luminosities for $\Delta u \Delta \bar{u}$ and $\delta u \delta \bar{u}$ essentially coincide with the one for $u\bar{u}$ at central rapidities, with $\nu$ variation bands of very similar size. These observations
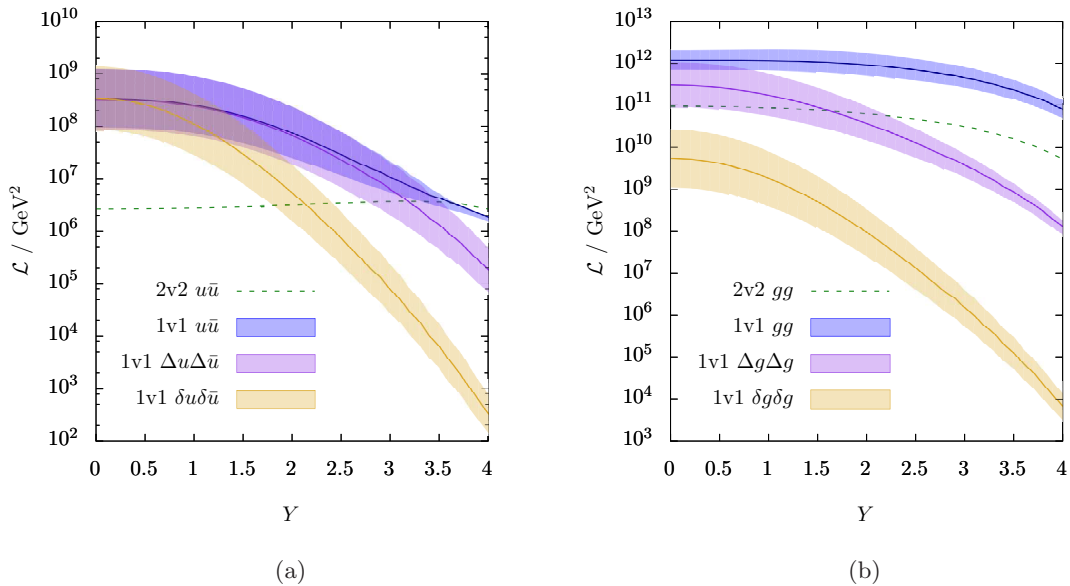
**Figure 19**. As in figure 12, but with two polarised 1v1 luminosity bands included in each case (and the unpolarised 2v1 band omitted).

can be explained by the fact that at $Y = 0$ we initialise all three DPDs using the same expression up to a minus sign, with $P_{g \to q\bar{q}} = |P_{g \to \Delta q \Delta \bar{q}}| = |P_{g \to \delta q \delta \bar{q}}|$ at $v = 1/2$, and that evolution has a rather weak effect on all three DPDs for central rapidities. This close agreement extends out to higher values of $Y$ for the $u\bar{u}$ and $\Delta u \Delta \bar{u}$ curves, owing to the fact that $P_{g \to q\bar{q}} = |P_{g \to \Delta q \Delta \bar{q}}|$ for general $v$. By contrast, the $\delta u \delta \bar{u}$ curve clearly falls below the other two for larger $Y$, because $P_{g \to q\bar{q}} > |P_{g \to \delta q \delta \bar{q}}|$ for $v \neq 1/2$.

In figure 19b, the $\delta g \delta g$ luminosity lies well below the one for $\Delta g \Delta g$, which in turn lies somewhat below the one for $gg$. This is mostly driven by the differences in initialisation expressions for the DPDs: at $v = 1/2$ (relevant at $Y = 0$) the relevant splitting function satisfy for instance $P_{g \to gg} : P_{g \to \Delta g \Delta g} : P_{g \to \delta g \delta g} = 9 : 7 : 1$. Ignoring evolution effects, one would then expect the $gg$ luminosity at $Y = 0$ to be roughly 1.5 times bigger than the one for $\Delta g \Delta g$, which in turn should be roughly 50 times bigger than the one for $\delta g \delta g$. This expectation is quite close to the actual luminosity ratios, which are $\sim 4$ and $\sim 60$, respectively. The remaining difference is due to evolution effects, which increase $gg$ rather considerably, increase $\Delta g \Delta g$ to a lesser extent, and hardly change $\delta g \delta g$.

Overall, we see that the polarised 1v1 luminosities can be of the same order of magnitude as the unpolarised ones, so that one must in general take into account all possible polarisation combinations in the DPS contribution, together with the SPS term and subtraction (where the subtraction will also contain both unpolarised and polarised contributions).

## 9.3 Production of two scalars

In this section, we study an explicit example to test our general argument that $\sigma_{1v1,pt}$ is of similar size as the corresponding perturbative order of the SPS cross section. The process we investigate is an artificial one, chosen for ease of computation. We consider the hypothetical production of two identical massive scalar bosons $\phi$, which couple to quarks with a Yukawa term $c\phi\bar{q}q$, where $c$ is a constant. For simplicity we take only one light quark flavour; including further light flavours just multiplies all following results (SPS and subtraction) by $n_f$. We will not compare the subtraction term to the full SPS contribution, but rather to the $gg$ initiated part of $\sigma_{SPS}$, which is the piece that contains the lowest order DPS-type $gg \to \phi\phi$ box diagram. This piece is gauge invariant and can hence be meaningfully considered by itself.

In keeping with the notation of the paper, let us denote the mass of each produced boson $\phi$ by $Q$. Let $Y$ be the rapidity of the diboson system in the $pp$ centre-of-mass frame, $\hat{s}$ its squared invariant mass, and $\beta = \sqrt{1 - 4Q^2/\hat{s}}$ the velocity of one boson in the diboson centre-of-mass frame. The $gg \to \phi\phi$ part of the SPS cross section can be written as

$$
\frac{d\sigma_{SPS}}{dY\, d\beta} = x\bar{x}\, g(x)\, g(\bar{x}) \frac{2\beta}{1 - \beta^2}\, \hat{\sigma}_{gg \to \phi\phi}
$$

$$
= x\bar{x}\, g(x)\, g(\bar{x}) \frac{1}{128\pi} \frac{1 - \beta^2}{Q^4} \int_0^{q_{max}^2} \frac{dq^2}{\sqrt{1 - q^2/q_{max}^2}}\, \left|\mathcal{A}_{gg \to \phi\phi}\right|^2, \tag{9.16}
$$

where $q_{max} = \beta Q/\sqrt{1 - \beta^2}$ and the $gg \to \phi\phi$ matrix element squared, $\left|\mathcal{A}_{gg \to \phi\phi}\right|^2$, is averaged over the spin and colour of the incoming gluons. The momentum fractions $x$ and $\bar{x}$ in the gluon distributions are obtained from the constraints $\hat{s} = x\bar{x}s$ and $Y = \frac{1}{2}\log(x/\bar{x})$.

The matrix element squared for $gg \to \phi\phi$ can be obtained from the $gg \to HH$ matrix element squared given in [54] by making the replacement

$$
\frac{\pi\alpha_W m_q^2}{M_W^2} \to c^2\,. \tag{9.17}
$$

In particular, $\left|\mathcal{A}_{gg \to \phi\phi}\right|^2$ in (9.16) is related to the terms "gauge1" and "gauge2" in [54] according to

$$
\left|\mathcal{A}_{gg \to \phi\phi}\right|^2 = \frac{c^4\alpha_s^2}{2^{10}\pi^2 m_q^4} \left(|\text{gauge1}|^2 + |\text{gauge2}|^2\right)\,. \tag{9.18}
$$

The right hand side of (9.18) is given in [54] for general quark mass $m_q$. We evaluate this expression numerically for very small $m_q$, using analytic $m_q \to 0$ approximations where this is necessary to avoid numerical instabilities.

Now we turn to the subtraction term. The two elementary $q\bar{q} \to \phi$ cross sections in this term contain the kinematic constraint $2\pi\delta(x_i\bar{x}_i s - Q^2)$, which fixes the momentum fractions $x_i$ and $\bar{x}_i$ entering the two hard subprocesses at given $x, \bar{x}$ and $\beta$ up to a two-fold

ambiguity. The DPS subtraction term is given by

$$\frac{d\sigma_{1\text{v1,pt}}}{dY\,d\beta} = \frac{\pi^2}{16}\frac{1-\beta^2}{Q^8}\,x^2\bar{x}^2 \sum_{a_1a_2b_1b_2} \left|\mathcal{A}_{a_1b_1\to\phi}\right|^2 \left|\mathcal{A}_{a_2b_2\to\phi}\right|^2 \int d^2\boldsymbol{y}\,\Phi^2(y\nu) \sum_{R=1,8}$$

$$\times \left[{}^R F_{a_1a_2,\text{spl,pt}}\left(\tfrac{1}{2}x(1+\beta),\tfrac{1}{2}x(1-\beta),\boldsymbol{y}\right) {}^R F_{b_1b_2,\text{spl,pt}}\left(\tfrac{1}{2}\bar{x}(1-\beta),\tfrac{1}{2}\bar{x}(1+\beta),\boldsymbol{y}\right)\right.$$

$$\left. + \{\beta \to -\beta\}\right] + \{F \to I\}, \tag{9.19}$$

where we sum over all possible colour, spin, and quark number interference/correlation possibilities. The index $R$ on $F$ denotes the colour channel ($R = 1$ for colour singlet, $R = 8$ for colour octet), and in the sum over $a_1a_2b_1b_2$ we sum over both unpolarised (e.g. $q\bar{q}\bar{q}q$) and polarised (e.g. $\Delta q\Delta\bar{q}\Delta\bar{q}\Delta q$) combinations. By $\{F \to I\}$ we denote the same expression, but with the quark number diagonal distributions replaced by the quark number interference ones (see section 2 of [5]).

The splitting kernels for the relevant spin combinations are [5]

$${}^1P_{g\to q\bar{q}}(v) = -\,{}^1P_{g\to\Delta q\Delta\bar{q}}(v) = \frac{1}{2}\left[v^2 + (1-v)^2\right], \quad {}^1P_{g\to\delta q\delta\bar{q}}(v) = -\,\delta^{jj'}v(1-v) \tag{9.20}$$

and ${}^8P = -\,{}^1P/\sqrt{N_c^2-1}$, where $j,j'$ are indices for transverse quark polarisation. The term with interference DPDs $I$ gives the same contribution as that with $F$, since the corresponding diagrams are simply related by reversing the direction of fermion flow in one of the two quark loops, and this does not change the expression for the diagram. The squared subprocess amplitudes, including an average over colour in the initial state, read

$$\left|\mathcal{A}_{q\bar{q}\to\phi}\right|^2 = \left|\mathcal{A}_{\Delta q\Delta\bar{q}\to\phi}\right|^2 = \frac{c^2Q^2}{2N_c}, \qquad \left|\mathcal{A}_{\delta q\delta\bar{q}\to\phi}\right|^2 = \delta^{jj'}\frac{c^2Q^2}{2N_c}, \tag{9.21}$$

where $j,j'$ are the indices for transverse quark polarisation and where we have used the spin projection operators given in equation (2.90) of [5].

Inserting (9.20) and (9.21) into (9.19), we obtain

$$\frac{d\sigma_{1\text{v1,pt}}}{dY\,d\beta} = x\bar{x}\,g(x)\,g(\bar{x})\,\frac{1}{128\pi}\frac{1-\beta^2}{Q^4}\,(1+\beta^4)\frac{c^4\alpha_s^2}{N_c^2-1}\int_0^\infty \frac{dy^2}{y^4}\,\Phi^2(y\nu). \tag{9.22}$$

Note that both (9.16) and (9.22) contain a product of gluon PDFs evaluated at the same $x$ values, $g(x)\,g(\bar{x})$. For the comparison, we can divide the common PDF factor out of the two expressions, in order to avoid having to use an explicit parameterisation. We also divide out various factors appearing in both expressions and compare the quantity

$$\Sigma(\beta) = \frac{d\sigma}{dY\,d\beta}\frac{Q^2}{x\bar{x}\,g(x)\,g(\bar{x})}\frac{128\pi(N_c^2-1)}{c^4\alpha_s^2}. \tag{9.23}$$

It is straightforward to show that $\Sigma$ is a function of the variable $\beta$ only. We plot $\Sigma$ for SPS and the subtraction term in figure 20, where the curve for $\Sigma_{1\text{v1,pt}}$ corresponds to the central choice $\nu = Q$ of the cutoff scale.
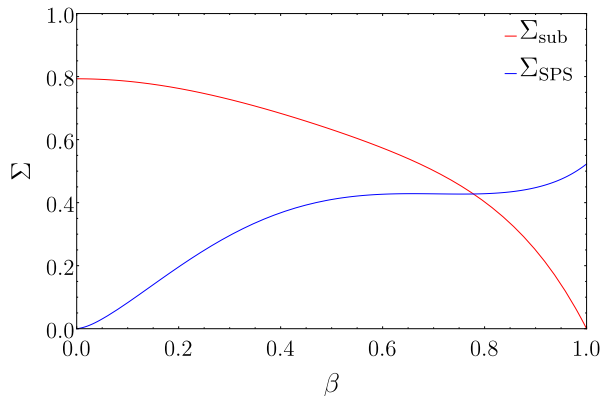
**Figure 20**. The scaled cross section $\Sigma$ defined in (9.23), plotted for the $gg \to \phi\phi$ contribution to SPS ($\Sigma_{\text{SPS}}$), and the subtraction term ($\Sigma_{\text{1v1,pt}}$).

We see that $\Sigma_{\text{1v1,pt}}(\beta)$ is indeed generally of the same order of magnitude as $\Sigma_{\text{SPS}}(\beta)$. The agreement is perhaps surprisingly good, given that $\Sigma_{\text{1v1,pt}}(\beta)$ involves integrating a low $q$ approximation to the matrix element squared outside the region where the approximation is valid. The agreement gets worse towards the endpoints $\beta \to 0$ and $\beta \to 1$, which correspond to the high energy limit and the threshold limit, respectively. It is to be expected that the agreement is especially bad at these points, since in the subtraction term we effectively assume that the integration over the squared transverse momentum $q^2$ of the scalar particles goes from zero to values of order $Q^2$. For both $\beta \to 0$ and $\beta \to 1$ this assumption becomes a poor one: near threshold the phase space in $q^2$ shrinks to zero, whilst in the high energy limit, $q^2$ can go up to size $\hat{s} \gg Q^2$.

## 10  Summary

Consistently incorporating the perturbative splitting of one parton into two is a highly nontrivial problem for the theoretical description of double parton scattering. DPS graphs in which such splittings occur in both protons (1v1 graphs) overlap with loop corrections to single parton scattering. Another type of graph, typically referred to as 2v1, in which one parton pair arises from a perturbative splitting, and the other pair is an "intrinsic" one already existing at the nonperturbative level, overlaps with twist-four contributions to the cross section. Finally there is an overlap between DPS contributions where a splitting occurs in both protons only in the amplitude or its conjugate, and SPS/DPS interference graphs.

We have presented a scheme to compute DPS and to consistently merge its contribution to the cross section with SPS and the other terms just mentioned. The scheme works in a similar manner for collinear and for TMD factorisation. Ultraviolet divergences that arise from perturbative splitting in a naive treatment of DPS are regulated by a cutoff function $\Phi(y\nu)$ in transverse position space. This avoids modification of the position space DPDs, which are defined via operator matrix elements in close analogy to single-parton

distributions. In collinear factorisation, these DPDs hence evolve according to a homogeneous DGLAP equation, whilst their TMD counterparts satisfy a generalisation of the renormalisation group equations for single-parton TMDs. No modification of hard scattering cross sections computed for standard collinear or TMD factorisation is necessary in our scheme. Collins-Soper type equations describe the rapidity evolution of transverse-momentum dependent DPDs and of collinear DPDs in colour non-singlet channels [21].

The problem of double counting between DPS and other contributions — notably between DPS and SPS — is solved by subtraction terms as specified in (4.16) and (4.22), which are obtained in a simple way from $\sigma_{\text{DPS}}$ by replacing the DPDs with their appropriate short-distance limits. This paves the way for using the scheme at higher orders in $\alpha_s$, with calculations being considerably simpler for the subtraction terms compared with the full hard scattering process at the corresponding order. With a suitable choice of starting conditions and scales, specified in section 6.1, the DPS part of the cross section correctly resums DGLAP logarithms that are not included in the fixed order twist-four contributions.

Our scheme is naturally formulated with position-space DPDs $F(x_i, \boldsymbol{y})$, but it is possible to relate the Fourier transform of $F(x_i, \boldsymbol{y}) \, \Phi(y\nu)$ to DPDs $F(x_i, \boldsymbol{\Delta})$ that are renormalised in transverse momentum space and satisfy an inhomogeneous DGLAP equation rather than a homogeneous one. This relation has the form of a perturbative matching equation, see (7.15), and is somewhat similar to the matching between PDFs defined in different schemes such as the $\overline{\text{MS}}$ and the DIS scheme. The momentum space representation also allows us to show that for the 2v1 contribution to DPS our scheme is equivalent to the ones in [12, 13] and in [14, 15] to leading logarithmic accuracy.

For collinear DPDs, one can make a model ansatz consisting of two terms which in the limit $y \ll 1/\Lambda$ respectively give the perturbative splitting and the intrinsic part of the distribution. With such an ansatz, the DPS cross section naturally splits into 1v1, 2v1 and 2v2 terms, where 2v2 refers to contributions in which the parton pairs from both protons are intrinsic. A crucial question is how large DPS is compared with SPS at the perturbative order where graphs contribute to both mechanisms. This is especially acute in collinear factorisation, where DPS is power suppressed with respect to SPS. Note that only in very few channels (notably pair production of electroweak gauge bosons) SPS calculations are available at the required order. We argue in section 6.2 that in our scheme the variation with $\nu$ of the 1v1 term in DPS provides an order-of magnitude estimate for the SPS contribution $\sigma_{\text{SPS}}$ (at the appropriate perturbative order), as it involves the same PDFs, overall coupling constants and kinematic region (small $\boldsymbol{y}$, corresponding to large transverse momenta and virtualities of internal lines). An alternative estimate is provided by the double counting subtraction term $\sigma_{\text{1v1,pt}}$, which by construction is dominated by small $\boldsymbol{y}$. For the hypothetical process of scalar boson pair production from two gluons, we have shown that the latter estimate works well within about a factor of two, provided that one stays away from the extreme kinematic regions where the velocity $\beta$ of a boson in the boson pair rest frame is close to 0 or 1.

We constructed explicit (collinear) DPD input forms using the model ansatz just described, restricting ourselves to three quark flavours for simplicity and ease of implementation. These inputs were then numerically evolved to other scales using a code that

implements the homogeneous double DGLAP equation. We used the resultant DPDs to compute so-called DPS luminosities (DPS cross sections omitting the process-dependent hard parts) and plotted these under various conditions. We observed that generically, the 1v1 contributions to the luminosity (both unpolarised and polarised) are comparable to or larger than the 2v1 and 2v2 contributions, with a large dependence on the cutoff $\nu$. This demonstrates that, when including DPS in a cross section calculation, one must in general include $\sigma_{\mathrm{SPS}}$ up to the order that contains DPS-like double box graphs, together with the associated subtraction term (with unpolarised and polarised partons). Otherwise, one would have an uncertainty on the overall cross section that is as large as, or larger than the DPS term itself. We also confirmed that the $\nu$ variation of the 1v1 DPS contribution is indeed comparable to the central value of the associated double counting subtraction term, so that either of them may be used as an estimate for the SPS contribution.

We identified several processes and scenarios where the $\nu$ variation of the 1v1 DPS luminosity is considerably smaller than the central value. As we argued above, one may then justifiably neglect the order of $\sigma_{\mathrm{SPS}}$ containing the first DPS-like double box, as well as the associated subtraction term, compared with $\sigma_{\mathrm{DPS}}$. One scenario of this kind is when the flavour indices in both DPDs are $u\bar{d}$ (the luminosity with this parton flavour combination appears in $W^+W^+$ production). The suppression of the DPS-like double box in $\sigma_{\mathrm{SPS}}$ is in this case related to the fact that perturbative splitting in the $u\bar{d}$ DPD starts at order $\alpha_s^2$ rather than at order $\alpha_s$. Further scenarios in which the $\nu$ variation in the 1v1 contribution is reduced are when $\sqrt{s}$ becomes very large compared to the hard scales $Q_i$, or when the rapidity separation between the produced hard systems is large. Both of these scenarios involve small $x$ values in the DPDs — in the first, both $x$ values in each DPD are small, whilst in the second, one $x$ value in each DPD becomes much smaller than the other. Such processes and kinematic regions are the most promising ones to make useful calculations and measurements for DPS. In fact, several measurements investigating DPS have already been made in kinematics with $Q_i \ll \sqrt{s}$ or with large rapidity differences. It will be interesting in future work to make more complete and comprehensive predictions for such processes and kinematic regions in our framework, including for instance the full flavour dependence and contributions from all partonic channels for a considered final state.

## A  Fourier integrals

In this appendix we collect a number of results for Fourier transforms in 2 or $2 - 2\epsilon$ dimensions.

**Fourier transform of a fractional power.**  To compute the Fourier transform of $(\boldsymbol{k}^2)^{-\lambda}$ in $2 - 2\epsilon$ dimensions we write $(\boldsymbol{k}^2)^{-\lambda} = \Gamma^{-1}(\lambda) \int_0^\infty d\alpha\, \alpha^{\lambda-1} e^{-\alpha \boldsymbol{k}^2}$. One can then perform the Gaussian integration over $\boldsymbol{k}$, and subsequently the integral over $\alpha$. The result is

$$\int d^{2-2\epsilon}\boldsymbol{k}\; e^{-i\boldsymbol{y}\boldsymbol{k}}\,(\boldsymbol{k}^2)^{-\lambda} = \frac{\Gamma(1-\epsilon-\lambda)}{\Gamma(\lambda)}\,\pi^{1-\epsilon}\left(\frac{\boldsymbol{y}^2}{4}\right)^{-1+\epsilon+\lambda} \tag{A.1}$$

As a corollary we obtain

$$\int d^{2-2\epsilon}\boldsymbol{k}\, e^{i\boldsymbol{yk}}\frac{\boldsymbol{k}^j}{k^2} = \frac{i}{2}\Gamma(1-\epsilon)\,(4\pi)^{1-\epsilon}\frac{\boldsymbol{y}^j}{y^{2-2\epsilon}} \tag{A.2}$$

using $\boldsymbol{k}^j\, e^{i\boldsymbol{yk}} = -i\partial/(\partial \boldsymbol{y}^j)\, e^{i\boldsymbol{yk}}$.

**Integrals involving $\Phi$.** We now compute the integrals in (4.19), which are also needed for (7.6) and (7.7), If $\Phi$ is the step function $\Theta(u-b_0)$ then

$$2\int_{b_0}^{\infty}\frac{du}{u}\, J_0(ur) = \log\frac{1}{r^2} + \frac{b_0^2\, r^2}{4}\,{}_2F_3\left(1,1;2,2,2;-\frac{b_0^2\, r^2}{4}\right), \tag{A.3}$$

where $b_0$ is defined in (4.3) and ${}_2F_3(1,1;2,2,2;z)$ is a generalised hypergeometric function, which can be expanded as $1+\mathcal{O}(z)$ for small arguments. For $\Phi(u) = 1 - e^{-u^2/4}$ we can proceed as follows:

$$2\int_0^{\infty}\frac{du}{u}\, J_0(ur)\left[1 - e^{-u^2/4}\right] = 2\int_0^{\infty}\frac{du}{u}\, J_0(ur)\, u^2\int_0^{1/4}d\tau\, e^{-\tau u^2} = \int_0^{1/4}\frac{d\tau}{\tau}\,\exp\left[-\frac{r^2}{4\tau}\right]$$

$$= E_1(r^2), \tag{A.4}$$

where $E_1$ is the exponential integral. For small arguments one has $E_1(r^2) = \log(1/r^2) - \gamma + \mathcal{O}(r^2)$. The corresponding integrals with $J_2$ instead of $J_0$ are given by

$$2\int_{b_0}^{\infty}\frac{du}{u}\, J_2(ur) = \frac{2J_1(b_0 r)}{b_0 r} \tag{A.5}$$

and

$$2\int_0^{\infty}\frac{du}{u}\, J_2(ur)\left[1 - e^{-u^2/4}\right] = \int_0^{\infty}du\,\frac{J_1(ur)}{ur}\, u e^{-u^2/4} = \frac{1 - e^{-r^2}}{r^2}, \tag{A.6}$$

where in the second case we have used integration by parts. Both (A.5) and (A.6) behave like $1 + \mathcal{O}(r^2)$ for small $r$.

**Connection between the Fourier-Bessel transform and a cutoff.** Consider the integral (8.7) with one term of the series (8.9). We will show that for $\Delta \ll \nu$

$$\int_{b_0/\nu}^{\infty}\frac{dy}{y}\, J_0(y\Delta)\log^n\left(\frac{y\nu}{b_0}\right) = \int_{b_0/\nu}^{b_0/\Delta}\frac{dy}{y}\log^n\left(\frac{y\nu}{b_0}\right)$$

$$+ \mathcal{O}\left(\frac{\Delta^2}{\nu^2}\right) + \mathcal{O}\left(\log^{n-2}\frac{\nu}{\Delta}\right). \tag{A.7}$$

The following argument is similar to the derivation given in appendix B of [55]. The integral on the r.h.s. is readily performed and gives

$$\int_1^{\nu/\Delta}dv\,\frac{1}{v}\log^n v = \frac{1}{1+n}\log^{n+1}\left(\frac{\nu}{\Delta}\right). \tag{A.8}$$

The expression on the l.h.s. of (A.7) can be rewritten as

$$\int_{b_0\Delta/\nu}^{\infty} du\, J_0(u)\, \frac{1}{u} \log^n\left(\frac{u\nu}{b_0\Delta}\right) = \frac{1}{n+1} \int_{b_0\Delta/\nu}^{\infty} du\, J_1(u) \log^{n+1}\left(\frac{u\nu}{b_0\Delta}\right) + \mathcal{O}\left(\frac{\Delta^2}{\nu^2}\right) \qquad (A.9)$$

using integration by parts. Since $J_1(u) = \mathcal{O}(u)$ at small $u$, the integral on the r.h.s. can be extended down to $u = 0$ with an accuracy of $\Delta^2/\nu^2$. Rewriting the logarithm using the binomial series, we obtain

$$\frac{1}{n+1} \sum_{k=0}^{n+1} \binom{n+1}{k} \log^{n+1-k}\left(\frac{\nu}{\Delta}\right) \int_0^{\infty} du\, J_1(u)\, \log^k\left(\frac{u}{b_0}\right)$$

$$= \frac{1}{n+1} \left[ \log^{n+1}\left(\frac{\nu}{\Delta}\right) + \sum_{k=3}^{n+1} \binom{n+1}{k} d_k\, \log^{n+1-k}\left(\frac{\nu}{\Delta}\right) \right]. \qquad (A.10)$$

In the last step we performed the integrals for $k = 0, 1, 2$ explicitly (the ones for $k = 1, 2$ are zero) and replaced the remaining ones by coefficients $d_k$ whose values are not important here. Comparison with (A.8) gives the desired result (A.7).

## Acknowledgments

## References

[1] R. Astalos et al., *Proceedings of the Sixth International Workshop on Multiple Partonic Interactions at the Large Hadron Collider*, 1506.05829.

[2] H. Jung, D. Treleani, M. Strikman and N. van Buuren, eds., *Proceedings of the 7th International Workshop on Multiple Partonic Interactions at the LHC (MPI@LHC 2015)*, DESY-PROC 2016-01, 2016.

[3] N. Paver and D. Treleani, *Multi - Quark Scattering and Large $p_T$ Jet Production in Hadronic Collisions*, *Nuovo Cim.* **A70** (1982) 215.

[4] M. Diehl and A. Schäfer, *Theoretical considerations on multiparton interactions in QCD*, *Phys. Lett.* **B698** (2011) 389–402, [1102.3081].

[5] M. Diehl, D. Ostermeier and A. Schäfer, *Elements of a theory for multiparton interactions in QCD*, *JHEP* **03** (2012) 089, [1111.0910].

[6] R. Kirschner, *Generalized Lipatov-Altarelli-Parisi Equations and Jet Calculus Rules*, *Phys. Lett.* **B84** (1979) 266–270.

[7] V. P. Shelest, A. M. Snigirev and G. M. Zinovev, *The Multiparton Distribution Equations in QCD*, *Phys. Lett.* **B113** (1982) 325.

[8] A. M. Snigirev, *Double parton distributions in the leading logarithm approximation of perturbative QCD*, *Phys. Rev.* **D68** (2003) 114012, [`hep-ph/0304172`].

[9] 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`].

[10] F. A. Ceccopieri, *An update on the evolution of double parton distributions*, *Phys. Lett.* **B697** (2011) 482–487, [`1011.6586`].

[11] M. Cacciari, G. P. Salam and S. Sapeta, *On the characterisation of the underlying event*, *JHEP* **04** (2010) 065, [`0912.4926`].

[12] B. Blok, Yu. Dokshitser, L. Frankfurt and M. Strikman, *pQCD physics of multiparton interactions*, *Eur. Phys. J.* **C72** (2012) 1963, [`1106.5533`].

[13] B. Blok, Yu. Dokshitzer, L. Frankfurt and M. Strikman, *Perturbative QCD correlations in multi-parton collisions*, *Eur. Phys. J.* **C74** (2014) 2926, [`1306.3763`].

[14] M. G. Ryskin and A. M. Snigirev, *A Fresh look at double parton scattering*, *Phys. Rev.* **D83** (2011) 114047, [`1103.3495`].

[15] M. G. Ryskin and A. M. Snigirev, *Double parton scattering in double logarithm approximation of perturbative QCD*, *Phys. Rev.* **D86** (2012) 014018, [`1203.2330`].

[16] A. V. Manohar and W. J. Waalewijn, *What is Double Parton Scattering?*, *Phys. Lett.* **B713** (2012) 196–201, [`1202.5034`].

[17] J. R. Gaunt, *Single Perturbative Splitting Diagrams in Double Parton Scattering*, *JHEP* **01** (2013) 042, [`1207.0480`].

[18] J. Collins, *Foundations of perturbative QCD*. Cambridge University Press, 2013.

[19] T. C. Rogers, *An Overview of Transverse Momentum Dependent Factorization and Evolution*, *Eur. Phys. J.* **A52** (2016) 153, [`1509.04766`].

[20] M. Diehl, J. R. Gaunt, D. Ostermeier, P. Plößl and A. Schäfer, *Cancellation of Glauber gluon exchange in the double Drell-Yan process*, *JHEP* **01** (2016) 076, [`1510.08696`].

[21] M. Buffing, M. Diehl and T. Kasemets, *Transverse momentum in double parton scattering: factorisation, evolution and matching, preprint in preparation, DESY-2017-014, NIKHEF-2016-028* (2017) .

[22] A. V. Manohar and W. J. Waalewijn, *A QCD Analysis of Double Parton Scattering: Color Correlations, Interference Effects and Evolution*, *Phys. Rev.* **D85** (2012) 114009, [`1202.3794`].

[23] H.-M. Chang, A. V. Manohar and W. J. Waalewijn, *Double Parton Correlations in the Bag Model*, *Phys. Rev.* **D87** (2013) 034009, [`1211.3132`].

[24] M. Rinaldi, S. Scopetta and V. Vento, *Double parton correlations in constituent quark models*, *Phys. Rev.* **D87** (2013) 114021, [`1302.6462`].

[25] M. Rinaldi, S. Scopetta, M. Traini and V. Vento, *Double parton correlations and constituent quark models: a Light Front approach to the valence sector*, *JHEP* **12** (2014) 028, [`1409.1500`].

[26] 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].

[27] M. Rinaldi and F. A. Ceccopieri, *Relativistic effects in model calculations of double parton distribution function*, *Phys. Rev.* **D95** (2017) 034040, [1611.04793].

[28] W. Broniowski and E. Ruiz Arriola, *Valence double parton distributions of the nucleon in a simple model*, *Few Body Syst.* **55** (2014) 381–387, [1310.8419].

[29] W. Broniowski, E. Ruiz Arriola and K. Golec-Biernat, *Generalized Valon Model for Double Parton Distributions*, *Few Body Syst.* **57** (2016) 405–410, [1602.00254].

[30] T. Kasemets and A. Mukherjee, *Quark-gluon double parton distributions in the light-front dressed quark model*, *Phys. Rev.* **D94** (2016) 074029, [1606.05686].

[31] V. M. Braun, A. N. Manashov and J. Rohrwild, *Baryon Operators of Higher Twist in QCD and Nucleon Distribution Amplitudes*, *Nucl. Phys.* **B807** (2009) 89–137, [0806.2531].

[32] V. M. Braun, A. N. Manashov and J. Rohrwild, *Renormalization of Twist-Four Operators in QCD*, *Nucl. Phys.* **B826** (2010) 235–293, [0908.1684].

[33] Y. Ji and A. V. Belitsky, *Renormalization of twist-four operators in light-cone gauge*, *Nucl. Phys.* **B894** (2015) 161–222, [1405.2828].

[34] A. P. Bukhvostov, G. V. Frolov, L. N. Lipatov and E. A. Kuraev, *Evolution Equations for Quasi-Partonic Operators*, *Nucl. Phys.* **B258** (1985) 601–646.

[35] R. L. Jaffe, *Spin, twist and hadron structure in deep inelastic processes*, in *The spin structure of the nucleon. Proceedings, International School of Nucleon Structure, 1st Course, Erice, Italy, August 3-10, 1995*, pp. 42–129, 1996. hep-ph/9602236.

[36] P. J. Mulders and J. Rodrigues, *Transverse momentum dependence in gluon distribution and fragmentation functions*, *Phys. Rev.* **D63** (2001) 094021, [hep-ph/0009343].

[37] R. L. Jaffe and X.-D. Ji, *Chiral odd parton distributions and Drell-Yan processes*, *Nucl. Phys.* **B375** (1992) 527–560.

[38] J. C. Collins, D. E. Soper and G. F. Sterman, *Transverse Momentum Distribution in Drell-Yan Pair and W and Z Boson Production*, *Nucl. Phys.* **B250** (1985) 199–224.

[39] M. Diehl, T. Kasemets and S. Keane, *Correlations in double parton distributions: effects of evolution*, *JHEP* **05** (2014) 118, [1401.1233].

[40] F. Caola, M. Dowling, K. Melnikov, R. Röntsch and L. Tancredi, *QCD corrections to vector boson pair production in gluon fusion including interference effects with off-shell Higgs at the LHC*, *JHEP* **07** (2016) 087, [1605.04610].

[41] A. V. Belitsky and A. V. Radyushkin, *Unraveling hadron structure with generalized parton distributions*, *Phys. Rept.* **418** (2005) 1–387, [hep-ph/0504030].

[42] M. Diehl and J. R. Gaunt, *Double parton scattering in the ultraviolet: addressing the double counting problem*, in *Proceedings, 7th International Workshop on Multiple Partonic Interactions at the LHC (MPI@LHC 2015): Miramare, Trieste, Italy, November 23-27, 2015*, pp. 121–125, 2016. 1603.05468.

[43] M. Diehl and J. R. Gaunt, *Double parton scattering in the ultraviolet: addressing the double counting problem*, in *Proceedings, QCD Evolution Workshop (QCD 2016): Amsterdam, Netherlands, May 30-June 3, 2016*, 2016. 1611.01316.

[44] Z. Ligeti, I. W. Stewart and F. J. Tackmann, *Treating the b quark distribution function with reliable uncertainties*, *Phys. Rev.* **D78** (2008) 114014, [`0807.1926`].

[45] R. Abbate, M. Fickinger, A. H. Hoang, V. Mateu and I. W. Stewart, *Thrust at $N^3LL$ with Power Corrections and a Precision Global Fit for alphas(mZ)*, *Phys. Rev.* **D83** (2011) 074021, [`1006.3080`].

[46] I. W. Stewart, F. J. Tackmann, J. R. Walsh and S. Zuberi, *Jet $p_T$ resummation in Higgs production at $NNLL' + NNLO$*, *Phys. Rev.* **D89** (2014) 054001, [`1307.1808`].

[47] M. Mekhfi and X. Artru, *Sudakov Suppression of Color Correlations in Multiparton Scattering*, *Phys. Rev.* **D37** (1988) 2618–2622.

[48] A. D. Martin, W. J. Stirling, R. S. Thorne and G. Watt, *Parton distributions for the LHC*, *Eur. Phys. J.* **C63** (2009) 189–285, [`0901.0002`].

[49] J. C. Collins and D. E. Soper, *Back-To-Back Jets: Fourier Transform from B to K-Transverse*, *Nucl. Phys.* **B197** (1982) 446–476.

[50] J. Collins and T. Rogers, *Understanding the large-distance behavior of transverse-momentum-dependent parton densities and the Collins-Soper evolution kernel*, *Phys. Rev.* **D91** (2015) 074020, [`1412.3820`].

[51] 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 $c\bar{c}c\bar{c}$*, *Phys. Rev.* **D90** (2014) 054017, [`1407.5821`].

[52] M. Diehl and T. Kasemets, *Positivity bounds on double parton distributions*, *JHEP* **05** (2013) 150, [`1303.0842`].

[53] T. Kasemets and M. Diehl, *Angular correlations in the double Drell-Yan process*, *JHEP* **01** (2013) 121, [`1210.5434`].

[54] E. W. N. Glover and J. J. van der Bij, *Higgs Boson Pair Production Via Gluon Fusion*, *Nucl. Phys.* **B309** (1988) 282–294.

[55] A. Bacchetta, D. Boer, M. Diehl and P. J. Mulders, *Matches and mismatches in the descriptions of semi-inclusive processes at low and high transverse momentum*, *JHEP* **08** (2008) 023, [`0803.0227`].

[56] D. Binosi and L. Theussl, *JaxoDraw: A Graphical user interface for drawing Feynman diagrams*, *Comput. Phys. Commun.* **161** (2004) 76–86, [`hep-ph/0309015`].

[57] D. Binosi, J. Collins, C. Kaufhold and L. Theussl, *JaxoDraw: A Graphical user interface for drawing Feynman diagrams. Version 2.0 release notes*, *Comput. Phys. Commun.* **180** (2009) 1709–1715, [`0811.4113`].